doi:10.1016/j.jmb.2006.08.024 J. Mol. Biol. (2006) 363, 531–544

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

The Paradoxical Behavior of a Highly Structured Misfolded Intermediate in RNA Folding

Rick Russell[1], Rhiju Das[2], Hyejean Suh[2], Kevin J. Travers[2] Alain Laederach[3], Mark A. Engelhardt[2] and Daniel Herschlag[2]*

[1]*Department of Chemistry and Biochemistry, Institute for Cellular and Molecular Biology University of Texas at Austin Austin, TX 78712, USA*

[2]*Department of Biochemistry Stanford University, Stanford CA 94305-5307, USA*

[3]*Department of Genetics Stanford University, Stanford CA 94305-5307, USA*

**Corresponding author*

Like many structured RNAs, the *Tetrahymena* group I ribozyme is prone to misfolding. Here we probe a long-lived misfolded species, referred to as M, and uncover paradoxical aspects of its structure and folding. Previous work indicated that a non-native local secondary structure, termed alt P3, led to formation of M during folding *in vitro*. Surprisingly, hydroxyl radical footprinting, fluorescence measurements with site-specifically incorporated 2-aminopurine, and functional assays indicate that the native P3, not alt P3, is present in the M state. The paradoxical behavior of alt P3 presumably arises because alt P3 biases folding toward M, but, after commitment to this folding pathway and before formation of M, alt P3 is replaced by P3. Further, structural and functional probes demonstrate that the misfolded ribozyme contains extensive native structure, with only local differences between the two states, and the misfolded structure even possesses partial catalytic activity. Despite the similarity of these structures, re-folding of M to the native state is very slow and is strongly accelerated by urea, Na^+, and increased temperature and strongly impeded by Mg^{2+} and the presence of native peripheral contacts. The paradoxical observations of extensive native structure within the misfolded species but slow conversion of this species to the native state are readily reconciled by a model in which the misfolded state is a topological isomer of the native state, and computational results support the feasibility of this model. We speculate that the complex topology of RNA secondary structures and the inherent rigidity of RNA helices render kinetic traps due to topological isomers considerably more common for RNA than for proteins.

Keywords: group I ribozyme; RNA topology; 2-aminopurine fluorescence; hydroxyl radical footprinting; DMS footprinting

Introduction

Beginning with studies of tRNA in the 1960s, it has become clear that RNA folding is rife with kinetic traps.[1–3] In the process of adopting specific three-dimensional structures, RNAs have a strong tendency to form intermediates that include non-native structure and must partially unfold to allow continued folding to the native state. Although these non-native intermediates typically are not the most thermodynamically favorable states, they can lie deep in local energetic minima, giving slow unfolding transitions that can be rate-limiting for the overall folding process. Since the early days of RNA folding studies, kinetically trapped folding intermediates have been identified for nearly every RNA whose folding has been studied.[4–13]

Because the presence of kinetic traps can obscure other folding processes, such as conformational searches and secondary and tertiary structure formation, there has been recent emphasis on identifying and studying RNAs whose folding is not rate-limited by re-folding from kinetic traps.[14–16] These efforts have yielded two RNAs, a domain of RNase P

Present address: R. Das, Department of Biochemistry, University of Washington, Seattle, WA 98195, USA.

Abbreviations used: DMS, dimethyl sulfate; M, the long-lived misfolded conformation of the *Tetrahymena* ribozyme; N, the native conformation of the *Tetrahymena* ribozyme; S, oligonucleotide substrate ($CCCUCUA_5$); 2AP, 2-aminopurine.

E-mail address of the corresponding author: herschla@stanford.edu

and a group II intron, that appear to avoid kinetic traps. The study of such RNAs is likely to continue to provide insights into the manner in which RNA forms complex three-dimensional structures.[16,17]

Nevertheless, the challenge remains to obtain a deep understanding of the RNA folding landscape and the kinetic traps that are encountered. Interest in the properties of kinetic traps arises from both physical and biological perspectives. Because kinetic traps are so prevalent in RNA folding, an understanding of the physical principles that govern RNA folding processes must include understanding how kinetic traps form during folding, what molecular features are required for their maintenance, and how they ultimately re-fold to their native structures. Indeed, even the "trap-free" RNAs noted above may transiently form misfolded species that are not detected because emergence from them is not rate-limiting or because the misfolded species does not possess the properties that are diagnostic for the presence of a kinetic trap.[18] From a biological perspective, it appears that Nature has taken advantage of the ability of RNA to adopt alternative structures for control and regulation of processes like transla-

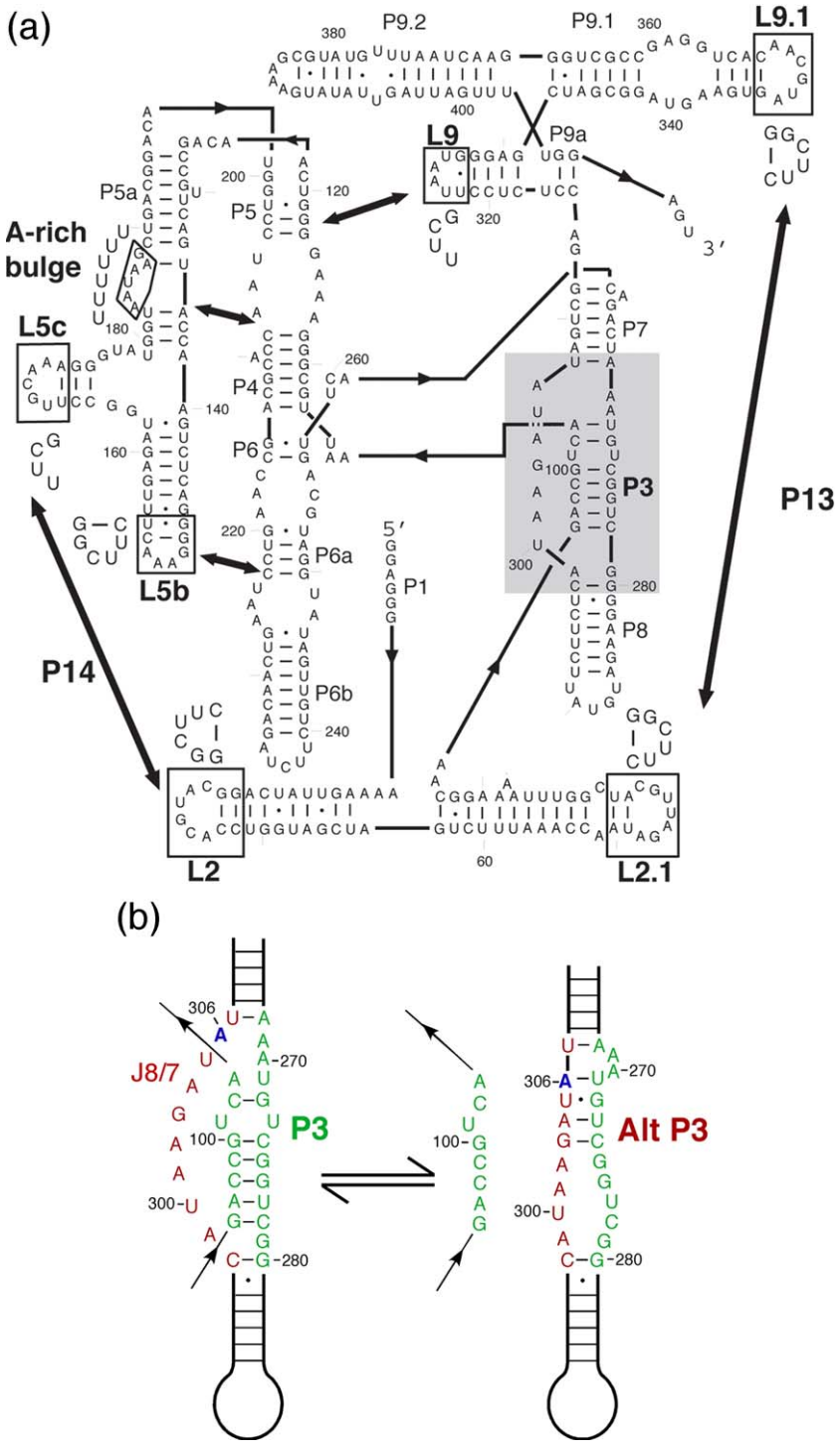


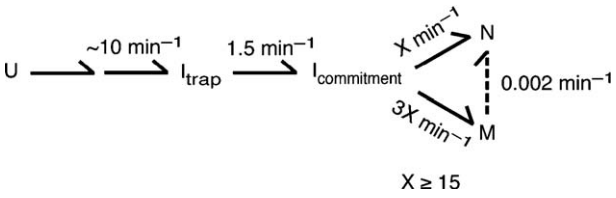
Figure 1. Secondary structure and alternative base-pairs of the *Tetrahymena* ribozyme. (a) Secondary structure. The region that can exchange between the P3 and alt P3 structures is highlighted in gray. Regions that were mutated to abolish tertiary contacts are boxed and the substituted residues are shown adjacent to each box (see Figure 4(b)). The five long-range contacts are indicated by arrows. (b) Exchange of P3 and alt P3. The strands that pair to form P3 are shown in green, and J8/7, which is unpaired when P3 is formed but is paired in alt P3, is shown in red. Residue 306 within J8/7, which was substituted to 2AP for fluorescence experiments (see Figure 3), is labeled and shown in blue.

tion and plasmid maintenance,[19,20] and these alternative structures may be akin to the kinetic traps observed in folding. Further, while it has been suggested that kinetic traps are avoided or resolved *in vivo* by RNA chaperones,[21–24] understanding how they are avoided or resolved in the cellular milieu will require a deeper understanding of the properties of these misfolded species.

The group I RNA from *Tetrahymena thermophila* has been a valuable model of RNA folding, structure, and function (Figure 1(a)). Early studies of its folding provided clear evidence for the formation of one or more kinetically trapped species.[6,7,25] Further work demonstrated that the ability of the RNA to form a non-native secondary structure, termed alt P3, which replaces the long-range P3 base-pairs in the conserved core, favored formation of a kinetic trap and led to a model in which alt P3 was present in the kinetic trap (Figure 1(b)).[26–28] Other work indicated that native contacts within a separate folding domain (P4-P6) are disrupted during continued folding of at least one trapped intermediate, suggesting that native structure stabilizes the alternative helix.[29]

Efforts to further probe the properties of folding intermediates and steps suggested that changes in nucleotide sequence and folding conditions profoundly impacted which intermediates were formed and which folding pathways were traversed.[27,30,31] Nevertheless, kinetic investigations using the onset of catalytic activity defined a minimal kinetic framework for folding under a defined set of conditions, revealing a branched folding pathway and identifying the positions of two distinct kinetically trapped intermediates within the folding pathway (Scheme 1).[28] Under commonly used conditions, essentially the entire population forms the first intermediate, I_{trap}, which then folds to a transient intermediate from which folding partitions to give native (N) and misfolded (M) ribozyme. M re-folds slowly to N, on the timescale of hours under typical *in vitro* conditions.

Elucidation of this folding pathway and the resulting ability to selectively accumulate either of these folding intermediates now allows the properties of the intermediates to be probed individually. Here we probe M, the longer-lived of the two kinetically trapped intermediates, using chemical footprinting, fluorescence, mutagenesis, enzymatic activity assays, and molecular mechanics modeling. The results yield two paradoxes. First, although the alternative secondary structure alt P3 causes misfolding, it is not present in the misfolded state; apparently the presence of alt P3 biases folding to a pathway that gives M, but alt P3 is exchanged during folding with the native P3 structure, and this exchange occurs prior to formation of M. The second paradox is that structural probes suggest only limited, local differences between the misfolded and native ribozyme species, but onset of enzymatic activity assays show that the folding transition from M to the native state involves extensive unfolding that includes all of the native contacts on the molecule's periphery; consideration of the structure and interconnections of the native ribozyme leads to a model in which the misfolded structure has a distinct topology from the native fold so that a topological barrier to rearrangement causes very slow re-folding of M to the native state.



Scheme 1.

Results and Discussion

The misfolded species, M, is highly structured and nearly identical to the native ribozyme

Global and local similarity between the misfolded and native states

Using conditions established to give predominantly the long-lived misfolded state M (90%),[28] we accumulated this species and used chemical footprinting to probe its structure and to compare it with the unfolded and native conformers. Hydroxyl radicals cause strand scission at any residue with a solvent-exposed sugar backbone;[32–34] residues within the interior of a structured RNA are protected.[35–38] Dimethyl sulfate (DMS) modification of the base-pairing faces of A and C residues is detected by reverse transcription and gives information about the accessibility of these surfaces.[39,40]

Extensive regions of the misfolded ribozyme were protected from hydroxyl radical-mediated cleavage (Figure 2(a); see also Figure S1), indicating the presence of a well-defined interior. Importantly, all of the regions known to form five long-range tertiary contacts (arrows in Figure 1(a)) show the same patterns of protections and enhancements in the native and misfolded states relative to the unfolded state (Figure 2(a) and (b)). These results suggest that all of the native peripheral tertiary contacts are formed in the M species. Further evidence for the presence of all of the long range tertiary contacts comes from analogous similarities in DMS accessibility in the native and misfolded states and the large effects of mutations within these contacts on the rate of conversion of M to N (see Figure 4(b) below).

The overall hydroxyl radical protection patterns that we observe for M and N are similar to those obtained in previous experiments in which these species were not distinguished.[35,36,41] Regions of the conserved core, including P4, P6, P6a, P3, P8, and J8/7, are protected relative to the unfolded ribozyme, as would be expected for segments that are buried within the globular structure or closely packed against its surface.[42,43] In contrast, helices that are expected to protrude from the globular structure into solution are even more exposed in the

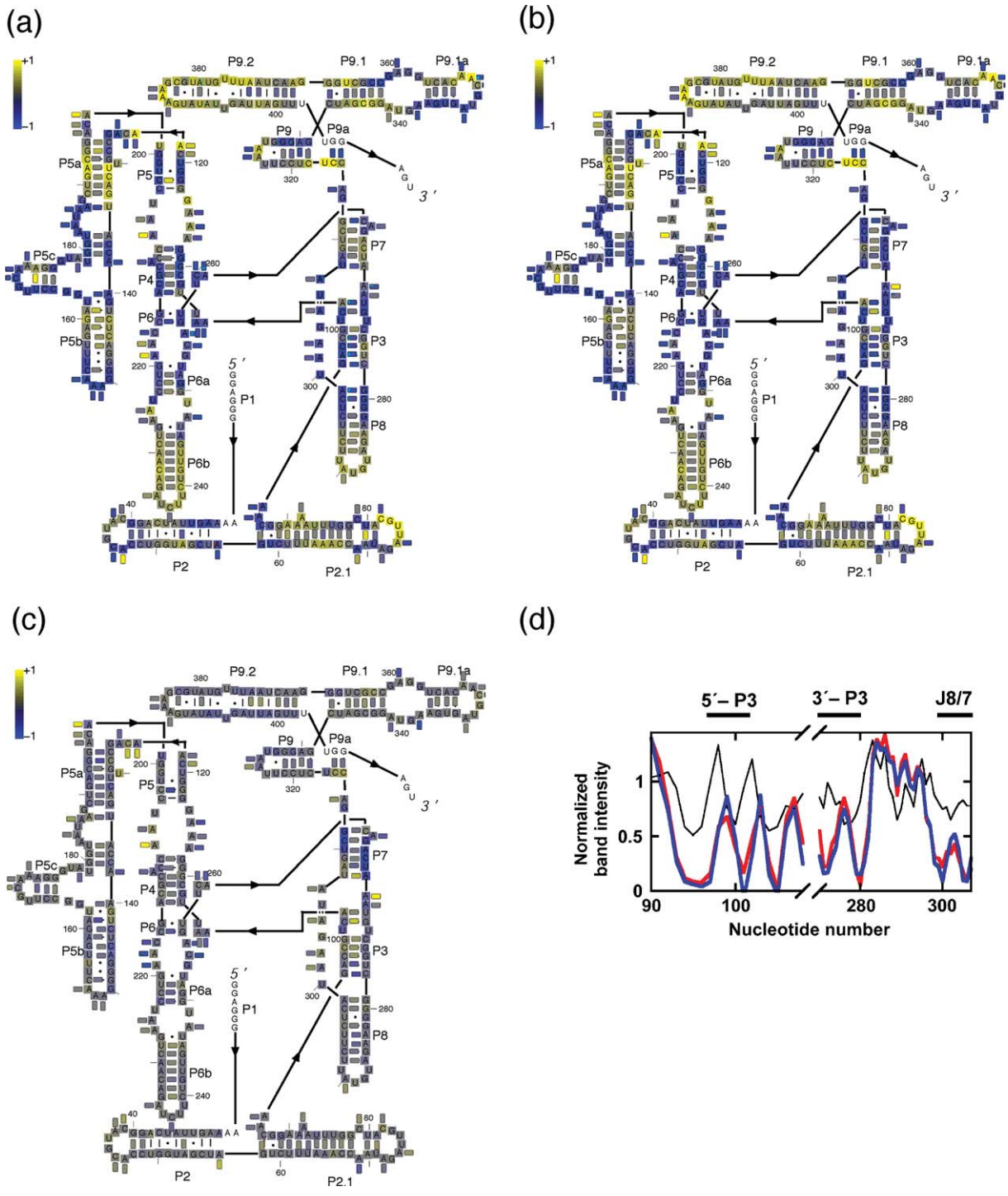


Figure 2. Chemical footprinting of misfolded and native ribozyme. (a) Footprinting of the misfolded ribozyme relative to the ribozyme in the absence of Mg^{2+} (unfolded). (b) Footprinting of the native ribozyme relative to the unfolded ribozyme. (c) Footprinting of the native ribozyme relative to the misfolded ribozyme. The boxes underneath the residue labels depict results of hydroxyl radical footprinting and the boxes adjacent to the labels depict DMS footprinting results. The color of each box corresponds to the difference in modification of the residue between the species of the comparison. For example, a blue box in (a), corresponding to a negative intensity value, indicates that the boxed residue was modified to a greater extent for the unfolded ribozyme than for the misfolded ribozyme. (d) Comparison of band intensities in hydroxyl radical footprinting for the P3 region of the unfolded state (black), the misfolded state (red), and the native state (blue). All results shown represent the averages of at least three independent determinations.

folded than unfolded species (P6b, P8, and P9.2), presumably because these helices transiently adopt positions in the unfolded ensemble that give at least weak protection from solvent. Further, the quantitative protection patterns of M and N are identical at nearly all residues (Figure 2(c); for discussion of local differences, see "A topological barrier to folding" below). Although DMS modification

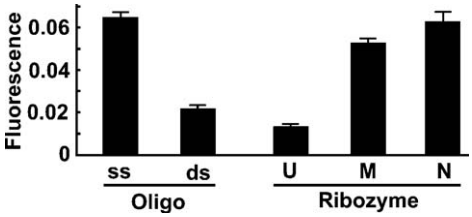


Figure 3. P3 formation monitored by 2AP fluorescence. Fluorescence of the A306(2AP) ribozyme in the unfolded state (U), the misfolded state (M), and the native state (N). Relative fluorescence levels are also shown for the 2AP-containing oligonucleotide used to construct the ribozyme (residues 298–311) free in solution (ss), and the same oligonucleotide upon duplex formation with a complementary oligonucleotide (ds). All fluorescence values are shown relative to that for the same concentration of free 2AP.

patterns are more difficult to interpret quantitatively, the patterns for the N and M species were also highly similar (Figure 2(c)). These results strongly suggest that most of the structure in the misfolded species is the same as in the native ribozyme.

The first paradox: although an alternative secondary structure (the alt P3 helix) promotes formation of the misfolded state, this misfolded state contains the native P3 helix

Woodson and colleagues provided strong evidence for a link between the P3 helix and the formation of a misfolded state (see Figure 1(b)).[26,27] Modification interference assays, in which the precursor RNA was chemically modified prior to initiating folding and then native and misfolded conformers were separated by native gel electrophoresis, showed that modifications to residues that form P3 promoted misfolding. Inspection of the intron sequence led to the proposal that one of the P3 strands could pair with the J8/7 strand to form a local helix termed alt P3. Subsequent mutagenesis of residues involved in P3 and alt P3 provided strong evidence that alt P3 formation causes misfolding. Nevertheless, alt P3 is not expected to form a highly stable duplex, as it contains only three Watson–Crick base-pairs and one wobble base-pair; this recognition and the observation by Williamson and colleagues that native structure can slow folding[29] was combined with the alt P3 results to give a model that has been widely accepted: the misfolded form has the alt P3 helix formed instead of P3, and this non-native secondary structure is stabilized, at least in part, by native interactions.[11,28,44] To our surprise, the hydroxyl radical footprinting and additional fluorescence experiments described below suggested that this misfolded species has the normal P3 helix rather than the alt P3 helix. These results are in full agreement with the previous data but not the model; i.e. the data showed that disruption of alt P3 or stabilization of P3 prior to folding favored native folding but did not show that alt P3 is present in the misfolded state.

As noted above, hydroxyl radical footprinting provides information about solvent accessible regions of structured RNAs and is insensitive to simple changes from unstructured single strands to duplexes in solution. Nevertheless, rearrangement of a segment of RNA from single-stranded to double-stranded within the core of a structured RNA might be expected to give changes in solvent exposure and thus in hydroxyl radical cleavage. The exception to this expectation would be for a region that is so thoroughly buried that no cleavage above background is observed so that, in essence, no information is provided about a rearrangement. Fortunately, the P3/ alt P3 region of the *Tetrahymena* RNA exhibits a "saw tooth" pattern of protection, suggesting that some of the residues are partially solvent accessible and that this accessibility varies from position to position (Figure 2(d)). Further, molecular modeling of a folded ribozyme with alt P3 replacing P3 gives changes in exposed surface area for the residues involved in the P3 and alt P3 helices, as well as residues in other regions (Figure S2). While the details of this model would not be expected to be accurate, the observation of significant changes in surface exposure suggests that hydroxyl radical footprinting experiments would likely detect a change from alt P3 to P3. However, the native and misfolded conformers gave indistinguishable hydroxyl radical protection patterns within the P3 and alt P3 regions (Figure 2(c); see also Figure S1). The simplest interpretation of these results is that the native P3, not alt P3, is present in the M state.

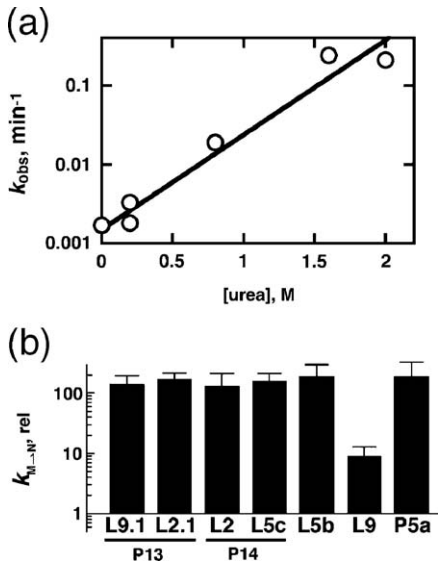


Figure 4. Ribozyme re-folding from the misfolded state to the native state monitored by enzymatic activity.[28] (a) Dependence of the re-folding rate on urea concentration under standard conditions (37 °C, 10 mM Mg^{2+}, 50 mM Na-Mops (pH 7.0)). (b) Re-folding of ribozyme variants with abrogated tertiary contacts. Rate constants for re-folding are shown relative to the wild-type ribozyme (0.0018 min^{-1} under standard conditions). As indicated below each bar, a region was substituted to prevent formation of a tertiary contact. The nucleotide changes are shown in Figure 1(a).

To test this interpretation *via* an independent assay, we constructed a ribozyme with a single 2-aminopurine (2AP) substitution within the P3/alt P3 region. 2AP is a fluorescent base analog of adenine, and its fluorescence is highly sensitive to its surrounding environment. Base stacking interactions strongly quench its fluorescence, such that 2AP can provide a sensitive probe for duplex formation.[45,46]

2AP was substituted for A306, as this residue is base-paired in alt P3 but single-stranded when P3 is formed (see Figure 1(b)). The modified ribozyme behaved similarly to the wild-type ribozyme in folding, giving a long-lived misfolded conformation that ultimately re-folded to the native state (see Folding of A306(2AP) ribozyme in Supplementary Data). The unfolded ribozyme gave a fluorescence value near background, only 1.3% of that for free 2AP (Figure 3). This value is similar to that of 2AP in a model duplex formed by the same 2AP-containing oligonucleotide used to construct the ribozyme (2.2%), consistent with the presence of alt P3 in the unfolded ribozyme. When the ribozyme was folded to the native state, the relative fluorescence level increased to 6.3%. This value is the same within error as that of the single-stranded 2AP-containing oligonucleotide free in solution, indicating that residue 306 is unpaired in the native state, as expected upon formation of P3. The fluorescence of the misfolded state (5.3%) is similar to that for the native state and is considerably higher than expected for 2AP in a duplex, based on the fluorescence of this and other 2AP-containing oligonucleotides in duplexes (Figure 3 and data not shown).[47,48] Thus, these results indicate that residue 306 is not stacked in a duplex in the M state, providing further evidence against the presence of alt P3.

Based on the fluorescence and hydroxyl radical footprinting data we conclude that P3 is present in the misfolded ribozyme. Indeed, recent oligonucleotide hybridization experiments have been interpreted to suggest that P3 replaces alt P3 prior to forming M.[49] Further strong support for this conclusion comes from the result, described below, that the misfolded state has catalytic activity.

We investigated whether the misfolded ribozyme possesses any catalytic activity because we were intrigued by the finding that it contains extensive native structure. Although previous experiments had shown that M does not significantly cleave the standard oligonucleotide substrate that mimics the 5′-splice site region,[10,25,28] it was possible that the active site remains intact but that a local rearrangement prevents docking of the P1 duplex, containing this substrate, into the ribozyme's catalytic core. Subsequent direct binding assays indeed showed that the P1 duplex remains predominantly undocked, and thus out of the active site, in the M state (P. Tijerina, H. Bhaskaran & R.R., unpublished results).

We therefore measured cleavage of a 3′-splice site mimic (equation (1)):

$$E + UCGA_5 \rightarrow E + UCG + A_5 \qquad (1)$$

a reaction in which the substrate binds in the guanosine site and is cleaved in the absence of a docked P1 duplex, presumably by hydrolysis with a hydroxide ion taking the place of the 5′-exon nucleophile in a reaction analogous to the second step of self-splicing.[50] While this reaction does not require P1 duplex docking, it does require that the active site be arranged appropriately for reaction.[51]

The misfolded ribozyme cleaved the 3′-splice site mimic, reacting with essentially the same efficiency as the native ribozyme. Misfolded ribozyme gave a second order rate constant of 550(±60) M^{-1} min^{-1} from three independent determinations, compared to 670(±60) M^{-1} min^{-1} for the native ribozyme (37 °C, 50 mM Mg^{2+}). To confirm that the misfolded ribozyme was not re-folding to N while the reaction was being followed (4–5 h, <10% re-folding expected), with the native ribozyme giving the cleavage attributed to M, we determined the fraction of native ribozyme at the end of the 3′-splice site cleavage reaction by determining the fraction of added 5′-splice site substrate that was rapidly cleaved in the presence of guanosine. The fractions of misfolded and native ribozyme were as expected, with little re-folding occurring during the 3′-splice site cleavage reaction (see Cleavage of 3′-splice site mimic and Figure S3A in Supplemental Data).

We performed two additional sets of control experiments to confirm that the guanosine binding site is utilized in this reaction. We first compared the rate constants for cleavage of two additional 3′-splice site mimics (GA_5 and $CUGA_5$), which are unable to form the P9.0 base-pairs and would therefore be expected to react less efficiently in the ribozyme active site than $UCGA_5$. Consistent with this expectation, these oligonucleotides were cleaved less efficiently than $UCGA_5$ by both the native and misfolded ribozyme (Figure S3B). The energetic penalties resulting from the lack of P9.0 were the same within error for the native and misfolded species, underscoring the apparent similarity of the active site between the native and misfolded species.

Second, we performed 3′-splice site cleavage reactions using ribozyme variants. The ribozyme variant lacking P5abc, which is severely compromised for catalytic activity under these conditions,[52] gave no detectable cleavage of $UCGA_5$, indicating that its cleavage reaction is at least 70-fold less efficient than that of the wild-type ribozyme. We then tested a ribozyme variant with an altered G-site (G264A/C311U), which was shown previously to exclude G from the active site but to accommodate 2-aminopurine ribonucleoside efficiently.[53] The wild-type ribozyme reacted poorly with an oligonucleotide in which G was replaced by 2AP, whereas this specificity was reversed for the ribozyme variant, with the 2AP-containing oligonucleotide reacting much more efficiently than the standard $UCGA_5$ (Figure S3C). The variant ribozyme cleaved each oligonucleotide with rate constants that were the same within error whether the ribozyme was folded to the native or misfolded conformers. The finding that the misfolded ribozyme changes its specificity

upon mutation of residues that form the G site indicates that this cleavage reaction takes place within the active site. Apparently the active site for phosphoryl transfer is intact in the misfolded species despite nearby structural differences, including the disruption of the binding site for the P1 duplex.

In addition to underscoring the structural similarity of the native and misfolded species, the result that the misfolded ribozyme possesses an intact active site provides further strong evidence against the presence of alt P3 in the misfolded ribozyme. Two conserved adenosine residues within J8/7, which would be severely constrained by base-pairing as part of alt P3, adopt highly unusual conformations in the native crystal structure and serve as ligands for catalytic metal ions.[54,55] Thus, catalytic activity would appear to be unlikely with alt P3 formed instead of the native P3.

A second folding paradox: Despite the enormous structural similarities between M and N, misfolded ribozyme converts to the native state very slowly and with extensive unfolding

The results described above indicate that the misfolded and native states are highly similar. Despite this similarity, the conversion of M to the native state is very slow, with a half time of 7 h under standard conditions of 37 °C and 10 mM Mg^{2+}†. [28] In contrast, the entire ribozyme can fold with a half time of a second or less.[44,56] The effects of solution conditions and mutations on the rate of conversion from M to N, described below, indicate that the ribozyme unfolds extensively in the course of this reaction.

Conversion from M to N was followed by the gain of enzymatic activity.[28] We first folded the ribozyme under defined conditions to give predominantly M and then varied the conditions to determine the effects of the changes in conditions on the re-folding rate. Previous results had shown, in addition to slow re-folding to N, that higher Mg^{2+} concentration further slowed re-folding, suggesting a requirement for disruption of Mg^{2+}-stabilized structure.[28] We probed the transition further by measuring the dependence on urea and monovalent ion concentration and on temperature.

Because urea destabilizes RNA secondary and tertiary structure,[57,58] an increase in the rate of a folding step by urea provides strong evidence for unfolding during the transition. We found that the rate constant for re-folding from M increased sharply with increasing urea concentration, giving an *m*-value of −1.67 kcal mol^{-1} M^{-1} (Figure 4(a)). The increase in rate constant with increasing urea concentration indicates that unfolding is required and the large magnitude of the *m*-value suggests that the unfolding is extensive. Measurements of RNA stability as a function of urea concentration gave a rough empirical correlation of the energetic effect with exposed surface area,[57] and this correlation suggests that the equivalent of 23 base-pairs become exposed to solvent as the misfolded ribozyme unfolds to the transition state for re-folding to N. This large degree of unfolding is supported by an enormous solvent isotope effect of 10–15 on the re-folding rate, suggesting that many hydrogen bonds are broken or perturbed (S. Wang, R.R. & D.H., unpublished results).

The re-folding rate was also strongly dependent on temperature, increasing more than 1000-fold from 25 to 50 °C to give an apparent activation enthalpy of 53 kcal mol^{-1} (Figure S4A). Further, the rate increased with increasing concentrations of Na^+ and K^+ (Figure S4B and data not shown). Monovalent ions may facilitate displacement of specifically bound or delocalized Mg^{2+} ions from the misfolded state or may interact preferentially with less structured species, stabilizing the more open transition state relative to the misfolded state.

To further probe the model of extensive unfolding, we determined the effect of a series of mutations on the rate constant for re-folding to N.[28] Strikingly, mutation of any of the five long-range tertiary contacts increased the rate‡ (Figure 4(b)). In each

† Despite the very slow folding transition from the misfolded to the native state, the ultimate accumulation of the native ribozyme in the absence of any input energy indicates that the native conformation is more stable than the misfolded conformation. Indeed, comparisons of the binding affinity of the peripheral element P5abc for the native and misfolded conformations of a P5abc-deleted ribozyme suggest an energetic difference of ~6 kcal/mol between the native and misfolded conformations of the wild-type ribozyme under standard conditions.[76]

‡ We infer that all of the variant ribozymes fold to the same or closely related misfolded species for the following reasons. First, all of the variant ribozyme gave quantitatively similar partitioning between the native and misfolded species, as well as similar dependences of re-folding rate on changes in Mg^{2+} concentration and temperature, suggesting a common folding pathway. DMS footprinting demonstrated that the misfolded species for one of the variants, the L5b variant, is the same or closely related to the M species of the wild-type ribozyme (R. R. and D. H., unpublished results), and the other variants were assumed to form similar M species because of their similar folding properties. Interestingly, the L5b variant ribozyme has been studied previously and was reported to avoid M, instead folding exclusively to N.[77] Upon folding under the conditions of this previous study (37 °C, 10 mM Tris-Cl (pH 8.0), 10 mM Mg^{2+}), we also do not detect the accumulation of M. However, M is readily formed and accumulated when this variant folds in the presence of 10 mM Mops (pH 8.0) at either lower temperature or higher Mg^{2+} concentration (data herein and K. Travers and D. H., unpublished results). Further, upon forming M at these conditions and then changing the conditions to those of the previous study, M re-folds rapidly to N (Figure 4(b)). Surprisingly, the Tris buffer used in the previous experiments itself increased the rate of re-folding by a factor of 50 (from 0.001 min^{-1} to 0.046 min^{-1}) relative to the same concentration of Mops buffer. Under the conditions of previous work, re-folding from M is even faster than the initial formation of N from unfolded ribozyme for the fraction that avoids M. Thus, it is possible that M is also formed by most of the ribozyme as a transient intermediate even under conditions that do not allow its detection, as it would be expected to re-fold to N with a larger rate constant than it is formed.

case the increase was substantial, with effects of 100–200-fold for all mutations except one. The smaller effect of ~tenfold with the L9 mutation may be a consequence of that interaction not being fully ablated by the mutation, that interaction not fully dissociating in the transition state for re-folding of the wild-type ribozyme, or that interaction contributing less to the stability of the M state than the other contacts. Remarkably, the results suggest that all of the tertiary interactions are broken, or weakened by at least 1–2 kcal/mol, in the re-folding transition state.

A model to account for the paradoxically slow transition from M to N: A topological barrier to folding

The striking similarity of the M and N states raises the question: how can such limited changes be so slow and require such extensive and widespread unfolding? The misfolded state appears to have all of the native long-range tertiary interactions, and chemical footprinting suggests that the structures are identical in nearly all regions. Nevertheless, there must also be differences, as the species are functionally distinct. Indeed, local differences between M and N were detected in our footprinting experiments (Figure 2(c) and Figure S1). The major difference observed by hydroxyl radical footprinting is in the conserved core and is centered around P7, where the misfolded species is less protected than the native species. Further, the M species is less protected from DMS modification within the junction between P4-P6 and P3-P8, and also in the internal loop of P9.1, which packs against P7. Considered together, the data suggest incomplete or improper packing of P4-P6 and P9 against the core element P7 in the misfolded species. Interestingly, the radius of gyration (R_g) of the M state is approximately 10% larger than that of the native state,[59] which could arise as a direct consequence of incomplete packing and expansion of the misfolded core, or because changes in packing redirect one or more of the helices that extend out of the core in a direction leading further from the molecule's center of mass. As significant expansion of the core would likely lead to reduced hydroxyl radical protection throughout the core, rather than the highly localized differences that are observed, we favor a model in which one or more helices are redirected as a result of changes in packing of structural elements within the core.

What types of structural differences are possible within the misfolded core and could account for the data? As the secondary structure is apparently the same in the two species, we are left to consider changes that involve re-positioning of one strand or structural element relative to another. One possible type of re-positioning would be movement of one strand or duplex along another, producing a change in the register of a contact within the core. Such a model could account for the slow re-folding of M, provided that the contact was stabilized enough by other contacts to exchange slowly, and the changes in packing could re-orient a helix projecting from the core to give the observed increased radius of gyration. However, the model requires *ad hoc* assumptions of a stable alternative "docking" of tertiary contacts and what would be a surprising requirement for global unfolding to re-orient local interactions; the simplest expectation from such a model would be that re-folding would be much more sensitive to the disruption of nearby tertiary contacts, instead of exhibiting near uniform effects from disruption of the long range tertiary contacts. Further, we might expect to observe a change in register of certain regions of the hydroxyl radical protection pattern, instead of the local, idiosyncratic changes observed.

We suggest that a more likely model for the M structure is that a component of the core, most likely one of the single-stranded linker segments, is moved "through" another strand or element instead of along it to produce a topological isomer of the native RNA. All of the observations obtained with multiple experimental approaches can be readily accounted for by such a model. It can easily be seen why M would be long-lived: disruption of the periphery and extensive unfolding would be needed to "untangle" the topological isomer. Only then could the strands that were improperly crossed in the misfolded isomer re-cross correctly to give N. Further, nearly all of the individual contacts could be identical, whereas the altered core topology could redirect peripheral helices to give the observed increase in R_g. Although we emphasize that we have no direct evidence for a topological isomer, we favor this model because it can readily account for the seemingly paradoxical results, whereas other models require *ad hoc* assumptions and rationalizations. Further, as we describe below, the model is sterically feasible, and it can explain the first paradox as well: the finding that alt P3 biases folding to give M, yet is not present in M.

To determine whether the topological isomer model is sterically reasonable, we generated four topological isomers of the folded RNA computationally (Figure 5 and Figure S5A). The models were evaluated based on two criteria: first, whether a minimal structural rearrangement would allow the isomerized structure to avoid steric overlap, and second, whether the rearrangement could occur at a sufficiently local level to be consistent with the highly similar overall patterns of chemical protection in the M and N states. All four topological isomers were readily rearranged manually and then refined to remove steric clashes, and in all cases it was possible to relax the structure to a local minimum similar in energy to that of the native topology. Only small local differences were observed in the predicted surface accessibilities and therefore in the predicted hydroxyl radical footprinting profiles (Figure S5B-E), indicating that extensive rearrangements are not required to accommodate any of the isomerizations.

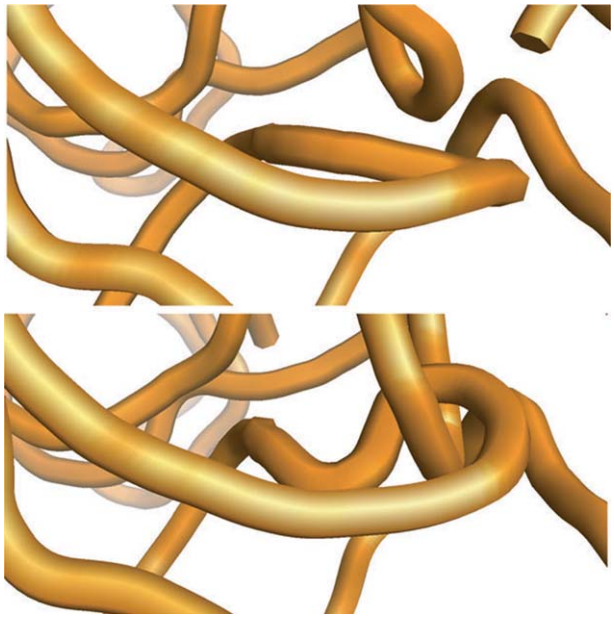


Figure 5. Molecular modeling of strand-crossings. Top panel, native strand topology. Bottom panel, crossed strand topology. The model represents the minimal conformational change required to accommodate the crossing of the strands at nucleotides C260 and U305 while satisfying all steric and molecular mechanics constraints (see Supplementary Data).

Thus, topological isomers with each of the strand crossings tested provide reasonable models to account for the data. In none of the modeled isomers are the changes in exposed surface area as distributed in different structural regions as in the experimental data (compare Figure S5B-E with Figure 2(c)), nor is the radius of gyration increased to the value of the M state (data not shown). However, this difference likely reflects a limitation of the modeling rather than a limitation of the models themselves. Because of the limited time scale of the simulations, there is a strong bias for local rearrangements, and thus the minimization protocol defines the smallest conformational change needed to accommodate the isomerization. The actual changes may be more distributed and could diminish packing and change the angular direction of some helices, thereby giving the observed changes in chemical protection (Figure 2(c)) and the observed increase in the radius of gyration.[59]

The topological model also provides a reasonable model to explain the first paradox: how the early presence of alt P3 can bias folding to give misfolding without itself remaining present in the M state (Figure 6). In this model, early formation of the native P3 favors adoption of the correct topology, which gives folding to the native state. On the other hand, the presence of alt P3 at the outset of folding allows an incorrect topology to be formed, and this incorrect topology is maintained even after alt P3 is replaced later in folding with the native P3 base-pairs. If the incorrect topology is indeed maintained through folding to M, this implies that upon reaching a partition point late in folding (abbreviated I_c in Figure 6 [28,44]), a small fraction of the ribozyme changes its topology to give the native state, while the rest maintains its topology and continues folding to M.

Conclusions and Implications

The prevalence of misfolded intermediates in RNA folding necessitates a thorough understanding of the properties that govern their formation and their stability. The misfolded species of the *Tetrahymena* ribozyme is especially tractable because it is formed by a large fraction of the population and is very long-lived.[10,28] Here we have probed the structural properties of the misfolded species and its folding transition to the native state.

Hydroxyl radical and DMS protections showed that the misfolded state is highly similar to the native state. These results, combined with fluorescence intensity measurements with a 2AP-substituted ribozyme and the partial enzymatic activity of the misfolded state, provide strong evidence for the presence of the native secondary structural element P3 in M, even though the alternative secondary structural element alt P3 is present in the unfolded state and places the RNA in a folding channel that leads to a partition point between M and N.[26,28,44]

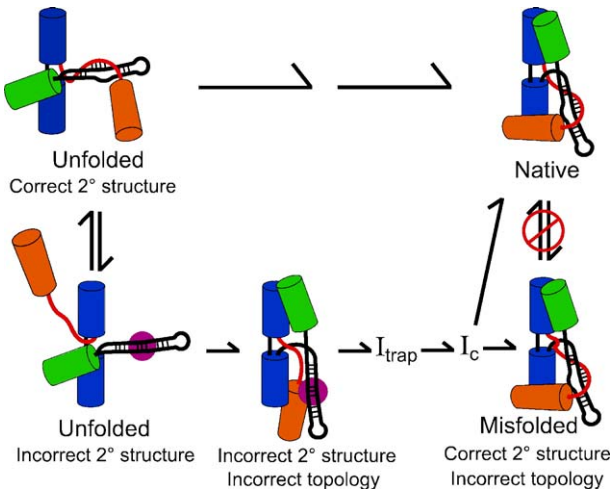


Figure 6. Model for the influence of early secondary structure on adoption of topology (see the text). The non-native secondary structure alt P3 is highlighted with a purple circle. The cartoon depicts the core elements of the ribozyme and their connections, with the peripheral elements simplified or omitted for clarity. The P4-P6 domain is blue, the P3-P8 domain is shown as black strands, and the 5′-strand of P3 is red. Simplified versions of the P2 and P9 domains are shown in orange and green, respectively. The topological difference depicted represents one of several possibilities for the misfolded ribozyme.

Despite the small, localized differences between the structures of M and N, experiments following re-folding of the misfolded RNA to the native state indicate that extensive unfolding is required. Monovalent ions, increased temperature, and urea give large increases in the re-folding rate, and mutations that disrupt the peripheral contacts greatly accelerate the re-folding transition, indicating that the peripheral contacts are broken during re-folding. All of the results can be accounted for by a difference in topology for the native and misfolded states. With a topological difference, the M and N states can be nearly identical, yet interconvert slowly and with extensive unfolding that includes the peripheral tertiary contacts.

Although the folding processes and even the types of energetic barriers encountered are certain to vary from one RNA to another, considerations of the properties of RNA suggest that topological barriers may often be present during folding of large RNAs and are likely to be more significant than for proteins of similar size. The rapid and stable formation of secondary structure makes RNA effectively a branched chain, which increases the topological complexity during folding. Further, these stable secondary structure elements, RNA duplexes, are very rigid compared with segments of a folding protein chain, which can readily lose and then re-form their secondary structures. Because of this difference, it is probably more difficult for RNA structural elements to change their relative positions within a folding intermediate, i.e. to adopt a new topology during folding. The formation of fortuitous contacts, which is common in RNA folding,[18,60] may further exacerbate topological barriers by hindering internal motions. In contrast, effects of non-native contact formation in proteins are substantially reduced because local structure is typically unstable in the absence of the cooperativity that arises only in the globally folded structure. The problems for RNA topology are expected to be progressively greater with larger RNAs because the greater structural complexity presents more opportunities for incorrect topologies to arise and for fortuitous or native contacts to stabilize them. Indeed, it is difficult to imagine the RNAs of the ribosome or spliceosome folding to their native structures without adopting incorrect topologies along the way.

Although there is no evidence for misfolding of the *Tetrahymena* RNA *in vivo*, the involvement of DExD/H-box proteins in folding and splicing of several fungal group I and group II introns strongly suggests that misfolded conformations do arise *in vivo* for these RNAs and can be sufficiently long-lived to offer a selective advantage for mechanisms that speed their re-folding.[61–63] More generally, it has been suggested that all or nearly all structured RNAs may require chaperones to fold efficiently and accurately *in vivo*.[22,61,64] It will be interesting to determine whether topological barriers are often among the barriers encountered during RNA folding, and if so, to explore the measures that nature has taken to avoid or resolve them.

Materials and Methods

Materials

Ribozyme genes encoding tertiary contact variants were prepared from the plasmid pT7L-21[65] using the QuikChange protocol (Stratagene) with oligonucleotide primers (Integrated DNA Technologies, Coralville, IA) encoding the desired changes. Complete nucleotide sequences of all ribozyme genes were confirmed by sequencing. Wildtype and variant ribozymes were prepared by runoff transcription from ScaI-linearized plasmid and purified.[10] The A306(2AP) ribozyme, containing 2-aminopurine at position 306, was produced by splint-mediated ligation.[66] Oligonucleotide substrates (Dharmacon, Lafayette, CO) and primers for reverse transcription reactions were 5′-end-labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase and purified by non-denaturing polyacrylamide gel electrophoresis.[65]

Hydroxyl radical footprinting with Fe(II)-EDTA

The ribozyme was ^{32}P-labeled at the 5′ or 3′-end using published protocols,[67,68] purified by 6% (w/v) denaturing polyacrylamide gel electrophoresis and exchanged into water using a microconcentrator (Millipore). The ribozyme was misfolded by performing a limited incubation with Mg^{2+} (10–50 mM Mg^{2+}, 50 mM Na-Mes (pH 6.0), 30 min at 25 °C), or folded to the native state by incubation at 50 °C for 30 min after addition of 10 mM Mg^{2+}.[69] After these initial folding procedures, the Mg^{2+} concentration was increased to 50 mM, if desired, or was maintained at 10 mM. (Cleavage patterns of native and misfolded ribozyme were unaffected by Mg^{2+} concentration from 10–50 mM; see Figure S1C.) Fenton reagents were added to 2 mM $(NH_4)_2Fe(II)(SO_4)_2$, 2.5 mM Na-EDTA, 6 mM ascorbic acid, and footprinting reactions were quenched after 10 min at 25 °C by addition of a half volume of 30 mM thiourea. Cleavage products and a control sample cleaved by ribonuclease T1 were separated by 8% and 20% denaturing polyacrylamide gel electrophoresis with different running times to resolve different regions of the RNA, imaged using a Phosphorimager (Amersham Biosciences, Piscataway, NJ), and quantified using single-band fitting program SAFA.[70] Data shown in Figure 2 and Figure S1A and B represent the averages of at least four independent determinations.

DMS footprinting

Misfolded and native ribozyme were prepared as above except that misfolded ribozyme was prepared by incubating in the presence of 10 mM Mg^{2+}, 50 mM Na-Mops (pH 7.0), for 15 min at 25 °C. Footprinting was performed using 2 μM ribozyme in 25 μl of standard buffer conditions (50 mM Na-Mops (pH 7.0), 10 mM $MgCl_2$) at 15 °C. DMS (1 μl of 16% DMS in ethanol) was added to a final concentration of 0.64% and allowed to react for 1 min. Further reaction was blocked by addition of 475 μl of 4 M β-mercaptoethanol, 0.3 M Na-acetate. Modified residues within the ribozyme were detected by primer extension using ^{32}P-labeled primers and AMV reverse transcriptase (Amersham) essentially as described[39] using five primers

with 3′-ends complementary to residues 120, 200, 240, 311, and 390 of the ribozyme. Reverse transcription products were separated using 8% denaturing polyacrylamide gel electrophoresis, imaged using a Phosphorimager, and quantified by boxing each band by hand using ImageQuant. Data shown represent the averages of at least three independent determinations. Re-analysis of selected DMS footprinting data using SAFA[70] gave similar results.

Fluorescence measurements of a ribozyme containing 2AP at nucleotide 306

To construct the dye-labeled ribozyme, RNAs corresponding to nucleotides 22–297 and 312–409 of the ribozyme were produced by *in vitro* transcription from PCR-generated templates and ligated with an oligonucleotide corresponding to nucleotides 298–311, with 2-aminopurine substituted at position 306. The oligonucleotide (Dharmacon, Lafayette, CO) was checked for homogeneity by gel electrophoresis and used without further purification. Ligation reactions, which contained this oligonucleotide (10 μM), RNA transcripts (10 μM), and complementary DNA oligonucleotides (10 μM), were annealed in the presence of 100 mM NaCl, 1 mM Tris-Cl (pH 8.0), 0.1 mM EDTA, by heating to 95 °C for 5 min and slowly cooling to room temperature. The annealed complexes were diluted to 1 μM in the presence of 66 mM Tris-Cl (pH 7.6), 8 mM $MgCl_2$, 10 mM DTT, 1.3 mM ATP, and 0.004% TritonX-100, and incubated ~12 h with 8 units/ml RNasin and 1 μM T4 DNA ligase. Ligated products were purified on denaturing polyacrylamide gels and eluted with 10 mM Tris-Cl (pH 8.0), 1 mM EDTA. The two large RNAs were ligated to the 2AP-containing oligonucleotide sequentially. The first ligation was between the 3′-fragment of the ribozyme and the oligonucleotide and included a complementary oligonucleotide of sequence 5′-CTAGCTCCCATTAAGGAGAGGTCCGACTATATCTTATG. The ligated product was purified and then ligated to the 5′-fragment of the ribozyme in a reaction that contained a complementary oligonucleotide of sequence 5′-GAGAGGTCCGACTATATCTTATGAGAAGAATACATCTTCCCCGACCG.

Fluorescence measurements were made using a SPEX fluorimeter (Jobin-Yvon). Excitation was at 310 nm (4 nm slit width), and emission was measured at 372 nm (8 nm slit width). All measurements were made at 10 °C in 50 mM K-Mops (pH 7.0). $MgCl_2$, when present, was 10 mM. RNAs were pre-annealed by heating to 95 °C and cooling to room temperature. The native and misfolded species were produced by incubating the ribozyme with 10 mM Mg^{2+} for 30 min at 50 °C or 25 °C, respectively.

Activity measurements of folding

A population of ~90% misfolded ribozyme was generated by adding 10 mM Mg^{2+} to the ribozyme at 25 °C and incubating 10–20 min under standard conditions (50 mM Na-Mops (pH 7.0), 10 mM Mg^{2+}). Solution conditions and temperature were then adjusted as desired to follow re-folding to the native state. To monitor the reaction, trace amounts of ^{32}P-labeled oligonucleotide substrate ($CCCUCUA_5$, S*) were added to aliquots of the re-folding reaction at various times. For ribozyme re-folding reactions under conditions that give rapid re-folding of misfolded ribozyme, a modified procedure was used. Aliquots of the re-folding reaction were first added to a quench solution, raising the Mg^{2+} concentration to 50 mM and diluting urea or Na^+ if necessary to block further re-folding, and then S* was added. For all conditions, the fraction of ribozyme that was native was determined by measuring the fraction of S* that was rapidly cleaved to the shorter oligonucleotide product (within 1 min). Cleavage reactions were quenched by addition of two volumes of stop solution (90% formamide, 20 mM EDTA, 0.01% (w/v) xylene cyanol, 0.01% (w/v) bromophenol blue), and substrate and product were separated by 20% polyacrylamide/7M urea gel electrophoresis.

The fraction of substrate cleaved using this procedure is a good measure of the fraction of native ribozyme because (1) S* binds to native and misfolded ribozyme with similar rate constants,[28] so the fraction of S* that is bound by the native ribozyme is approximately equal to the fraction of ribozyme that is native and (2) S* is rapidly cleaved by native ribozyme (10 min^{-1}) but is released slowly from misfolded ribozyme (0.02 min^{-1}), preventing significant cleavage of S* that binds to the misfolded ribozyme within the 1 min that was allowed for the cleavage reaction. We confirmed that these conditions hold for all the ribozyme variants (data not shown), such that their re-folding could be followed using the same experimental protocol.

Determination of the urea *m*-value for re-folding of misfolded ribozyme

Rate constants for re-folding to the native state in the presence of various concentrations of urea were determined as described above. The dependence of the free energy change for a folding transition on the concentration for denaturant, termed the *m*-value, has been used extensively as a diagnostic tool in studies of proteins[71,72] and has more recently been applied to RNA.[8,15,16,57,73] For both RNA and protein, linear relationships have been observed between the free energy changes of folding transitions and the concentration of denaturants such as urea (equation (2)), in which ΔG_0 represents the free energy change for the folding transition in the absence of denaturant:

$$\Delta G = \Delta G_0 + m[\text{urea}] \quad (2)$$

The rate constant for a folding transition can described as an equilibrium between the ground and transition states using transition-state theory, and can therefore be related analogously to denaturant concentration (equation (3)), where R is the gas constant (1.98 cal mol^{-1} K^{-1}) and T is absolute temperature. Thus, equation (3) was fit to the re-folding data to determine the urea *m*-value for re-folding of the misfolded ribozyme:

$$\ln k = \ln k_0 - \frac{m}{RT}[\text{urea}] \quad (3)$$

Cleavage of 3′-splice site mimic

Cleavage of trace concentrations of ^{32}P-labeled $UCGA_5$ by subsaturating concentrations of misfolded and native ribozyme was measured to determine values of (k_{cat}/K_M). After generating misfolded or native ribozyme as described above, reactions were performed at 37 °C in 50 mM Na-Mops (pH 7.0), 50 mM Mg^{2+}.

Molecular modeling

Molecular models of topological isomers were generated with the CNS (Crystallography and NMR System) molecular mechanics package[74] using the 2.8 Å

crystal structure of the truncated *Tetrahymena* ribozyme (PDB ID: 1GRZ) as initial atomic coordinates.[42] Four nucleotide pairs were identified (G100-A306, A105-C260, C260-U305, A261-G313) as potential sites for strand crossings based on two criteria. First, at least one backbone atom was required to be within 6 Å of a backbone atom from the potential crossing partner in the native structure. Second, the nucleotides were chosen from regions expected to be outside of Watson–Crick double helices in the native structure. The ribozyme structure was manipulated in the program PyMOL by rotating and translating the nucleotides until the two strands were crossed. Conjugated gradient minimization was used to relax any resulting steric overlap. The accessible surface areas (ASA) of the five sugar carbon atoms were computed and summed at each nucleotide for the native and crossed topologies. The ratio of the ASA values for the native and crossed structures was used to predict the expected relative change in hydroxyl radical footprinting given a conformational rearrangement.[35,75]

Acknowledgements

This work was funded by National Institutes of Health Grant P01-GM066275 (to D.H.), National Institutes of Health Grant R01-GM070456 and Welch Foundation Grant F-1563 (to R.R.), and through the NIH Roadmap for Medical Research grant U54 GM072970. Information on the National Centers for Biomedical Computing can be obtained§. A.L. is a Damon Runyon Cancer Research Foundation Fellow; K.J.T. was a Beckman Scholar, and R.D. was supported by an Abbott Laboratories Stanford Graduate Fellowship.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2006.08.024

References

1. Lindahl, T., Adams, A. & Fresco, J. R. (1966). Renaturation of transfer ribonucleic acids through site binding of magnesium. *Proc. Natl Acad. Sci. USA*, **55**, 941–948.
2. Adams, A., Lindahl, T. & 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.
3. Cole, P. E. & Crothers, D. M. (1972). Conformational changes of transfer ribonucleic acid. Relaxation kinetics of the early melting transition of methionine transfer ribonucleic acid (*Escherichia coli*). *Biochemistry*, **11**, 4368–4374.
4. Walstrum, S. A. & Uhlenbeck, O. C. (1990). The self-splicing RNA of *Tetrahymena* is trapped in a less active conformation by gel purification. *Biochemistry*, **29**, 10573–10576.
5. Woodson, S. A. & Cech, T. R. (1991). Alternative secondary structures in the 5′ exon affect both forward and reverse self-splicing of the *Tetrahymena* intervening sequence RNA. *Biochemistry*, **30**, 2042–2050.
6. Zarrinkar, P. P. & Williamson, J. R. (1994). Kinetic intermediates in RNA folding. *Science*, **265**, 918–924.
7. Downs, W. D. & Cech, T. R. (1996). Kinetic pathway for folding of the *Tetrahymena* ribozyme revealed by three UV-inducible crosslinks. *RNA*, **2**, 718–732.
8. Pan, J., Thirumalai, D. & Woodson, S. A. (1997). Folding of RNA involves parallel pathways. *J. Mol. Biol.* **273**, 7–13.
9. Pan, T. & Sosnick, T. R. (1997). Intermediates and kinetic traps in the folding of a large ribozyme revealed by circular dichroism and UV absorbance spectroscopies and catalytic activity. *Nature Struct. Biol.* **4**, 931–938.
10. Russell, R. & Herschlag, D. (1999). New pathways in folding of the *Tetrahymena* group I RNA enzyme. *J. Mol. Biol.* **291**, 1155–1167.
11. Treiber, D. K. & Williamson, J. R. (1999). Exposing the kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.* **9**, 339–345.
12. Chadalavada, D. M., Knudsen, S. M., Nakano, S. & Bevilacqua, P. C. (2000). A role for upstream RNA structure in facilitating the catalytic fold of the genomic hepatitis delta virus ribozyme. *J. Mol. Biol.* **301**, 349–367.
13. Pichler, A. & Schroeder, R. (2002). Folding problems of the 5′ splice site containing the P1 stem of the group I thymidylate synthase intron: substrate binding inhibition in vitro and mis-splicing *in vivo*. *J. Biol. Chem.* **277**, 17987–17993.
14. Treiber, D. K. & Williamson, J. R. (2001). Beyond kinetic traps in RNA folding. *Curr. Opin. Struct. Biol.* **11**, 309–314.
15. Fang, X. W., Pan, T. & Sosnick, T. R. (1999). Mg2+-dependent folding of a large ribozyme without kinetic traps. *Nature Struct. Biol.* **6**, 1091–1095.
16. Fang, X. W., Thiyagarajan, P., Sosnick, T. R. & Pan, T. (2002). The rate-limiting step in the folding of a large ribozyme without kinetic traps. *Proc. Natl Acad. Sci. USA*, **99**, 8518–8523.
17. Swisher, J. F., Su, L. J., Brenowitz, M., Anderson, V. E. & Pyle, A. M. (2002). Productive folding to the native state by a group II intron ribozyme. *J. Mol. Biol.* **315**, 297–310.
18. Bartley, L. E., Zhuang, X., Das, R., Chu, S. & Herschlag, D. (2003). Exploration of the transition state for tertiary structure formation between an RNA helix and a large structured RNA. *J. Mol. Biol.* **328**, 1011–1026.
19. Yanofsky, C. (1981). Attenuation in the control of expression of bacterial operons. *Nature*, **289**, 751–758.
20. Franch, T., Gultyaev, A. P. & Gerdes, K. (1997). Programmed cell death by hok/sok of plasmid R1: processing at the hok mRNA 3′-end triggers structural rearrangements that allow translation and antisense RNA binding. *J. Mol. Biol.* **273**, 38–51.
21. Karpel, R. L., Miller, N. S. & Fresco, J. R. (1982). Mechanistic studies of ribonucleic acid renaturation by a helix-destabilizing protein. *Biochemistry*, **21**, 2102–2108.
22. Herschlag, D. (1995). RNA chaperones and the RNA folding problem. *J. Biol. Chem.* **270**, 20871–20874.
23. Lorsch, J. R. (2002). RNA chaperones exist and DEAD box proteins get a life. *Cell*, **109**, 797–800.
24. Schroeder, R., Barta, A. & Semrad, K. (2004). Strategies for RNA folding and assembly. *Nature Rev. Mol. Cell. Biol.* **5**, 908–919.

§ http://nihroadmap.nih.gov/bioinformatics

25. Emerick, V. L. & Woodson, S. A. (1994). Fingerprinting the folding of a group I precursor RNA. *Proc. Natl Acad. Sci. USA*, **91**, 9675–9679.
26. Pan, J. & Woodson, S. A. (1998). Folding intermediates of a self-splicing RNA: mispairing of the catalytic core. *J. Mol. Biol.* **280**, 597–609.
27. Pan, J., Deras, M. L. & Woodson, S. A. (2000). Fast folding of a ribozyme by stabilizing core interactions: Evidence for multiple folding pathways in RNA. *J. Mol. Biol.* **296**, 133–144.
28. Russell, R. & Herschlag, D. (2001). Probing the folding landscape of the *Tetrahymena* ribozyme: commitment to form the native conformation is late in the folding pathway. *J. Mol. Biol.* **308**, 839–851.
29. Treiber, D. K., Rook, M. S., Zarrinkar, P. P. & Williamson, J. R. (1998). Kinetic intermediates trapped by native interactions in RNA folding. *Science*, **279**, 1943–1946.
30. Rook, M. S., Treiber, D. K. & Williamson, J. R. (1998). Fast folding mutants of the *Tetrahymena* group I ribozyme reveal a rugged folding energy landscape. *J. Mol. Biol.* **281**, 609–620.
31. Pan, J. & 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.
32. Hertzberg, R. P. & Dervan, P. B. (1984). Cleavage of DNA with methidiumpropyl-EDTA-iron(II): reaction conditions and product analyses. *Biochemistry*, **23**, 3934–3945.
33. Tullius, T. D. & Dombroski, B. A. (1985). Iron(II) EDTA used to measure the helical twist along any DNA molecule. *Science*, **230**, 679–681.
34. Balasubramanian, B., Pogozelski, W. K. & Tullius, T. D. (1998). DNA strand breaking by the hydroxyl radical is governed by the accessible surface areas of the hydrogen atoms of the DNA backbone. *Proc. Natl Acad. Sci. USA*, **95**, 9738–9743.
35. Latham, J. A. & Cech, T. R. (1989). Defining the inside and outside of a catalytic RNA molecule. *Science*, **245**, 276–282.
36. Celander, D. W. & Cech, T. R. (1991). Visualizing the higher order folding of a catalytic RNA molecule. *Science*, **251**, 401–407.
37. Murphy, F. L. & Cech, T. R. (1993). An independently folding domain of RNA tertiary structure within the *Tetrahymena* ribozyme. *Biochemistry*, **32**, 5291–5300.
38. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E. *et al.* (1996). Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science*, **273**, 1678–1685.
39. Inoue, T. & Cech, T. (1985). Secondary structure of the circular form of the *Tetrahymena* rRNA intervening sequence: a technique for RNA structure analysis using chemical probes and reverse transcriptase. *Proc. Natl Acad. Sci. USA*, **82**, 648–652.
40. Moazed, D., Stern, S. & Noller, H. F. (1986). Rapid chemical probing of conformation in 16 S ribosomal RNA and 30 S ribosomal subunits using primer extension. *J. Mol. Biol.* **187**, 399–416.
41. Sclavi, B., Sullivan, M., Chance, M. R., Brenowitz, M. & Woodson, S. A. (1998). RNA folding at millisecond intervals by synchrotron hydroxyl radical footprinting. *Science*, **279**, 1940–1943.
42. Golden, B. L., Gooding, A. R., Podell, E. R. & Cech, T. R. (1998). A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. *Science*, **282**, 259–264.
43. Adams, P. L., Stahley, M. R., Kosek, A. B., Wang, J. & Strobel, S. A. (2004). Crystal structure of a self-splicing group I intron with both exons. *Nature*, **430**, 45–50.
44. Russell, R., Zhuang, X., Babcock, H. P., Millett, I. S., Doniach, S., Chu, S. & Herschlag, D. (2002). Exploring the folding landscape of a structured RNA. *Proc. Natl Acad. Sci. USA*, **99**, 155–160.
45. Menger, M., Tuschl, T., Eckstein, F. & Porschke, D. (1996). Mg(2+)-dependent conformational changes in the hammerhead ribozyme. *Biochemistry*, **35**, 14710–14716.
46. Law, S. M., Eritja, R., Goodman, M. F. & Breslauer, K. J. (1996). Spectroscopic and calorimetric characterizations of DNA duplexes containing 2-aminopurine. *Biochemistry*, **35**, 12329–12337.
47. Ward, D. C., Reich, E. & Stryer, L. (1969). Fluorescence studies of nucleotides and polynucleotides. I. Formycin, 2-aminopurine riboside, 2,6-diaminopurine riboside, and their derivatives. *J. Biol. Chem.* **244**, 1228–1237.
48. Rachofsky, E. L., Seibert, E., Stivers, J. T., Osman, R. & Ross, J. B. (2001). Conformation and dynamics of abasic sites in DNA investigated by time-resolved fluorescence of 2-aminopurine. *Biochemistry*, **40**, 957–967.
49. Ohki, Y., Ikawa, Y., Shiraishi, H. & Inoue, T. (2002). Mispaired P3 region in the hierarchical folding pathway of the *Tetrahymena* ribozyme. *Genes Cells*, **7**, 851–860.
50. Inoue, T., Sullivan, F. X. & Cech, T. R. (1986). New reactions of the ribosomal RNA precursor of *Tetrahymena* and the mechanism of self-splicing. *J. Mol. Biol.* **189**, 143–165.
51. Shan, S., Yoshida, A., Sun, S., Piccirilli, J. A. & Herschlag, D. (1999). Three metal ions at the active site of the *Tetrahymena* group I ribozyme. *Proc. Natl Acad. Sci. USA*, **96**, 12299–12304.
52. Engelhardt, M. A., Doherty, E. A., Knitt, D. S., Doudna, J. A. & Herschlag, D. (2000). The P5abc peripheral element facilitates preorganization of the *Tetrahymena* group I ribozyme for catalysis. *Biochemistry*, **39**, 2639–2651.
53. Michel, F., Hanna, M., Green, R., Bartel, D. P. & Szostak, J. W. (1989). The guanosine binding site of the *Tetrahymena* ribozyme. *Nature*, **342**, 391–395.
54. Adams, P. L., Stahley, M. R., Gill, M. L., Kosek, A. B., Wang, J. & Strobel, S. A. (2004). Crystal structure of a group I intron splicing intermediate. *RNA*, **10**, 1867–1887.
55. Stahley, M. R. & Strobel, S. A. (2005). Structural evidence for a two-metal-ion mechanism of group I intron splicing. *Science*, **309**, 1587–1590.
56. Zhuang, X., Bartley, L. E., Babcock, H. P., Russell, R., Ha, T., Herschlag, D. & Chu, S. (2000). A single molecule study of RNA catalysis and folding. *Science*, **288**, 2048–2051.
57. Shelton, V. M., Sosnick, T. R. & Pan, T. (1999). Applicability of urea in the thermodynamic analysis of secondary and tertiary RNA folding. *Biochemistry*, **38**, 16831–16839.
58. Ralston, C. Y., He, Q., Brenowitz, M. & Chance, M. R. (2000). Stability and cooperativity of individual tertiary contacts in RNA revealed through chemical denaturation. *Nature Struct. Biol.* **7**, 371–374.
59. Russell, R., Millett, I. S., Doniach, S. & Herschlag, D. (2000). Small angle X-ray scattering reveals a compact intermediate in RNA folding. *Nature Struct. Biol.* **7**, 367–370.
60. Karbstein, K. & Herschlag, D. (2003). Extraordinarily slow binding of guanosine to the *Tetrahymena* group I

ribozyme: implications for RNA preorganization and function. *Proc. Natl Acad. Sci. USA*, **100**, 2300–2305.
61. Mohr, S., Stryker, J. M. & Lambowitz, A. M. (2002). A DEAD-box protein functions as an ATP-dependent RNA chaperone in group I intron splicing. *Cell*, **109**, 769–779.
62. Seraphin, B., Simon, M., Boulet, A. & Faye, G. (1989). Mitochondrial splicing requires a protein from a novel helicase family. *Nature*, **337**, 84–87.
63. Huang, H. R., Rowe, C. E., Mohr, S., Jiang, Y., Lambowitz, A. M. & Perlman, P. S. (2005). The splicing of yeast mitochondrial group I and group II introns requires a DEAD-box protein with RNA chaperone function. *Proc. Natl Acad. Sci. USA*, **102**, 163–168.
64. Clodi, E., Semrad, K. & Schroeder, R. (1999). Assaying RNA chaperone activity *in vivo* using a novel RNA folding trap. *EMBO J.* **18**, 3776–3782.
65. Zaug, A. J., Grosshans, C. A. & Cech, T. R. (1988). Sequence-specific endoribonuclease activity of the *Tetrahymena* ribozyme: enhanced cleavage of certain oligonucleotide substrates that form mismatched ribozyme-substrate complexes. *Biochemistry*, **27**, 8924–8931.
66. Moore, M. J. & Sharp, P. A. (1992). Site-specific modification of pre-mRNA: the 2′-hydroxyl groups at the splice sites. *Science*, **256**, 992–997.
67. Donis-Keller, H., Maxam, A. M. & Gilbert, W. (1977). Mapping adenines, guanines, and pyrimidines in RNA. *Nucl. Acids Res.* **4**, 2527–2538.
68. Huang, Z. & Szostak, J. W. (1996). A simple method for 3′-labeling of RNA. *Nucl. Acids Res.* **24**, 4360–4361.
69. Herschlag, D. & Cech, T. R. (1990). Catalysis of RNA cleavage by the *Tetrahymena* thermophila ribozyme. 1. Kinetic description of the reaction of an RNA substrate complementary to the active site. *Biochemistry*, **29**, 10159–10171.
70. Das, R., Laederach, A., Pearlman, S. M., Herschlag, D. & Altman, R. B. (2005). SAFA: Semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. *RNA*, **11**, 344–354.
71. Tanford, C. (1970). Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Advan. Protein Chem.* **24**, 1–95.
72. Pace, C. N. (1975). The stability of globular proteins. *CRC Crit. Rev. Biochem.* **3**, 1–43.
73. Fang, X., Pan, T. & Sosnick, T. R. (1999). A thermodynamic framework and cooperativity in the tertiary folding of a Mg^{2+}-dependent ribozyme. *Biochemistry*, **38**, 16840–16846.
74. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
75. Tullius, T. D. & Greenbaum, J. A. (2005). Mapping nucleic acid structure by hydroxyl radical cleavage. *Curr. Opin. Chem. Biol.* **9**, 127–134.
76. Johnson, T. H., Tijerina, P., Chadee, A. B., Herschlag, D. & Russell, R. (2005). Structural specificity conferred by a group I RNA peripheral element. *Proc. Natl Acad. Sci. USA*, **102**, 10176–10181.
77. Treiber, D. K. & Williamson, J. R. (2001). Concerted kinetic folding of a multidomain ribozyme with a disrupted loop-receptor interaction. *J. Mol. Biol.* **305**, 11–21.

Edited by J. Doudna

(Received 26 May 2006; received in revised form 29 July 2006; accepted 9 August 2006)
Available online 15 August 2006