RNA (1999), *5*:1308–1325. Cambridge University Press. Printed in the USA.

Relationship between internucleotide linkage geometry and the stability of RNA

GARRETT A. SOUKUP and RONALD R. BREAKER

Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520-8103, USA

ABSTRACT

The inherent chemical instability of RNA under physiological conditions is primarily due to the spontaneous cleavage of phosphodiester linkages via intramolecular transesterification reactions. Although the protonation state of the nucleophilic 2′-hydroxyl group is a critical determinant of the rate of RNA cleavage, the precise geometry of the chemical groups that comprise each internucleotide linkage also has a significant impact on cleavage activity. Specifically, transesterification is expected to be proportional to the relative in-line character of the linkage. We have examined the rates of spontaneous cleavage of various RNAs for which the secondary and tertiary structures have previously been modeled using either NMR or X-ray crystallographic data. Rate constants determined for the spontaneous cleavage of different RNA linkages vary by almost 10,000-fold, most likely reflecting the contribution that secondary and tertiary structures make towards the overall chemical stability of RNA. Moreover, a correlation is observed between RNA cleavage rate and the relative in-line fitness of each internucleotide linkage. One linkage located within an ATP-binding RNA aptamer is predicted to adopt most closely the ideal conformation for in-line attack. This linkage has a rate constant for transesterification that is ~12-fold greater than is observed for an unconstrained linkage and was found to be the most labile among a total of 136 different sites examined. The implications of this relationship for the chemical stability of RNA and for the mechanisms of nucleases and ribozymes are discussed.

Keywords: 2′ hydroxyl; in-line attack; near-attack conformation; ribozyme; RNA structure; spontaneous RNA cleavage; transesterification

INTRODUCTION

Many ribonucleases and ribozymes cleave RNA by an intramolecular phosphoester transfer reaction. The mechanism for this reaction (Fig. 1) involves the nucleophilic attack of a 2′ oxygen on the adjacent phosphorus center (Kuimelis & McLaughlin, 1998; Oivanen et al., 1998; Raines, 1998; Zhou & Taira, 1998). The 5′- and 3′-cleavage fragments produced by this S_N2-like reaction carry 2′,3′-cyclic phosphate and 5′-hydroxyl termini (structures 3 and 4), respectively. In RNA, the nucleophilic oxygen of the 2′-hydroxyl group inherently is held in proximity to the phosphodiester linkage and is the basis for the substantial decrease in stability of RNA compared to DNA. This single chemical difference causes the average RNA linkage to be ~100,000-fold less stable than a corresponding DNA linkage under physiological conditions (Li & Breaker, 1999).

A number of chemical events can accelerate the cleavage of RNA by this phosphoester transfer pathway. For example, deprotonation of the 2′-hydroxyl group by specific base catalysis accelerates transesterification by increasing the fraction of the more nucleophilic 2′-oxyanion group relative to the 2′-hydroxyl group. Under acidic conditions, specific acid catalysis becomes a significant factor that influences the rate of RNA transesterification. Here, protonation of a nonbridging phosphate oxygen and protonation of the 5′-oxyanion leaving group both accelerate the transesterification reaction (Oivanen et al., 1998). Another parameter that is known to affect the rate of RNA transesterification is the position of the attacking nucleophile relative to the leaving group (Westheimer, 1968; Usher, 1969). The nucleophilic group must be in a “near-attack conformation” for the reaction to proceed (Lightstone & Bruice, 1996; Bruice & Lightstone, 1999). For RNA, transesterification is permitted only when the attacking 2′-oxygen

Reprint requests to: Ronald R. Breaker, Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06520-8103, USA; e-mail: ronald.breaker@yale.edu.

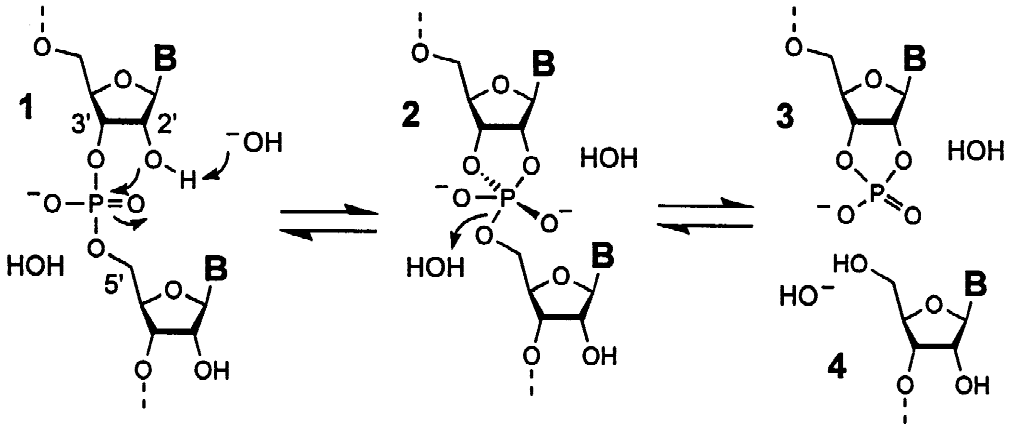

FIGURE 1. Mechanism for RNA cleavage by intramolecular phosphoester transfer. The RNA internucleotide linkage (**1**) is severed via a nucleophilic attack by the 2′ oxygen on the adjacent phosphorus center to form a pentacoordinate intermediate (**2**), with eventual departure of the 5′-oxyanion leaving group. This reaction produces RNA fragments with 2′,3′-cyclic phosphate (**3**) and 5′ hydroxyl (**4**) termini.

nucleophile is positioned in line with the 5′-oxyanion leaving group, such that the leaving group must be located directly on the opposing side of the target phosphorus center relative to the nucleophile.

Each of the mechanisms described above could be employed by RNA-cleaving enzymes to accelerate transesterification. For example, bovine pancreatic ribonuclease A (RNase A) is proposed to combine both general base catalysis and general acid catalysis to promote RNA transesterification in the first step of its catalytic mechanism (Walsh, 1979; Gerlt, 1993). In addition, the RNA phosphodiester substrate is positioned by the active site of the nuclease such that a near in-line conformation is established (Richards & Wyckoff, 1971; Carlson, 1976). When applied simultaneously by RNase A, these catalytic strategies provide a rate enhancement of nearly 10^{12}-fold over the uncatalyzed rate of RNA transesterification (Thompson et al., 1995).

The importance of the in-line orientation of nucleophile and leaving group for RNA transesterification is evident upon the examination of the chemical stability of single-stranded RNAs compared to RNA duplexes. Divalent lead ions, which accelerate the cleavage of RNA by catalyzing the 2′-hydroxyl-mediated transesterification reaction, preferentially promote cleavage in the variably structured regions of otherwise well-folded RNAs (Ciesiolka et al., 1992; Welch et al., 1997; Zagórowska et al., 1998). In the context of an A-form helix, the 2′-oxyanion group and the adjacent 5′-oxyanion leaving group of each 3′,5′-phosphodiester bond are precluded from adopting an in-line conformation, thereby making the linkage more resistant to cleavage by RNA transesterification. In contrast, single-stranded RNAs are more susceptible to spontaneous or metal-catalyzed transesterification because each linkage is free to explore conformational space and occasionally sample the reactive in-line structure. This phenomenon is made evident by the fact that the natural 3′,5′-phosphodiester bonds are more stable than their 2′,5′-phosphodiester analogs when positioned within an A-form helix (Usher & McHale, 1976). Whereas 3′,5′-phosphodiester linkages are stabilized by the helical conformation, 2′,5′-phosphodiester linkages are held in an in-line conformation that accelerates cleavage by transesterification.

RNA stability and the action of protein- and RNA-based nucleases are of fundamental importance to biological systems. Therefore, establishing the maximum enhancement that each individual catalytic mechanism could impart on the rate of RNA transesterification would greatly benefit the efforts to understand the processes of RNA degradation. For example, it has been established that the maximum rate constant that can be obtained for RNA transesterification of an average internucleotide linkage by total deprotonation of the 2′-hydroxyl group is $\sim$0.02 min^{-1} at 23 °C in 3.16 M KCl at pH 14 (Li & Breaker, 1999). This equates with an overall rate enhancement for RNA cleavage of $\sim$100,000-fold compared to the uncatalyzed rate of RNA cleavage under simulated physiological conditions. Similarly, in this study we set out to examine the magnitude of rate acceleration that can be achieved by positioning an RNA linkage for in-line nucleophilic attack to the exclusion of all other mechanisms.

RESULTS

Folded RNAs provide a range of internucleotide linkage conformations

The relative in-line character of a phosphodiester linkage in RNA is established by the combination of three structural parameters (Fig. 2A). The two most signifi-

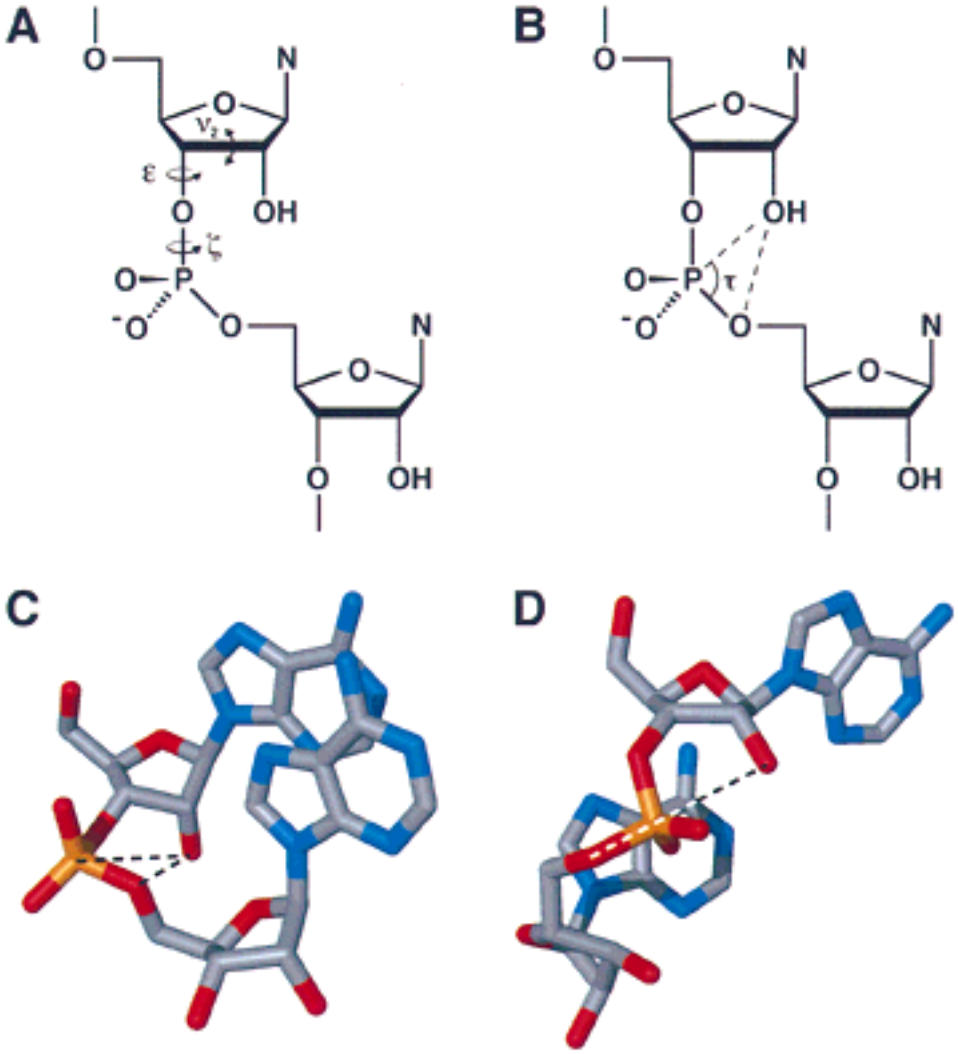

FIGURE 2. The internucleotide geometry of RNA. **A**: The relative in-line conformation of each RNA linkage is defined by the torsion angles centered on three bonds identified as ε, ζ, and ν_2. Free rotation is possible at ε and ζ, whereas more restricted rotation around ν_2 dictates the sugar pucker. **B**: For any given structure, the τ angle can be derived from a triangle whose sides are defined by the O2′-P and O2′-O5′ interatomic distances (dashed lines), and the P-O5′ bond length. A τ angle of 180° is optimal for in-line attack. Because of steric constraints, a τ angle of ~45° approximates the most unfavorable orientation that can be achieved by RNA. Likewise, an attack distance of 3.0 Å is used as the closest approach that can be achieved between the 2′ oxygen and the phosphorus center in the ground state. **C**: Model of a linkage formed by the dinucleotide ApA in a non-in-line conformation. This representative linkage resides within a standard A-form helix that establishes a τ angle of 66°, an attack distance of 3.91 Å, and an *F* value of 0.07. **D**: Model of an ApA dinucleotide with a near perfect in-line fitness. NMR analysis indicates that this structure may be formed by the A12 and A13 nucleotides of an ATP aptamer (Sassanfar & Szostak, 1993). A τ angle of 171°, an attack distance of 3.0 Å, and an *F* value of 0.95 are predicted to form upon ligand binding to the aptamer (Jiang et al., 1996; structural model 7). *N* represents any nucleobase. Carbon, nitrogen, oxygen, and phosphorus atoms are represented as gray, blue, red, and orange, respectively.

cant contributions to the in-line character of an internucleotide linkage are made by rotation around the C3′-O3′ (ε) and the O3′-P (ζ) bonds. The torsion angles centered at these two bonds determine whether the 5′-oxygen leaving group will be positioned directly on the opposite side of the electrophilic phosphorus center relative to the nucleophilic 2′ oxygen. The third structural parameter is the alteration in the sugar pucker that varies with changes in the torsion angle centered on the bond joining the C2′ and C3′ carbons (ν_2). These more subtle changes in sugar pucker also influence the overall in-line character of the linkage.

At any given time, each of the bonds identified above can adopt one of a continuum of torsion angles. As a result, there may exist a continuum of structural states that range between a perfect non-in-line conformation (τ = 45°) and a perfect in-line conformation (τ = 180°) that defines the orientation of the 2′-oxygen, phosphorus, and 5′-oxygen centers relative to each other (Fig. 2B; see Materials and Methods for definitions and equations).

Presumably, adopting torsion angles that establish a more perfect in-line orientation will produce a faster rate of RNA transesterification. However, a perfect in-line orientation in itself is not expected to be sufficient to establish rapid RNA transesterification. In addition, the interatomic distance between the 2′-oxygen nucleophile and the phosphorus electrophile is expected to influence the rate of transesterification. Up to a point, shorter "attack distances" should increase the frequency of productive collisions between the nucleophile and the electrophile, which would produce a corresponding increase in the rate of the reaction. The same three torsion angles that determine in-line orientation also establish the attack distance between the 2′-hydroxyl group and the phosphorus atom. Variations in both the in-line orientation and the interatomic distances between reactive groups define the "in-line fitness" (F) of a particular RNA linkage.

To examine the contribution that in-line fitness makes to the overall rate of RNA transesterification, we needed to find a means by which the attack distances and the torsion angles centered on bonds ε, ζ, and ν_2 could be established and maintained at specific levels of fitness. Unfortunately, designing and synthesizing an adequate series of model chemical compounds to comprehensively survey different F values would be an enormous challenge. However, many RNA aptamers (Gold et al., 1995; Osborne & Ellington, 1997) and ribozymes (Altman & Kirsebom, 1999; Lambowitz et al., 1999; McKay & Wedekind, 1999) adopt complex tertiary folds that inherently establish a variety of linkage conformations that are held in place by the global RNA structure. Moreover, the recent advances in RNA structure determination using nuclear magnetic resonance (NMR) and X-ray crystallography techniques offer atomic-level structural models for specific RNA domains that provide explicit information about the fine structures of many internucleotide linkages (Cate et al., 1996; Feigon et al., 1996; Chow & Bogdan, 1997; Patel, 1997; Ferré-D'Amaré et al., 1998). For example, an ApA dinucleotide that resides within the context of a typical A-form RNA helix[1] adopts a chemically stable non-in-line conformation that inhibits cleavage by spontaneous phosphoester transfer (Fig. 2C). In contrast, this same dinucleotide is predicted to form a near perfect in-line arrangement (Fig. 2D) in the context of an ATP-binding RNA aptamer (Sassanfar & Szostak, 1993). This demonstrates the structural variation in phosphodiester linkages that are manifested in folded RNA structures.

We set out to examine the relationship between the in-line fitness of RNA linkages and the rate constants defining the spontaneous cleavage of RNA by intramolecular transesterification. Beginning with a database[2] of RNA structures, we determined the τ angles, O2′-P attack distances, and the corresponding in-line fitness values for a total of 136 internucleotide linkages that reside in one of several chosen RNA motifs (Fig. 3). This was achieved by using a simple algorithm that defines the τ angle and the attack distance using the atomic coordinates in the database (see Materials and Methods for details). The calculations yield values for F that range from 0 to 1, which represent in numerical form the least disposed and most disposed orientations for in-line attack, respectively. We find that a substantial distribution of

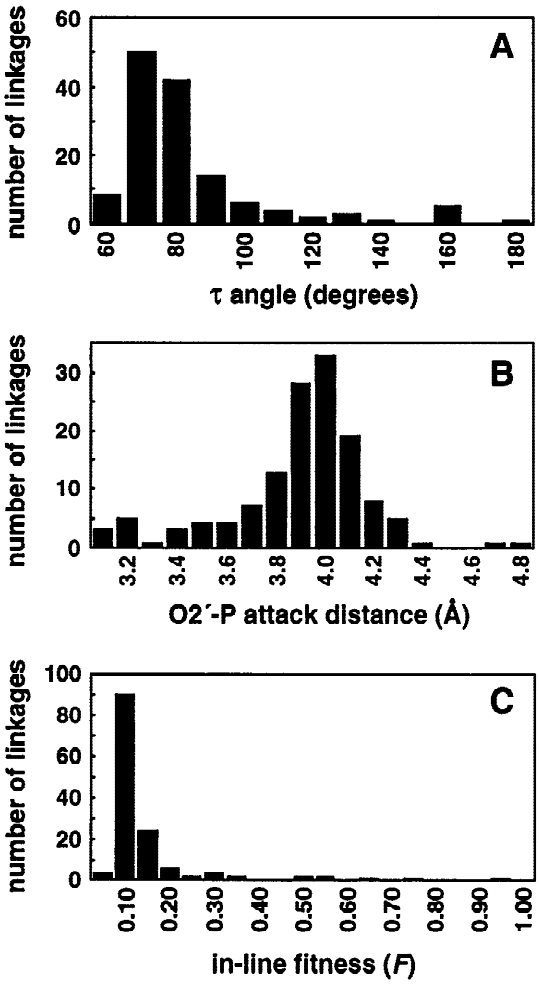

FIGURE 3. Distribution of values for τ angles (**A**), O2′-P attack distances (**B**), and in-line fitness classifications (**C**) for 136 RNA linkages examined in the context of one of several tertiary-structured RNAs. Geometric parameters for the selected RNA linkages (identified in Results) were determined using the atomic-level structures of the FMN aptamer (Fan et al., 1996), ATP aptamer (Jiang et al., 1996), P5abc (Cate et al., 1996), and the HDV ribozyme (Ferré-D'Amaré et al., 1998). The sums of the linkages that fall within each range of values specified are plotted for each parameter.

[1]Coordinates for a typical A-form helix were obtained from the MidasPlus graphics package from the Computer Graphics Laboratory, University of California, San Francisco (Ferrin et al., 1988).

[2]The RNA structural database can be accessed from the Protein Data Bank compiled at the Brookhaven National Laboratory (www.pdb.bnl.gov).

geometric parameters established by the RNA tertiary structures is present even among this relatively small sampling of RNA linkages. Therefore, we speculated that tertiary-folded RNAs with available atomic-level structural data could be used to establish the importance of in-line fitness for the cleavage of RNA by transesterification.

The structure of the FMN aptamer does not favor in-line attack

One of the RNAs used in this study is the FMN aptamer, which is an RNA sequence that was isolated by in vitro selection based on its ability to bind selectively to flavin mononucleotide (Burgstaller & Famulok, 1994). An atomic-level structure of a 35-nt version of the FMN–RNA complex has been generated using NMR (Fan et al., 1996), which confirms the presence of two stems that flank a ligand-binding pocket formed by a conserved bulge domain (Fig. 4A). Like many aptamers, this RNA exhibits adaptive binding to its corresponding ligand (Patel et al., 1997). The core of the FMN-binding RNA (nt 8–13 and 24–28) remains largely unstructured in the absence of ligand, but adopts a precise tertiary structure upon ligand binding. Because of the ligand-dependent conformational changes, the structural parameters of the core linkages as well as their cleavage activities should differ when incubated in the absence versus the presence of FMN. This ligand-induced alteration of RNA tertiary structure can be exploited as a control for our studies, because the effects of FMN on the in-line character of each linkage should be predictable.

As expected, we found that the uncatalyzed rates of RNA cleavage are dependent upon the addition of FMN to the aptamer. Incubation of a trace amount of 5′ ^{32}P-labeled aptamer RNA in the absence of FMN for 24 h gives a cleavage pattern that indicates more rapid RNA cleavage occurs within the regions that comprise the relatively unstructured ligand-binding core than in other more structured regions (Fig. 4B, lane 4, nt 8–13 and 24–28). In addition, linkages within the UUCG tetraloop that closes the hairpin structure are more susceptible to uncatalyzed RNA cleavage than are those within adjacent base-paired regions. These results are consistent with the hypothesis that helical structures preclude phosphoester transfer by preventing in-line attack, whereas linkages that reside in unstructured or heterogeneously structured regions are more susceptible to this form of cleavage. Also as predicted, cleavage within the two core regions of the aptamer is substantially suppressed when FMN is included in the incubation (Fig. 4B, lane 5). This indicates that adaptive binding of the ligand creates an RNA tertiary structure that impedes the adoption of linkage orientations with favorable in-line fitness values by favoring the tertiary structure that is predicted to form when FMN is bound.

The cleavage patterns observed with the FMN aptamer can be rationalized by examining the τ angles and the O2′-P attack distances for each linkage derived from the atomic-level structure of the RNA–ligand complex (Fig. 4C). In nearly all cases, τ angles are less than 90°, a value that is far below the optimal angle of 180° for in-line attack. Likewise, the attack distances in most cases are nearly 1 Å greater than the closest allowable approach of ~3 Å in the ground state structure. This indicates that FMN binding brings about a change in structure within the aptamer from a heterogeneously or variably structured (unbound) state that occasionally samples the in-line conformation to a homogeneous (bound) conformation that disfavors RNA cleavage.

A plot of the rate constants determined for each internucleotide linkage versus their predicted in-line fitness also reflects a relationship between linkage geometry and the rate constant (k_{obs}) for RNA cleavage (Fig. 4D). All linkages have *F* values of less than 0.3 and all display k_{obs} values that are below 10^{-6} min^{-1}. The average rate constant for an unconstrained linkage residing within a single-stranded RNA region under similar incubation conditions typically is greater than 10-fold more reactive than the linkages that comprise the ligand-bound FMN aptamer (Li & Breaker, 1999). These results are expected if the aptamer-ligand complex forms a stable structure that maintains RNA linkages in a non-in-line conformation. Therefore, the geometric parameters derived from the atomic-level structure are consistent with the observed dampening of RNA cleavage rates in the aptamer-ligand complex.

The ATP aptamer has a linkage with near perfect in-line fitness

We used an identical approach to examine the ATP aptamer that was isolated by Sassanfar and Szostak (1993). Like the FMN aptamer, this RNA undergoes adaptive binding when presented with ATP or with related analogs such as adenosine. In addition, two groups (Dieckmann et al., 1996; Jiang et al., 1996) have published atomic-level structural models for the aptamer–ligand complex based on NMR analysis. The two structural models are similar within the ligand-binding core, and both are consistent with a structure wherein ATP binds within a central bulge that is flanked by two helical domains (Fig. 5A).

As observed with the FMN aptamer, the uncatalyzed rates of RNA transesterification within the ATP aptamer are dependent on ligand binding. In the absence of ATP, the aptamer shows a general susceptibility to cleavage that is confined mostly to nt 6–17 (Fig. 5B, lane 4), which are positions that correspond to the relatively unstructured bulge region. The two stems are known to form even in the absence of ligand (Nonin et al., 1997),

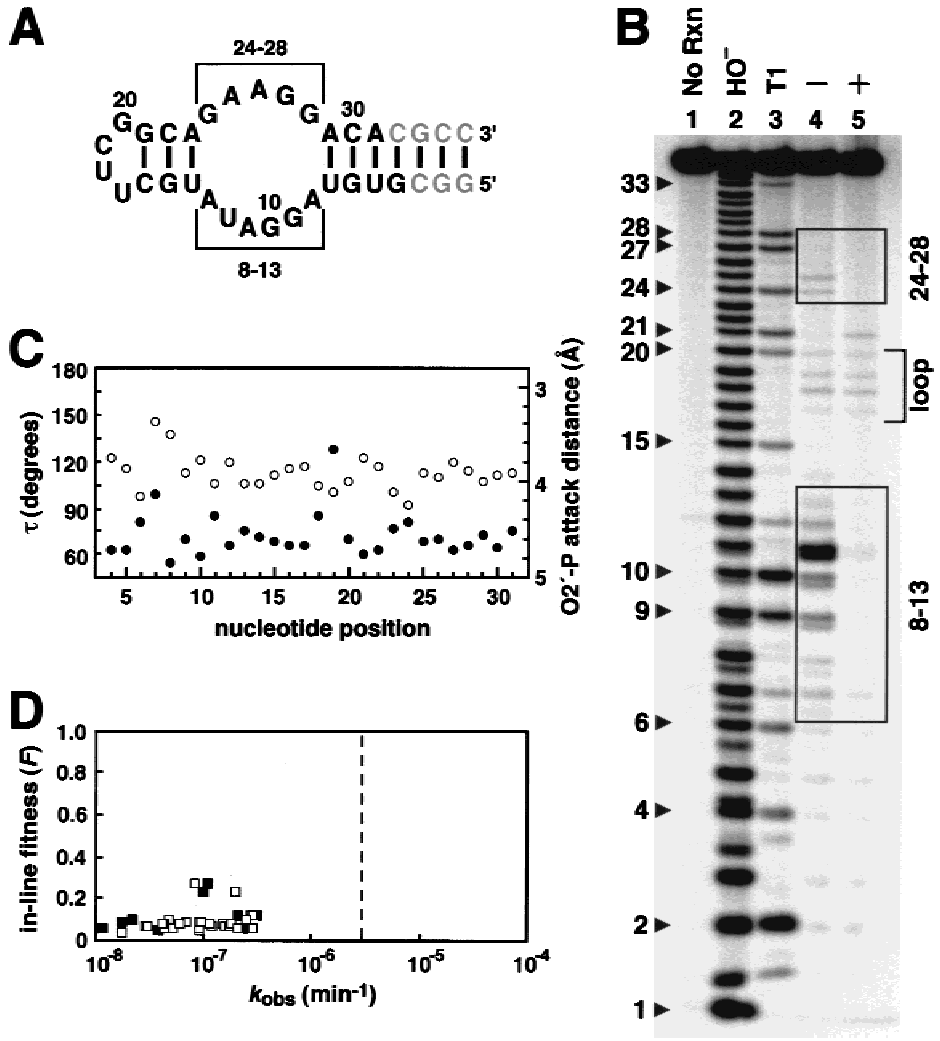

FIGURE 4. Uncatalyzed cleavage of the FMN aptamer by intramolecular phosphoester transfer. **A**: Sequence and secondary structure of the FMN aptamer. Internucleotide linkages located immediately 3′ relative to the nucleotide positions indicated by bold type were examined using structural and biochemical data to compare internucleotide geometry with the observed rate of phosphodiester cleavage. Brackets identify the two regions that comprise the ligand-binding bulge. Nucleotide numbering system used is that described by Fan et al. (1996). **B**: Analysis of ligand-dependent phosphodiester cleavage of the FMN aptamer. 5′ ^{32}P-labeled RNA was incubated in the absence (−, lane 4) or presence (+, lane 5) of FMN in reaction buffer and separated on a denaturing 15% polyacrylamide gel. Also shown are unreacted RNA (lane 1) and RNA ladders generated by hydroxide cleavage (lane 2) and RNase T1 (G-specific cleavage; lane 3). Boxes identify the two core regions (8–13 and 24–28) that display ligand-modulated cleavage patterns. Cleavage within the loop region (bracket; nt 17–20) is not affected by FMN binding. Doublet banding patterns apparent at positions 1–11 are due to resolution of the 2′,3′-cyclic phosphate termini (upper bands) and 2′(3′)-phosphate termini (lower bands) present on otherwise identical RNA cleavage products. Numbered arrowheads identify the bands corresponding to cleavage 3′ relative to guanosine residues as indicated. **C**: Geometric parameters determined for internucleotide linkages within the FMN aptamer complex. Shown are the τ angle (open circles) and O2′-P attack distance (filled circles) for each linkage located 3′ to indicated nucleotides within model 1 of the published atomic coordinates (Fan et al., 1996). The most favorable linkages for in-line attack are made evident when both data points for a given position reside near the top of the plot. In this case, all linkages are predicted to form a geometry that disfavors in-line attack. Note that nt 1–3 and 32–35 were not included in the analysis because structural heterogeneity due to helix fraying as well as interference from the precursor band complicate data analysis at these sites. **D**: Correlation between in-line fitness (F) and the rate constant (k_{obs}) for phosphodiester cleavage within the FMN aptamer in the presence of ligand. Open and filled squares represent data derived from two independent experiments. The dashed line indicates the average observed rate constant for cleavage of a phosphodiester bond within an unconstrained linkage as determined from nucleotides composing the relatively unstructured loop region (G6 through G17) of the ATP aptamer in the absence of ligand (see Fig. 5). This rate constant also is consistent with the rate constant predicted for an unconstrained linkage under similar reaction conditions (Li & Breaker, 1999).

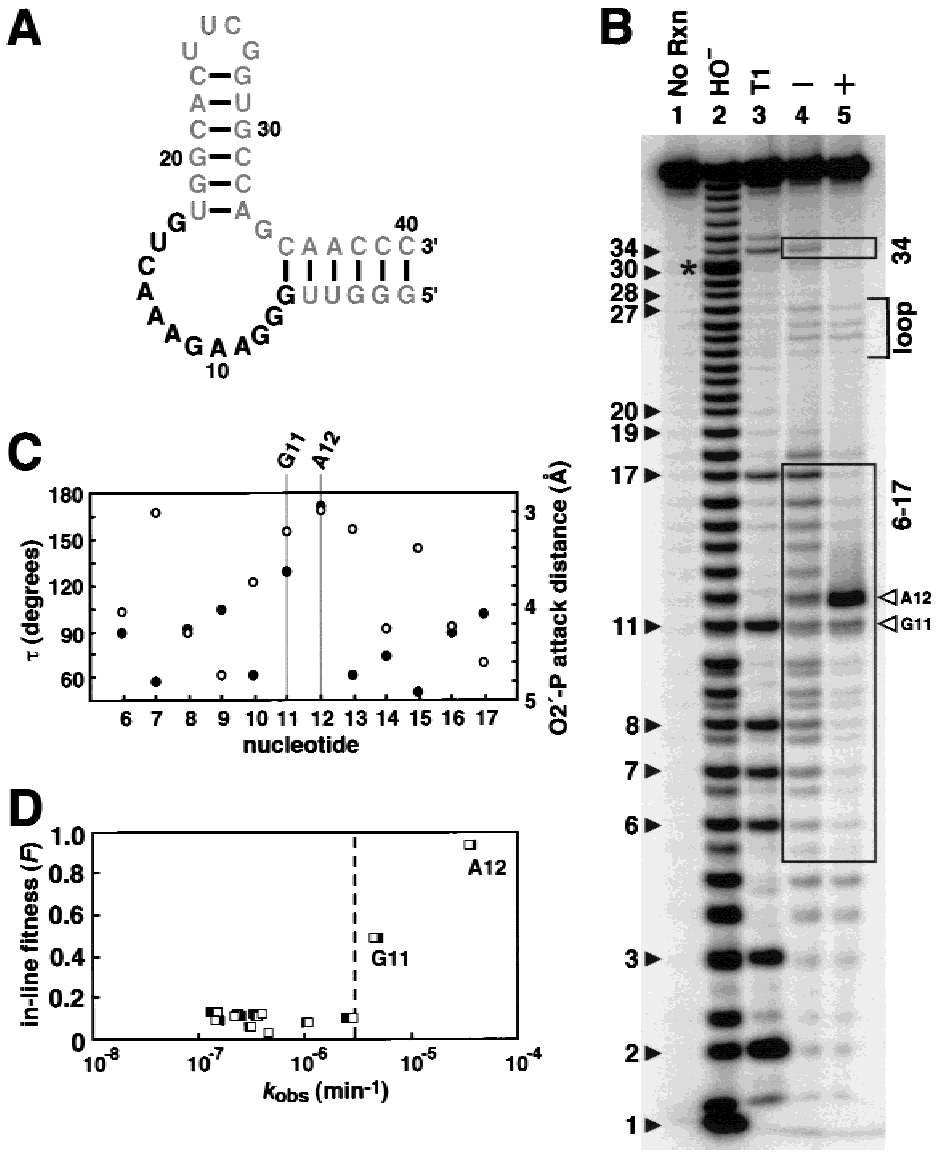

FIGURE 5. Analysis of the ATP aptamer complex. **A**: Sequence and secondary structure of the ATP aptamer. Internucleotide linkages located immediately 3′ relative to the nucleotide positions indicated by bold type were examined in this study. The nucleotide numbering system used is that described by Jiang et al. (1996). **B**: Analysis of ligand-dependent phosphodiester cleavage of the ATP aptamer. Analysis of 5′ ^{32}P-labeled RNA was conducted as described for Figure 4B using ATP as the ligand. Boxes identify bands corresponding to the bulge nucleotides (6–17 and 34) in the secondary structure of the ATP aptamer. Open arrowheads designate the ATP-dependent cleavages at nucleotides G11 and A12. The asterisk designates the position of compressions that prevent resolution of individual bands. Other annotations are as described for Figure 4B. **C**: Geometric parameters determined for internucleotide linkages within the ATP aptamer complex. Shown are the parameters determined for linkages within model 7 of the published atomic coordinates (Jiang et al., 1996). Linkages identified by the shaded lines (G11 and A12) are predicted to conform to an in-line geometry when bound to ATP. **D**: Correlation between *F* and the k_{obs} for phosphodiester cleavage of the designated linkages within the ATP aptamer in the presence of ligand. Details of the plot are presented in the legend to Figure 4D.

and this explains the lower level of spontaneous RNA cleavage in these regions. In contrast, cleavage is significantly diminished at most sites within the main bulge region and at a single bulge position (nt 34) when the RNA is incubated in the presence of ATP (Fig. 5B, lane 5). These results are consistent with a ligand-induced conformational change within the bulge that largely precludes the adoption of in-line conformational states. Cleavage of internucleotide linkages of the loop spanning nt 24–27 remains unaffected by the addition of ATP, indicating that this distal structural element is not affected by ligand binding.

In contrast, two sites (G11 and A12) within the ATP-binding core show significant levels of strand scission in the presence of ATP. The extent of RNA cleavage at position G11 in the absence and presence of ATP largely remains unchanged. This finding indicates that ligand binding either does not restrict conformational freedom of this linkage, or that ligand binding establishes a near in-line attack conformation that provides an equivalent rate enhancement. Furthermore, the cleavage activity at position A12 is substantially increased upon the addition of ATP, indicating that ligand binding brings about a significant increase in the in-line fitness. Linkages within the ATP-binding bulge present a wide range of τ angles and attack distances (Fig. 5C), but these two parameters both become favorable for RNA cleavage only at G11 and A12 upon ligand binding. This more detailed examination of the geometric parameters for the ATP aptamer linkages reveals that these two sites indeed are uniquely positioned for in-line attack, whereas all other linkages are predicted to be held in orientations that are unfavorable for intramolecular transesterification.

A plot of k_{obs} versus F (Fig. 5D) reveals that increasing in-line fitness correlates with increasing susceptibility to spontaneous RNA transesterification. Most linkages have F values below 0.2 and display rate constants of less than 10^{-6} min^{-1}, as was observed with the FMN aptamer (Fig. 4D). In comparison, positions G11 ($F = 0.50$) and A12 ($F = 0.95$) exceed the typical rate of cleavage for an unconstrained RNA linkage. A12 is of particular interest because this linkage is predicted to be most favorably positioned for in-line attack. This site cleaves ~12-fold faster ($k_{obs} = 3.6 \times 10^{-5}$ min^{-1}) than a typical unconstrained linkage (average $k_{obs} = 3.0 \times 10^{-6}$ min^{-1}), and cleaves more than 100-fold faster than most linkages that are constrained in an orientation that is unfavorable for in-line attack.

Examination of an independently folding domain of a large ribozyme

Investigating the importance of in-line fitness in RNA transesterification using the experimental approach described above can be fully implemented only when adequate structural data are available. Atomic-level structural data are available for other aptamers that bind theophylline (Zimmermann et al., 1997), tobramycin (Jiang et al., 1997) and citrulline or arginine (Yang et al., 1996). In addition, structures have been reported for hairpin loops such as UUGA (Butcher et al., 1997), UUCG (Varani et al., 1991; Dieckmann et al., 1996), and GAAA (Jucker et al., 1996). However, all the linkages within these structures are predicted by our methods to have low F values, and therefore are not expected to display significant transesterification activity. In accordance with our hypothesis, the theophylline aptamer was found to display diminished cleavage activity in the presence of its corresponding ligand relative to its absence (data not shown).

To further expand our data set, we have examined the atomic-level structures of several ribozymes and ribozyme fragments that have been reported recently. For example, P5abc RNA (Fig. 6A) is an independently folding fragment of the *Tetrahymena* group I ribozyme (Murphy & Cech, 1993). This RNA displays multiple cleavage sites that differ in the absence and presence of Mg^{2+} (Fig. 6B). This finding is consistent with previous data (Cate & Doudna, 1996) indicating that metal-binding sites are formed by P5abc, and that significant structural changes occur when divalent metals are presented (Celander & Cech, 1991; Wu & Tinoco, 1998). Like many RNA aptamers, the ribozyme fragment undergoes adaptive binding of divalent metal ligands, which is expected to alter the in-line character of individual RNA linkages.

In-line fitness values within the P5abc domain have been quantified using the atomic coordinates from the crystal structure of the larger P4–P6 domain of the group I ribozyme that was determined by Cate et al. (1996). Two sites in particular (C138 and G188) simultaneously display values for τ angles and attack distances that are favorable for in-line attack (Fig. 6C). Interestingly, the linkage at C138 is one of the most labile sites in the absence of divalent metals, but cleavage is not accelerated when Mg^{2+} is present (Fig. 6D). In contrast, the linkage at G188 is not distinctively labile in the absence of divalent metals, but cleavage accelerates in the presence of Mg^{2+}. Only the latter result is consistent with our hypothesis, and this point is examined in more detail in the Discussion section.

The RNA is also susceptible to cleavage at several sites that are predicted by linkage geometry to disfavor in-line attack (Fig. 6B, lane 5, loop, *i*, and *ii*). Linkages within the loop region defined by nt 168–172 of P5abc cleave with rate constants that are expected for an unconstrained single-stranded RNA. Similar cleavage effects are observed with both the FMN and ATP aptamers (Figs. 4A and 5A). The two bands identified by *i* in Figure 6B represent cleavage at U185 and U179. These linkages are formed by the dinucleotides UpA and UpG, the former of which has been shown to undergo greater spontaneous cleavage in certain cases when presented in single-stranded regions (Kierzek, 1992; Williams et al., 1995). The bands identified by *ii* in Figure 6B result from RNA cleavage at nucleotides (139, 140, 163, and 164) that are involved in the formation of noncanonical G•A base pairs at the terminus of stem P5b. These weaker pairing interactions should allow occasional sampling of linkage conformations that favor in-line attack, and are expected to cleave with higher frequency. Therefore, a rationale exists for accelerated cleavage at these particular sites that are otherwise predicted to be stabilized by the Mg^{2+}-induced RNA structure.

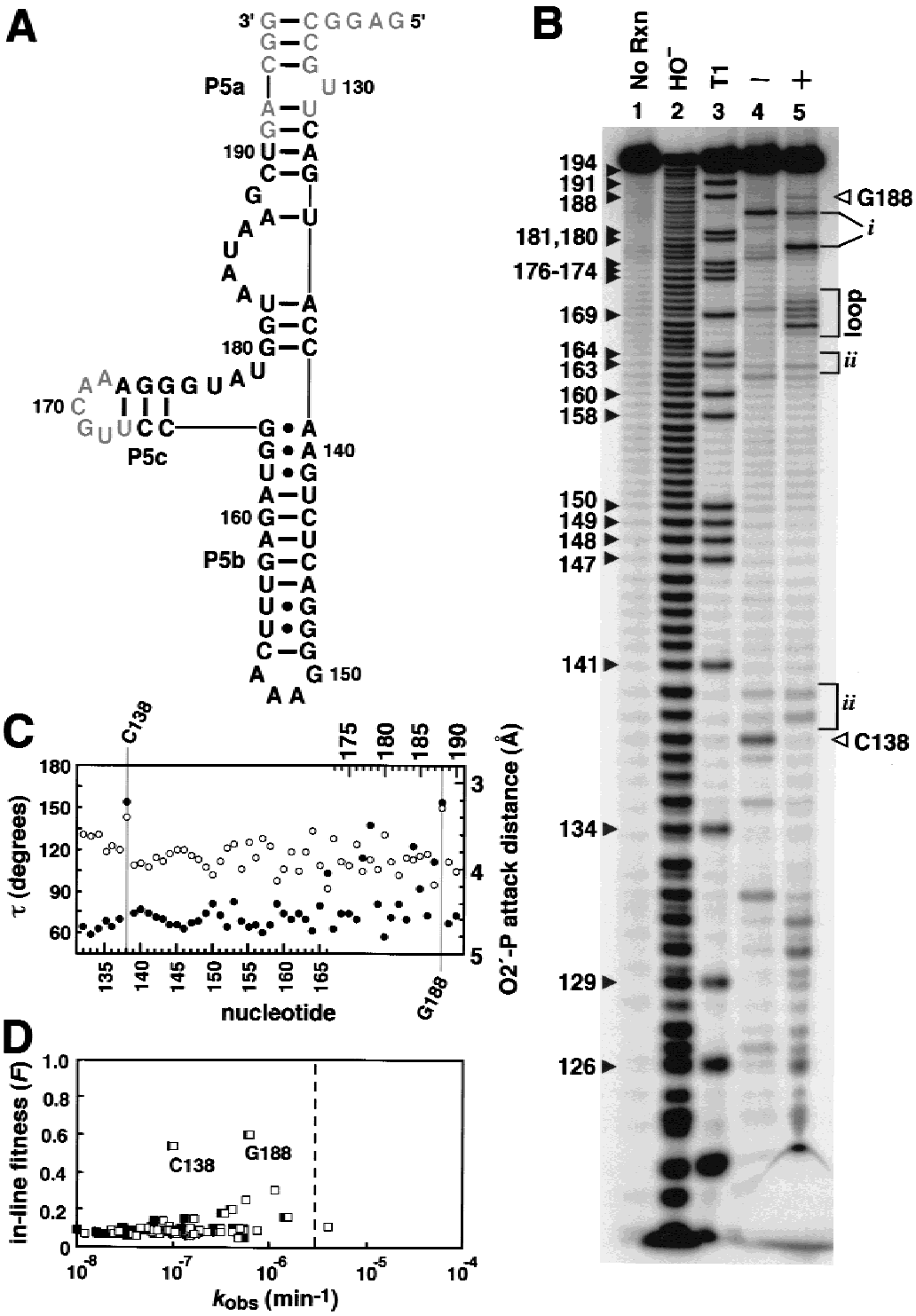

FIGURE 6. Analysis of P5abc RNA. **A**: Sequence and secondary structure of P5abc RNA. Internucleotide linkages located immediately 3′ relative to the nucleotide positions indicated by bold type were examined in this study. Nucleotide numbering system for P5abc corresponds to that used by Cate et al. (1996). **B**: Analysis of Mg^{2+}-dependent phosphodiester cleavage in P5abc RNA. 5′ ^{32}P-labeled RNA was incubated in the absence (−, lane 4) or presence (+, lane 5) of 20 mM Mg^{2+} in reaction buffer and separated on a 10% denaturing polyacrylamide gel. Lanes 1–3 were generated as described for Figure 4B. Regions of increased RNA cleavage that are not predicted by the in-line fitness values are denoted *i* and *ii*. Other annotations are as described for Figure 4B. **C**: Geometrical parameters determined for dinucleotide linkages within P5abc RNA. Shown are the parameters determined for linkages within chain A of the published atomic coordinates (Cate et al., 1996). Linkages identified by the shaded lines (C138 and G188) are predicted to conform to in-line geometry in the presence of Mg^{2+}. **D**: Correlation between *F* and the k_{obs} for phosphodiester cleavage of the designated linkages within P5abc RNA in the presence of Mg^{2+}. Details of the plot are presented in the legend to Figure 4D.

HDV ribozyme contains three linkages with in-line conformation

In addition to the P4–P6 structural model, the crystal structure of the hammerhead ribozyme (e.g., Scott et al., 1995) and the crystal structure of the self-cleaving ribozyme from the hepatitis delta virus (HDV; Ferré-D'Amaré et al., 1998) are available. The hammerhead ribozyme does not include any internucleotide linkages that are predicted to adopt an in-line conformation. However, the HDV RNA (Fig. 7A) structure contains several linkages that are favorably positioned for nucleophilic

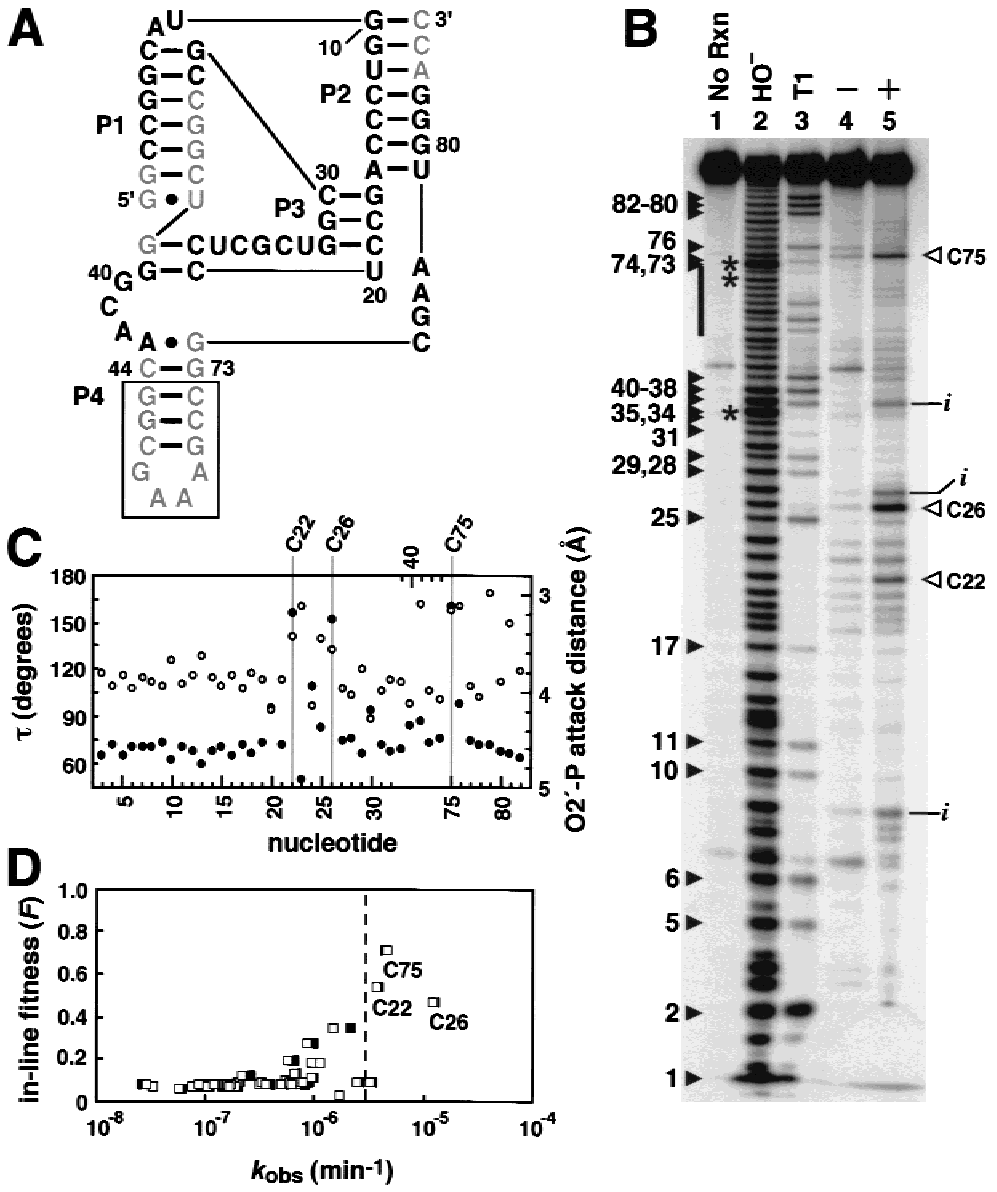

FIGURE 7. Analysis of HDV RNA. **A**: Sequence and secondary structure of HDV RNA. Internucleotide linkages located immediately 3′ relative to the nucleotide positions indicated by bold type were examined in this study. Nucleotide numbering system for HDV corresponds to that used by Ferré-D'Amaré et al. (1998) with exceptions for the replacement of stem-loop P4. This latter domain (boxed nucleotides) remains unnumbered. **B**: Analysis of Mg^{2+}-dependent phosphodiester cleavage in HDV RNA. End-labeled RNA was incubated in the absence (−, lane 4) or presence (+, lane 5) of 20 mM Mg^{2+} in reaction buffer and separated on a 10% denaturing polyacrylamide gel. Lanes 1–3 were generated as described for Figure 4B. Asterisks designate the positions of compressions that prevent resolution of individual bands. The vertical bar encompasses bands that correspond to the boxed nucleotides in **A**. Other annotations are as described for Figure 4B. **C**: Geometrical parameters determined for dinucleotide linkages within HDV RNA. Shown are the parameters determined from the published atomic coordinates (Ferré-D'Amaré et al., 1998). Linkages identified by the shaded lines (C22, C26, and C75) are predicted to conform to in-line geometry in the presence of Mg^{2+}. **D**: Correlation between F and k_{obs} for phosphodiester cleavage of the designated linkages within HDV RNA in the presence of Mg^{2+}. Details of the plot are presented in the legend to Figure 4D.

attack. The RNA cleavage patterns of the HDV RNA differ in the absence and presence of Mg^{2+}, indicating that conformational changes occur upon metal binding (Fig. 7B). This finding is in accordance with observations that the catalytic activity of HDV is dependent on divalent metal ions (Suh et al., 1993). Three of the sites that display accelerated cleavage (C22, C26, and C75) correspond to the three linkages that are modeled to be preorganized for in-line attack (Fig. 7C).

A plot of k_{obs} versus F (Fig. 7D) for this RNA again reveals that increasing in-line fitness correlates with increasing susceptibility to uncatalyzed RNA transesterification. However, several sites predicted to have F values less than 0.2 cleave with rate constants that approximate those measured for unconstrained linkages. Like similar linkages in P5abc RNA, two of these more reactive sites (Fig. 7B, lane 5, *i*) again correspond to purine–pyrimidine dinucleotides that are known to cleave preferentially (Kierzek, 1992; Williams et al., 1995). Therefore, in most instances the positions of accelerated cleavage in this RNA and in the P5abc RNA can be ascribed to structure-induced in-line orientation, unconstrained structural exploration, or previously observed sequence preferences for spontaneous phosphodiester cleavage.

DISCUSSION

In-line orientation as a determinant of RNA cleavage rates

Some single stranded RNAs are known to adopt random-coil states that simulate the helical structures of duplex RNA and DNA (Saenger, 1984). However, these relatively weak structures are largely free to sample many possible linkage conformations in a dynamic fashion and occasionally acquire an in-line orientation and a proximal relationship between a reactive 2′ oxygen and its adjacent phosphorus group. As a result, single stranded RNAs are more prone than duplex RNAs to uncatalyzed and metal-catalyzed cleavage by an intramolecular phosphoester transfer mechanism (Ciesiolka et al., 1992; Reynolds et al., 1996; Welch et al., 1997). Linkages near the ends of RNA helices are also more susceptible to cleavage (Usher & McHale, 1976), presumably due to fraying (Jin et al., 1993; Puglisi et al., 1994) of the stabilizing helical structure. In contrast, folded RNAs of greater complexity adopt secondary and tertiary structures that present a variety of distinct linkage geometries. The conformations either preclude or favor RNA cleavage via an intramolecular transesterification mechanism depending on their relative in-line character. These observations are consistent with the hypothesis that in-line character influences the rates of RNA cleavage as implied earlier (Usher & McHale, 1976).

Among the 136 internucleotide linkages examined in this study, we find that the majority have a geometry ($F < 0.2$) that disfavors S_N2 nucleophilic attack via an in-line mechanism. Correspondingly, most of these sites cleave between 1/10th and 1/1,000th the rate of a typical unconstrained linkage (Fig. 8). Only five of these 122 linkages cleave with rate constants that approximate an unstructured RNA. In three instances, these rare sites with unpredictably high cleavage rates have been identified as unusually predisposed to spontaneous scission. Taken together, these findings are consistent with the hypothesis that in-line orientation is necessary for transesterification, and that common folding elements in RNA (e.g., A-form helices) preclude the adoption of this critical geometric conformation.

Seven sites among the combined data points are predicted by the structural models to approach a near in-line configuration with F greater than 0.4, while seven

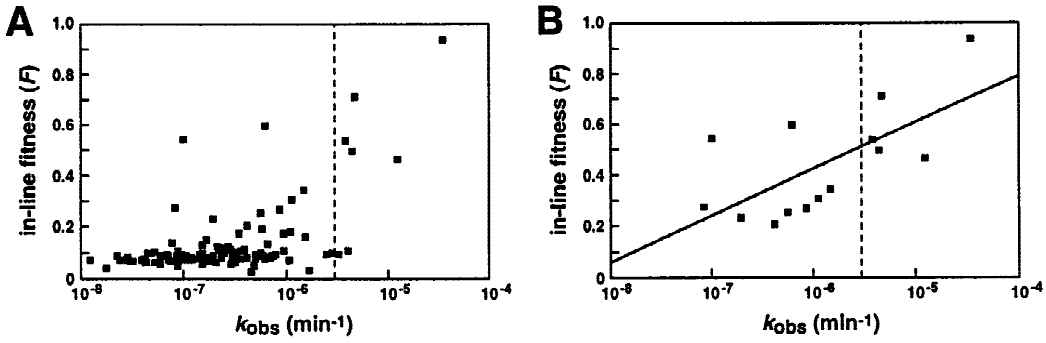

FIGURE 8. Composite plots depicting the relationship between in-line fitness and rate constants for RNA transesterification for 136 linkages distributed among four different RNA structures. **A:** The k_{obs} for cleavage at each phosphodiester linkage is shown relative to an in-line fitness value determined for each corresponding linkage of the four RNA structures examined in this study. The dashed line indicates the average observed rate constant for cleavage of a phosphodiester bond within an unrestrained dinucleotide linkage as described in the legend to Figure 4. **B**: Similar plot depicting the correlation between in-line fitness and the rate constant for RNA transesterification (slope = 0.18, $R = 0.65$) using data points with F values greater than 0.2.

additional linkages have an *F* value between 0.2 and 0.4. Only two sites within the first group fail to cleave with a rate constant that exceeds the typical value observed for an unconstrained linkage (Fig. 6). A plot of the 14 linkages with *F* values greater than 0.2 reveals a modest trend between in-line fitness and the rate of spontaneous RNA transesterification (Fig. 8B). It is important to note that several experimental parameters discussed in more detail below may impact the accuracy of these plots. However, with consideration for these issues, we conclude that the general relationship observed in this study is indicative of the importance of in-line orientation to RNA cleavage proceeding via a cyclizing mechanism.

Accuracy of the kinetic and structural data

Establishing an accurate value for the rate enhancement that can be derived from in-line positioning is of significant importance for developing a thorough understanding of RNA cleavage mechanisms and the mode of action of RNA-cleaving enzymes. Unfortunately, the precision of the data points that are used to generate the plots depicted in Figure 8 can be adversely affected by a variety of factors. First, measurements of the uncatalyzed cleavage of RNA must incorporate extended reaction times in a mixture that is free from contaminating nucleases. All materials and procedures employed in this study have been implemented with consideration for the prevention of nuclease contamination. Second, the cleavage products separated on high-resolution polyacrylamide gels must be quantitated with exactness to assure the accuracy of the rate constants. We have observed that replicate experiments with individual RNAs give highly reproducible results, indicating that the kinetic data reported herein are accurate. Third, the presence of RNA structural heterogeneity when a single defined structure is expected could yield misleading cleavage patterns and reaction kinetics. As a result, we have chosen to examine RNA motifs that are believed to form a single structured state in the presence of their small organic effector or metal-ion ligands.

We have not included an assessment of solvent accessibility in this study, given the small size of the RNAs examined. Accessibility at each internucleotide linkage is critical because deprotonation of the 2′-hydroxyl group by solvent to generate the more nucleophilic 2′ oxyanion is likely to be a major determinant of RNA transesterification rates (Li & Breaker, 1999). Chemical accessibility of phosphodiester linkages is known to be a factor for large RNA structures. For example, Celander and Cech (1991) demonstrated that interior linkages of a folded group I ribozyme are not accessible to cleavage by free radicals generated using Fe(II)-EDTA. Interestingly, position C138 of the P5abc RNA cleaves in the absence of Mg^{2+}, but fails to cleave as predicted when the divalent metal is present (Fig. 6). The cleavage patterns in this region indicate that the local RNA structure may be identical in the absence and the presence of Mg^{2+}, as evident by the fact that accelerated cleavage appears at the G•A base pairs in Figure 6B, lanes 4 and 5. Therefore, the robust cleavage observed at C138 in the absence of Mg^{2+} may indeed be due to a stable conformation that favors in-line attack. However, the C138 linkage resides immediately adjacent to a metal-binding site (Cate & Doudna, 1996), and metal binding to this site may contribute to the dampening of the RNA cleavage rate when Mg^{2+} is added. Rate reduction could be due to metal-induced steric exclusion of solvent or perhaps a shift in the pK_a of the 2′-hydroxyl group. However, the specific reasons why our predictive method fails to more accurately project the cleavage kinetics of this linkage remain unclear.

A significant source of complexity in the interpretation of the data is derived from the observation of enhanced cleavage at sites that are predicted to adopt a non-in-line orientation. We have identified several sites within the four RNAs examined in this study that cleave with unexpectedly high rates of cleavage. In most cases, however, increased cleavage can be ascribed to unstable structures (G•A base pairs; Fig. 6) or the presence of certain pyrimidine/purine dinucleotides that have been shown to favor spontaneous cleavage (Figs. 6 and 7). Each of these unexpected sites may form dynamic structures that cleave via an in-line mechanism; however, they may not be easy to predict by examining a static model of an RNA structure. The presence of these unpredictable cleavage sites casts doubt over any conclusions about sites that are predicted to show and indeed do display enhanced cleavage rates. Although we have no reason to believe that any of the 14 sites depicted in Figure 8B are spurious, it may be difficult in general to distinguish predicted in-line cleavage from enhanced cleavage due to another mechanism of predisposition.

Perhaps the most important component of the analysis used in this study is the resolution of the NMR and X-ray structural data. The accuracy of the in-line fitness values generated for each linkage is highly dependent on the correct placement of atoms in the RNA structural models. Misplacement of atoms even by sub-angstrom distances can significantly affect the calculated in-line fitness value for a linkage. As this is the primary predictive parameter for transesterification rate constants, the resulting error could be substantial. Unfortunately, the resolution of the structural models can be affected by a number of factors. For example, molecular perturbations that may result from the physical processes of structure determination may result in an incorrect prediction of the in-line fitness values for the RNA in solution. Structural alterations due to crystal-packing forces would be one possible source of error. In addi-

tion, the NMR methods used to generate the models for the FMN and ATP aptamers do not directly visualize the location of the phosphorus centers. The location of this critical atomic center is determined by the energy minimizing algorithms used to establish structural models from a given set of NMR distance constraints.

The precision of the NMR model-generating process can be assessed by examining the various τ angles and O2′-P attack distances for the FMN and ATP aptamers. Five related models for the FMN aptamer have been generated using the same set of NMR distance constraints (Fan et al., 1996). Considerable variability exists in the internucleotide geometries between different models (Fig. 9A). Variability, of course, will affect the accuracy of our in-line fitness projections, but in this case the RNA linkages do not generally maintain favorable attack angles or attack distances. This is consistent with the observed chemical stability of the FMN aptamer in the presence of ligand (Fig. 4). Likewise, the internucleotide geometries for the ATP aptamer vary between the eight structural models (Jiang et al., 1996) generated using NMR distance constraints (Fig. 9B). Relative heterogeneity of linkage parameters is observed for most linkages, with the exception of position G11. Here, all eight models consistently predict an orientation that favors in-line attack and this prediction is consistent with the enhanced cleavage rate observed at this position. A more complicated situation exists for A12, the other site of enhanced cleavage. In this case, seven of the eight models predict a linkage configuration that is not expected to favor in-line attack. However, model 7 predicts that the A12 linkage is held with near perfect in-line geometry. The main difference between model 7 and the remaining structures is the position of nucleotides A12 and A13, which are rotated toward the core of the ATP-binding bulge as opposed to facing outward. Based on the cleavage data depicted in Figure 5, we suggest that the ATP aptamer may favor a conformation most similar to model 7, or at least may sample this alternate structure on occasion such that in-line nucleophilic attack at position A12 is more frequent.

The influence of metal ions on RNA transesterification

Cationic forms of the alkali earth metals promote the nonspecific degradation of RNA via the transesterification mechanism depicted in Figure 1. The role of metal ions in this process is generally believed to be that of a general base catalyst involving a metal-hydrate complex, or that of a Lewis acid catalyst involving the direct coordination of the metal ion to one or more oxygen atoms that comprise the labile RNA linkage (Cech & Golden, 1999). Metal ions also may serve a structural role by stabilizing intricately folded RNAs (Pyle & Green, 1995; Draper, 1996). These characteristics are believed to be important for the catalytic function of most ribozymes that function as metalloenzymes (Pyle, 1993, 1996).

As discussed above, metal ions can directly accelerate RNA transesterification by functioning as a specific base catalyst or as a Lewis acid catalyst. In fact, the precise roles that metal ions perform in some ribozyme-

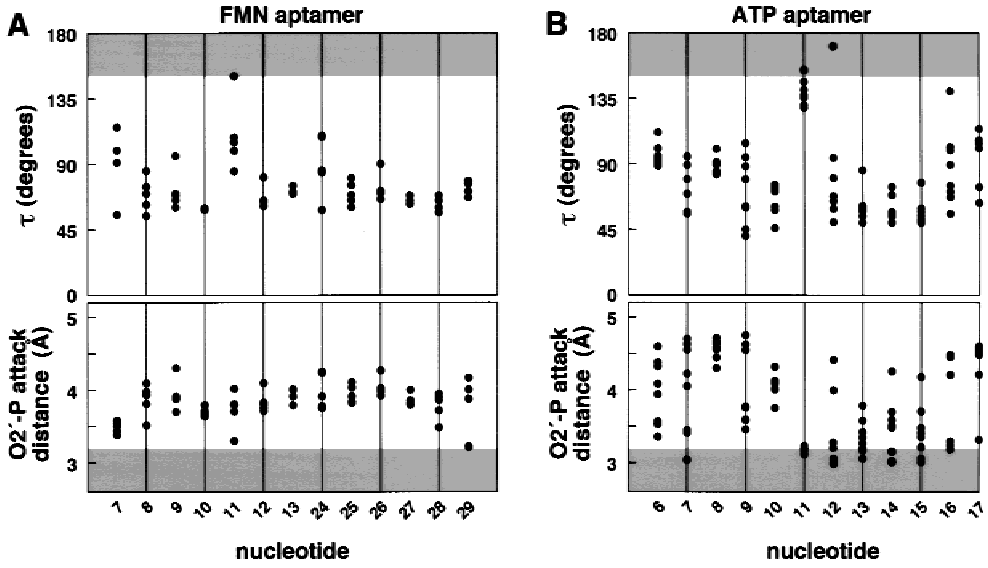

FIGURE 9. Relative uniformity of the structural data for the (**A**) FMN and (**B**) ATP aptamers. Values for τ and the attack distances are plotted for five structural models for the FMN aptamer (Fan et al., 1996) and eight structural models for the ATP aptamer (Jiang et al., 1996) that were derived using NMR constraints. Numbering systems for the FMN and ATP aptamers are according to Figures 4A and 5A, respectively. For the ATP aptamer, nucleotides 11 and 12 correspond to linkages G11 and A12, respectively. Note that some of the data points plotted are overlapping due to similar values.

mediated RNA cleavages are still a point of fervent debate (Kuimelis & McLaughlin, 1998; Zhou & Taira, 1998). The general perception of the catalytic roles of divalent metal cations largely has been shaped by the interaction of Pb^{2+} with $tRNA^{Phe}$ of yeast. RNA is particularly susceptible to nonspecific cleavage by Pb^{2+} ions, though the cleavage preferentially occurs in single-stranded regions when low concentrations of metal ions are used (Ciesiolka et al., 1992; Welch et al., 1997; Zagórowska et al., 1998). The tertiary structure of yeast $tRNA^{Phe}$ forms several binding sites for Pb^{2+} (Brown et al., 1983), and the occupation of the "Pb 1" binding site by lead ion induces site-specific RNA cleavage at a proximal phosphodiester linkage (between nt 17 and 18). Based on the crystal structure of the Pb^{2+}-$tRNA^{Phe}$ complex, this metal ion is believed to interact with the 2′-hydroxyl group of nt 17 via an intervening hydroxide ligand, which activates the nucleophile for attack on the adjacent phosphorus center. Hydrated Pb^{2+} ions are well suited to perform the role of a specific base catalyst because the pK_a of the water ligand is reduced ~7 units to near neutral pH (Feig & Uhlenbeck, 1999). Similarly, other metal–water complexes with lowered pK_a values could promote specific base catalysis of RNA transesterification. However, the labile linkage in the crystal structure of $tRNA^{Phe}$ is not well positioned for in-line attack ($F = 0.06$). Presumably relatively minor conformational adjustments in the linkage coupled with the catalytic power of Pb^{2+}-hydroxide are sufficient to generate the modest rate enhancements observed in this system. The importance of specific base catalysis of RNA cleavage by metal hydrates is further supported by the crystal structure of the leadzyme (Wedekind & McKay, 1999), a small Pb^{2+}-dependent ribozyme that was isolated using in vitro selection (Pan & Uhlenbeck, 1992).

With analogy to this metal-hydroxide cleavage mechanism, several groups have examined putative metal binding sites in RNAs by conducting RNA cleavage experiments similar to those described in this article (e.g., Kazakov & Altman, 1991; Polacek & Barta, 1998). In these metal-ion-probing studies, RNA cleavage patterns are examined in the absence and in the presence of various divalent metal ions. New cleavage sites that appear only when metals are added are attributed to the catalytic action of the ion when bound to a specific divalent cation binding site. We find that both the P5abc (Fig. 6) and the HDV (Fig. 7) RNAs show differential cleavage in the absence and presence of 20 mM Mg^{2+}. However, the extent of cleavage in the absence of Mg^{2+} compared to the cleavage in the presence of Mg^{2+} is nearly the same in each case, suggesting that metal ions are generally not necessary for the appearance of significant levels of localized spontaneous RNA cleavage. Moreover, not all cleavage sites observed in P5abc correspond with metal-binding sites that were determined by X-ray analysis of the crystallized P4–P6 domain (Cate & Doudna, 1996). Therefore, we propose that metal ions indeed associate with distinct pockets within each of these RNAs, but that local binding or even binding at a great distance from the site of cleavage simply may alter the tertiary structure and generate linkage conformations that favor in-line attack. In other words, metal ions may not be directly involved as general base catalysts or as Lewis acid catalysts, but may only serve to stabilize specific conformations that are inherently more susceptible to spontaneous cleavage. Data from metal-ion-probing studies (Kazakov & Altman, 1991; Polacek & Barta, 1998) do not provide conclusive evidence that metal-ion-binding pockets reside near spontaneous RNA cleavage sites and thus should be interpreted with caution.

The value of in-line fitness to the rate enhancement of RNA transesterification

The hypothesis that in-line attack is important for RNA cleavage reactions that proceed via an S_N2-like mechanism has been extensively discussed in the literature. However, little quantitative data exist concerning RNA cleavage rate enhancements that can be derived exclusively by positioning the reactive groups for in-line attack. Usher and McHale (1976) reported a ~900-fold rate enhancement for cleavage of a 2′,5′-phosphodiester linkage compared to the rate constant observed for the cleavage of the analogous 3′,5′-phosphodiester linkage when both are positioned within the polyadenylated strand of a U•A•U RNA triplex. When the linkages are unconstrained by helix formation, they both cleave at the same rate, indicating that the inherent chemical stabilities of the two isomers are identical. The rate differential is established by the helical structure, which positions the 2′,5′ linkage in an in-line conformation and positions the 3′,5′ linkage in a non-in-line conformation. Compared to the rates observed for the unconstrained 2′,5′ linkage, the helix accelerates RNA transesterification of the linkage by a modest 18-fold.

Also using unconstrained linkages as a reference point (average $k_{obs} = 3.0 \times 10^{-6}$ min^{-1}), we similarly observed that cleavage rates can be reduced at linkages that adopt a non-in-line conformation, whereas cleavage rate enhancements can be observed for several linkages with established in-line geometry. The maximum rate constant for cleavage ($k_{obs} = 3.6 \times 10^{-5}$ min^{-1}) was observed at position A12 of the ATP aptamer (Fig. 5). Of the 136 linkages examined, the A12 site is predicted to most closely approach perfect in-line geometry. Therefore, we conclude from our data that the rate enhancement that can be derived by organizing an RNA linkage for in-line attack is at least 12-fold. Interestingly, the rate enhancement that occurs at A12 corresponds with the value observed for transesterification of the 2′,5′ linkage when similarly posi-

tioned for in-line attack as discussed above (Usher & McHale, 1976). However, given that the conformations of three torsion angles influence the in-line fitness of an RNA linkage, we expect that the true maximum rate enhancement might be somewhat larger. Conceivably, one problem with the approach used in this study is the possibility that folded RNA structures that occur naturally or that are isolated by in vitro selection may avoid perfect in-line geometries. If the maximum rate enhancement is larger than 20-fold, then RNAs that contain linkages with perfect in-line orientation may be unstable to the extent that it becomes evolutionarily disadvantageous.

The importance of in-line positioning consequently is enigmatic. Our data and that of Usher and McHale (1976) indicate that the rate enhancement derived by perfectly positioning a linkage for in-line attack may be as little as 10- to 20-fold greater than the uncatalyzed rate measured for a typical unconstrained linkage. Considering the fact that complete deprotonation of the nucleophilic 2′-hydroxyl group contributes ~2 million-fold enhancement toward the rate of cleavage (Li & Breaker, 1999), the paltry contribution that in-line positioning makes to the overall rate enhancement of an enzyme seems unimportant. However, if the global structure of an RNA holds a linkage in a non-in-line conformation (e.g., an A-form helix) then the theoretical rate for RNA transesterification at this site would be zero, assuming that the structure is not transiently denatured and that no alternate nucleophilic attack directions are permitted. Therefore, the theoretical maximum rate enhancement derived from in-line positioning compared to the rate of zero for a non-in-line linkage would be infinity. Of course, attaining a perfectly stable non-in-line conformation under physiological conditions is not practical. The lowest rate constants observed in our study are $\sim 10^{-8}$ min^{-1} for linkages that are positioned in a non-in-line orientation. Compared to these experimentally derived values, in-line positioning provides rate enhancements of as much as 10^3 to 10^4.

In-line orientation and RNA-cleaving enzymes

Protein enzymes and ribozymes that cleave RNA via the transesterification reaction depicted in Figure 1 proceed via an in-line S_N2-like mechanism (Walsh, 1979; Kuimelis & McLaughlin, 1998). As a result, these enzymes are obligated to accommodate an in-line conformation at the site of cleavage. Several atomic-level structures are available for self-cleaving RNAs that provide details of the geometry at the cleavage sites (e.g., Brown et al., 1983; Scott et al., 1995; Wedekind & McKay, 1999). In all cases, the cleavage-site linkage is not well positioned for in-line attack. Presumably, the functional structure of the active site is transient and the in-line conformation of the target linkage is not captured by the physical methods used for structure determination (Murray et al., 1998).

One of the most studied RNA-cleaving enzymes is RNase A, which uses a combination of catalytic strategies to generate an overall rate enhancement of nearly 12 orders of magnitude above the uncatalyzed rate of RNA cleavage (Thompson et al., 1995). X-ray structural analysis of the related RNase S enzyme docked with the noncleavable substrate analog UpcA (Richards & Wyckoff, 1971; Carlson, 1976) has been performed with a resolution of 2 Å. UpcA is a relatively isosteric analog of the dinucleotide UpA that carries a 5′-deoxy-5′-phosphonomethylene linkage, which is expected to accurately simulate the conformation of the normal substrate when bound in the active site. The linkage geometry of UpcA in the active site of RNase A corresponds well with an in-line conformation ($F = 0.59$), suggesting that part of the catalytic strategy of this enzyme is to preorganize the substrate for in-line attack via an S_N2-like mechanism.

CONCLUSIONS

In-line orientation of the attacking 2′-oxygen nucleophile relative to the 5′-oxyanion leaving group is a critical aspect of the cyclizing mechanism for RNA cleavage. Consequently, a comprehensive evaluation of the function of enzymes and ribozymes that promote RNA transesterification can proceed only with a detailed understanding of the contributions that in-line orientation can make towards cleavage rate enhancements. Although intricately folded RNA structures inherently offer a source of precisely formed linkage geometries, the recent increase in the number of atomic-level structures being published has only now provided access to the details necessary for assessing the value of in-line orientation. The methods reported herein provide an experimental means by which the importance of in-line preorganization to RNA-cleaving enzymes can be assessed. Furthermore, this approach makes possible the experimental examination of theories relating to the importance of near-attack conformations (Bruice & Lightstone, 1999) for other chemical reactions.

In this study, we have established a continuum of in-line fitness values for various RNA linkages using several high-resolution structural models for RNA. Using this data set, we find that in-line orientation provides a minimum of 10- to 20-fold rate enhancement for the cleavage of RNA by an intramolecular transesterification mechanism. Establishing an accurate value for the maximum rate enhancement generated by in-line preorganization is difficult given the inherent imprecision of structural data and the complications encountered in establishing cleavage rate constants. Improvements in the resolution of these models would allow even more precise determinations to be made of the importance

of in-line orientation to RNA cleavage. However, it seems reasonable to assume that the expansion of the current data set using new RNA structures could provide sufficient guidance for approximating the maximum rate enhancement with reasonable accuracy. With this level of understanding, it becomes practical to comprehend in more detail the catalytic strategies used by various RNA-cleaving enzymes.

MATERIALS AND METHODS

Determination of τ and in-line fitness (F)

For each RNA linkage investigated, the angular degree of nucleophilic attack (τ) by the 2′ oxygen (O2′) on the phosphorus electrophile (P) relative to the 5′-oxygen leaving group (O5′) was determined from published atomic coordinates. Equation (1) represents the cosine rule that holds for any triangle with angles A, B, and C and sides opposite a, b, and c, respectively.

$$\cos A = \frac{b^2 + c^2 - a^2}{2bc} \quad (1)$$

Therefore, equation (2) can be used for a triangle whose sides are defined by the O2′-P and O2′-O5′ interatomic distances, and the P-O5′ bond length (Fig. 1B), where $d_{x,y}$ is the distance between the specified atoms x and y.

$$\cos \tau = \frac{d^2_{O2',P} + d^2_{P,O5'} - d^2_{O2',O5'}}{2 d_{O2',P} d_{P,O5'}} \quad (2)$$

The estimated in-line fitness (F) of each phosphodiester linkage for intramolecular phosphoester transfer was ranked from 0 (least disposed) to 1 (most disposed). Both τ and O2′-P distance in angstroms are considered in the fitness equation, where each factor is scaled by a minimum observed valued of 45° for τ, and 3 Å for O2′-P distance among the 136 sites examined in this study. These minimum values are established by the force fields (e.g., Cornell et al., 1995) used to compute the minimum-energy structural models with the structural constraints derived by NMR and X-ray crystallography. The distance-dependent reactivity of the nucleophile towards the electrophile will depend on the proximity distribution of the two groups relative to each other in space. In this case, the reactivity of O2′ with P is estimated to decrease as a function of the inverse cube of the distance. In-line fitness (F) may thus be estimated using equation (3), where the first term represents the contribution of the τ angle and the second term represents the contribution of the attack distance.

$$F = \frac{(\tau - 45)}{180 - 45} \times \frac{3^3}{d^3_{O2',P}} \quad (3)$$

An exact value for the factor that determines the distance-dependent reactivity of the 2′ oxygen and the phosphorus center is not known. The frequency of productive collisions is expected to decrease with increasing distance between the reactive groups. Above absolute zero, the distance between the two nuclei is not static. Therefore, both nuclei are clustered around some thermodynamically optimal distance, each with a distribution curve whose shape is dictated by the temperature of the reaction. The frequency of collision is a function of the shapes of the distance-distribution curves. Because the precise shapes of these curves are expected to vary with temperature and are otherwise difficult to define, it is not possible to assign a precise factor for the distance-dependent reactivity. To address the importance of the accuracy of the distance-dependent reactivity factor for equation 3, we have examined our data using exponents from one through six. In each instance, the trend of increasing F correlating with increasing rate constant was present. Therefore, we arbitrarily chose to display the data that was generated using the inverse cube of the distance, which represents an intermediate estimate for the distance-dependent reactivity factor.

Molecular graphics

Molecular graphics images were produced using the MidasPlus package from the Computer Graphics Laboratory, University of California, San Francisco (supported by National Institutes of Health grant P41 RR-01081). Models were constructed and examined using a Silicon Graphics O2 workstation.

Preparation of RNAs

RNAs were transcribed from double-stranded DNA templates using T7 RNA polymerase. DNA templates complementary to the FMN aptamer, ATP aptamer, and HDV RNA were prepared by standard solid phase methods (Keck Biotechnology Resource Laboratory, Yale University) and purified by denaturing (8 M urea) polyacrylamide gel electrophoresis (PAGE) as previously described (Tang & Breaker, 1997). Each synthetic DNA template was converted to double-stranded DNA by reverse transcriptase (RT) primer extension of the complementary oligonucleotide 5′-TAATACGACTCACTATAG, which encodes the T7 RNA polymerase promoter. RT extension reactions (50 μL) containing 200 pmol template, 300 pmol primer, 50 mM Tris-HCl (pH 8.3 at 23 °C), 75 mM KCl, 3 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 1 mM of each dNTP, and 400 U SuperScript II reverse transcriptase (Gibco BRL) were incubated at 37 °C for 30 min. Extension products were precipitated with ethanol and redissolved in deionized water (dH_2O). A double-stranded DNA template encoding P5abc RNA was the generous gift of J. Doudna. Transcription reactions (50 μL) containing ~40 pmol template, 40 mM Tris-HCl (pH 8.0 at 23 °C), 15 mM $MgCl_2$, 5 mM DTT, 50 μg/mL bovine serum albumin (BSA), 1 mM of each NTP, and ~1,000 U T7 RNA polymerase were incubated at 37 °C for 1 h. RNAs were purified by PAGE, isolated from the gel by crush-soak elution, precipitated with ethanol, and redissolved in dH_2O.

Prior to analysis, RNA was dephosphorylated using calf intestinal alkaline phosphatase (CIAP; Boehringer Mannheim) and 5′ end labeled using T4 polynucleotide kinase (T4 PNK; New England Biolabs). Phosphatase reactions (20 μL) containing ~50 pmol RNA, 50 mM Tris-HCl (pH 8.5 at 20 °C), 100 μM EDTA, 2 U CIAP were incubated at 50 °C for 1 h. Dephosphorylated RNA was precipitated with ethanol and redissolved in dH_2O. Kinase reactions (20 μL) containing

~10 pmol dephosphorylated RNA, 70 mM Tris-HCl (pH 7.6 at 25 °C), 10 mM $MgCl_2$, 5 mM DTT, ~10 pmol [γ-^{32}P]-ATP (~60 μCi), and 20 U T4 PNK were incubated at 37 °C for 1 h. End-labeled RNAs were purified by PAGE, isolated from the gel by crush-soak elution, precipitated with ethanol, redissolved in water, and quantitated by liquid scintillation counting.

Analysis of RNA cleavage kinetics

The rate constants for the spontaneous transesterification of RNA linkages were determined using 5′ ^{32}P-labeled RNAs. Labeled FMN and ATP aptamers were incubated in the absence or presence of 200 μM FMN or 1 mM ATP, respectively, in reactions (10 μL) containing ~200 fmol RNA, 50 mM Tris-HCl (pH 8.5 at 23 °C), and 20 mM $MgCl_2$. Likewise, labeled P5abc and HDV RNAs were incubated in the absence or presence of 20 mM $MgCl_2$. Reactions were terminated by the addition of an equal volume of gel loading buffer containing 10 M urea and 1.5 mM EDTA.

RNA cleavage ladders were generated by incubating ~200 fmol 5′ ^{32}P-labeled RNA in reactions (10 μL) containing 50 mM $NaHCO_3$/Na_2CO_3 (pH 9.0 at 23 °C) and 1 mM EDTA for 7 min at 90 °C (Knapp, 1989). G-specific sequencing ladders were generated by incubating ~200 fmol of RNA in reactions (10 μL) containing 25 mM sodium citrate (pH 5.0 at 23 °C), 7 M urea, 1 mM EDTA, and 1 U RNase T1 (Boehringer Mannheim) for 15 min at 55 °C. Reactions were quenched on ice with an equal volume of gel loading buffer. Reaction products were separated by PAGE and were visualized and quantitated using a PhosphorImager with ImageQuant software (Molecular Dynamics). The fraction of total RNA cleaved at the 3′ side of each nucleotide position was calculated and corrected for the fraction of cleavage observed at the analogous position of untreated RNA. Observed rate constants (k_{obs}) for cleavage at each nucleotide position of end-labeled RNAs was derived as the negative slope of a line that describes the natural logarithm of the fraction of RNA uncleaved at that position as a function of time.

ACKNOWLEDGMENTS

We thank Frederic Richards for his help with the RNase S structural data and we also thank Yingfu Li and other members of the Breaker laboratory for helpful discussions and for critical reading of the manuscript. This work was supported by a grant from Ribozyme Pharmaceuticals. G.A.S. received support from a Seessel postdoctoral fellowship through Yale University. R.R.B. is also supported by fellowships from the Hellman Family and from the David and Lucile Packard Foundation.

Received April 28, 1999; returned for revision June 21, 1999; revised manuscript received June 30, 1999

REFERENCES

Altman S, Kirsebom L. 1999. Ribonuclease P. In: Gesteland RF, Cech TR, Atkins JF, eds. *The RNA world.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 351–380.

Brown RS, Hingerty BE, Dewan JC, Klug A. 1983. Pb(II)-catalyzed cleavage of the sugar-phosphate backbone of yeast tRNA^{Phe}: Implications for lead toxicity and self-splicing RNA. *Nature 303*: 543–546.

Bruice TC, Lightstone FC. 1999. Ground state and transition state contributions to the rates of intramolecular and enzymatic reactions. *Acc Chem Res 32*:127–136.

Burgstaller P, Famulok M. 1994. Isolation of RNA aptamers for biological cofactors by in vitro selection. *Angew Chem Int Ed Engl 33*:1084–1087.

Butcher SE, Dieckmann T, Feigon J. 1997. Solution structure of the conserved 16 S-like ribosomal RNA UGAA tetraloop. *J Mol Biol 268*:348–358.

Carlson WD. 1976. X-ray diffraction studies of bovine pancreatic ribonuclease-S and *E. coli* alkaline phosphatase. Ph.D. Thesis, Yale University, New Haven, Connecticut. pp 141–201.

Cate JH, Doudna JA. 1996. Metal-binding sites in the major groove of a large ribozyme domain. *Structure 4*:1221–1229.

Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, Kundrot CE, Cech TR, Doudna JA. 1996. Crystal structure of a group I ribozyme domain: Principles of RNA packing. *Science 273*:1678–1685.

Cech T, Golden BL. 1999. Building a catalytic active site using only RNA. In: Gesteland RF, Cech TR, Atkins JF, eds. *The RNA world.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 321–349.

Celander DW, Cech TR. 1991. Visualizing the higher order folding of a catalytic RNA molecule. *Science 251*:401–407.

Chow CS, Bogdan FM. 1997. A structural basis for RNA-ligand interactions. *Chem Rev 97*:1489–1513.

Ciesiolka J, Lorenz S, Erdmann VA. 1992. Different conformational forms of *Escherichia coli* and rat liver 5S rRNA revealed by Pb(II)-induced hydrolysis. *Eur J Biochem 204*:583–589.

Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM Jr, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. 1995. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc 117*:5179–5197.

Dieckmann T, Suzuki E, Nakamura GK, Feigon J. 1996. Solution structure of an ATP-binding RNA aptamer reveals a novel fold. *RNA 2*:628–640.

Draper DE. 1996. Strategies for RNA folding. *Trends Biochem Sci 21*:145–149.

Fan P, Suri AK, Fiala R, Live D, Patel DJ. 1996. Molecular recognition in the FMN-RNA aptamer complex. *J Mol Biol 258*:480–500.

Feig AL, Uhlenbeck OC. 1999. The role of metal ions in RNA biochemistry. In: Gesteland RF, Cech TR, Atkins JF, eds. *The RNA world.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 287–319.

Feigon J, Dieckmann T, Smith FW. 1996. Aptamers from A to ζ. *Chem Biol 3*:611–617.

Ferré-D'Amaré AR, Zhou K, Doudna JA. 1998. Crystal structure of a hepatitis delta virus ribozyme. *Nature 395*:567–574.

Ferrin TE, Huang CC, Jarvis LE, Langridge R. 1988. The MIDAS display system. *J Mol Graphics 6*:13–27.

Gerlt JA. 1993. Mechanistic principles of enzyme-catalyzed cleavage of phosphodiester bonds. In: Linn SM, Lloyd RS, Roberts RJ, eds. *The nucleases.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 1–34.

Gold L, Polisky B, Uhlenbeck O, Yarus M. 1995. Diversity of oligonucleotide functions. *Annu Rev Biochem 64*:763–797.

Jiang F, Kumar RA, Jones RA, Patel D. 1996. Structural basis of RNA folding and recognition in an AMP-RNA aptamer complex. *Nature 382*:183–186.

Jiang L, Suri AK, Fiala R, Patel DJ. 1997. Saccharide-RNA recognition in an aminoglycoside antibiotic-RNA aptamer complex. *Chem Biol 4*:35–50.

Jin R, Chapman WH Jr, Srinivasan AR, Olson WK, Breslow R, Breslauer KJ. 1993. Comparative spectroscopic, calorimetric, and computational studies of nucleic acid complexes with 2′,5″- versus 3′,5″-phosphodiester linkages. *Proc Natl Acad Sci USA 90*: 10568–10572.

Jucker FM, Heus HA, Yip PF, Moors EH, Pardi A. 1996. A network of hydrogen bonds in GNRA tetraloops. *J Mol Biol 264*:968–980.

Kazakov S, Altman S. 1991. Site-specific cleavage by metal ion cofactors and inhibitors of M1 RNA, the catalytic subunit of RNase P from *Escherichia coli. Proc Natl Acad Sci USA 88*:9193–9197.

Kierzek R. 1992. Nonenzymatic hydrolysis of oligoribonucleotides. *Nucleic Acids Res 20*:5079–5084.
Knapp G. 1989. Enzymatic approaches to probing of RNA secondary and tertiary structure. *Methods Enzymol 180*:192–212.
Kuimelis RG, McLaughlin LW. 1998. Mechanisms of ribozyme-mediated RNA cleavage. *Chem Rev 98*:1027–1044.
Lambowitz AM, Caprara MG, Zimmerly S, Perlman PS. 1999. Group I and group II ribozymes as RNPs: Clues to the past and guides to the future. In: Gesteland RF, Cech TR, Atkins JF, eds. *The RNA world.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 451–485.
Li Y, Breaker RR. 1999. Kinetics of RNA degradation by specific base catalysis of transesterification involving the 2′-hydroxyl group. *J Am Chem Soc 121*:5364–5372.
Lightstone FC, Bruice TC. 1996. Ground state conformations and entropic and enthalpic factors in the efficiency of intramolecular and enzymatic reactions. 1. Cyclic anhydride formation by substituted glutharates, succinate, and 3,6-endoxo-D4-tetrahydrophthalate monophenyl esters. *J Am Chem Soc 118*:2595–2605.
McKay DB, Wedekind JE. 1999. Small ribozymes. In: Gesteland RF, Cech TR, Atkins JF, eds. *The RNA world.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 265–286.
Murphy FL, Cech TR. 1993. An independently folding domain of tertiary structure within the *Tetrahymena* ribozyme. *Biochemistry 32*:5291–5300.
Murray JB, Terwey DP, Maloney L, Karpeisky A, Usman N, Beigelman L, Scott WG. 1998. The structural basis of hammerhead ribozyme self-cleavage. *Cell 92*:665–673.
Nonin S, Jiang F, Patel DJ. 1997. Imino proton exchange and base-pair kinetics in the AMP-RNA aptamer complex. *J Mol Biol 268*: 359–374.
Oivanen M, Kuusela S, Lönnberg H. 1998. Kinetics and mechanisms for the cleavage and isomerization of the phosphodiester bonds of RNA by Brønsted acids and bases. *Chem Rev 98*:961–990.
Osborne SE, Ellington AD. 1997. Nucleic acid selection and the challenge of combinatorial chemistry. *Chem Rev 97*:349–370.
Pan T, Uhlenbeck OC. 1992. *Biochemistry 31*:3887–3895.
Patel DJ. 1997. Structural analysis of nucleic acid aptamers. *Curr Opin Chem Biol 1*:32–46.
Patel DJ, Suri AK, Jiang F, Jiang L, Fan P, Kumar RA, Nonin S. 1997. Structure, recognition and adaptive binding in RNA aptamer complexes. *J Mol Biol 272*:645–664.
Polacek N, Barta A. 1998. Metal ion probing of rRNAs: Evidence for evolutionarily conserved divalent cation binding pockets. *RNA 4*:1282–1294.
Puglisi EV, Puglisi JD, Williamson JR, RajBhandary UL. 1994. NMR analysis of tRNA acceptor stem microhelices: Discriminator base change affects tRNA conformation at the 3′ end. *Proc Natl Acad Sci USA 91*:11467–11471.
Pyle AM. 1993. Ribozymes: A distinct class of metalloenzymes. *Science 261*:709–714.
Pyle AM. 1996. Role of metal ions in ribozymes. In: Sigel A, Sigel H, eds. *Metal ions in biological systems*, vol. 32. New York: Marcel Dekker. pp 479–520.
Pyle AM, Green JB. 1995. RNA folding. *Curr Opin Struct Biol 5*:303–310.
Raines RT. 1998. Ribonuclease A. *Chem Rev 98*:1045–1065.
Reynolds MA, Beck TA, Say PB, Schwartz DA, Dwyer BP, Daily WJ, Vaghefi MM, Metzler MD, Klem RE, Arnold LJ Jr. 1996. Antisense oligonucleotides containing an internal, non-nucleotide-based linker promote site-specific cleavage of RNA. *Nucleic Acids Res 24*:760–765.
Richards FM, Wyckoff HW. 1971. Bovine pancreatic ribonuclease. In: Boyer P, ed. *The enzymes*, vol. IV. New York: Academic Press. pp 647–806.
Saenger W. 1984. *Principles of nucleic acid structure*. New York: Springer-Verlag.
Sassanfar M, Szostak JW. 1993. An RNA motif that binds ATP. *Nature 364*:550–553.
Scott WG, Finch JT, Klug A. 1995. The crystal structure of an all-RNA hammerhead ribozyme: A proposed mechanism for RNA catalytic cleavage. *Cell 81*:991–1002.
Suh Y-A, Kumar PKR, Taira K, Nishikawa S. 1993. Self-cleavage activity of the genomic HDV ribozyme in the presence of various divalent metal ions. *Nucleic Acids Res 21*:3277–3280.
Tang J, Breaker RR. 1997. Rational design of allosteric ribozymes. *Chem Biol 4*:453–459.
Thompson JE, Kutateladze TG, Schuster MC, Venegas FD, Messmore JM, Raines RT. 1995. Limits to catalysis by ribonuclease A. *Bioorg Chem 23*:471–481.
Usher DA. 1969. On the mechanism of ribonuclease action. *Proc Natl Acad Sci USA 62*:661–667.
Usher DA, McHale AH. 1976. Hydrolytic stability of helical RNA: A selective advantage for the natural 3′,5′-bond. *Proc Natl Acad Sci USA 73*:1149–1153.
Varani G, Cheong C, Tinoco I. 1991. Structure of an unusually stable RNA hairpin. *Biochemistry 30*:3280–3289.
Walsh C. 1979. *Enzymatic reaction mechanisms.* New York: WH Freeman. pp 199–207.
Wedekind JE, McKay DB. 1999. Crystal structure of a lead-dependent ribozyme revealing metal binding sites relevant to catalysis. *Nature Struct Biol 6*:261–268.
Welch M, Majerfeld I, Yarus M. 1997. 23S rRNA similarity from selection for peptidyl transferase mimicry. *Biochemistry 36*:6614–6623.
Westheimer FH. 1968. Pseudo-rotation in the hydrolysis of phosphate esters. *Acc Chem Res 1*:70–78.
Williams KP, Ciafré S, Tocchini-Valentini GP. 1995. Selection of novel Mg^{2+}-dependent self-cleaving ribozymes. *EMBO J 14*:4551–4557.
Wu M, Tinoco I Jr. 1998. RNA folding causes secondary structure rearrangement. *Proc Natl Acad Sci USA 95*:11555–11560.
Yang Y, Kochoyan M, Burgstaller P, Westhof E, Famulok M. 1996. Structural basis of ligand discrimination by two related RNA aptamers resolved by NMR spectroscopy. *Science 272*:1343–1347.
Zagórowska I, Kuusela S, Lönnberg H. 1998. Metal ion-dependent hydrolysis of RNA phosphodiester bonds within hairpin loops. A comparative kinetic study on chimeric ribo/2′-*O*-methylribo oligonucleotides. *Nucleic Acids Res 26*:3392–3396.
Zhou D-M, Taira K. 1998. The hydrolysis of RNA: From theoretical calculations to the hammerhead ribozyme-mediated cleavage of RNA. *Chem Rev 98*:991–1026.
Zimmermann GR, Jenison RD, Wick CL, Simorre JP, Pardi A. 1997. Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. *Nat Struct Biol 4*:644–649.