## Title

Distinct sites of phosphorothioate substitution interfere with folding and splicing of the Anabaena group I intron



## Journal

Nucleic Acids Research, 32(7)

## ISSN

0305-1048

## Authors

Lupták, Andrej
Doudna, Jennifer A

## Publication Date

2004-04-19

## DOI

10.1093/nar/gkh548

Published online April 23, 2004

2272–2280 Nucleic Acids Research, 2004, Vol. 32, No. 7
DOI: 10.1093/nar/gkh548

# Distinct sites of phosphorothioate substitution interfere with folding and splicing of the *Anabaena* group I intron

**Andrej Lupták[1] and Jennifer A. Doudna[2,3,*]**

[1]Department of Chemistry, [2]Department of Molecular Biophysics and Biochemistry and [3]Howard Hughes Medical Institute, Yale University, New Haven, CT 06520, USA

Received February 26, 2004; Revised and Accepted March 26, 2004

## ABSTRACT

**Although the active site of group I introns is phylogenetically conserved, subclasses of introns have evolved different mechanisms of stabilizing the catalytic core. Large introns contain weakly conserved 'peripheral' domains that buttress the core through predicted interhelical contacts, while smaller introns use loop–helix interactions for stability. In all cases, specific and non-specific magnesium ion binding accompanies folding into the active structure. Whether similar RNA–RNA and RNA–magnesium ion contacts play related functional roles in different introns is not clear, particularly since it can be difficult to distinguish interactions directly involved in catalysis from those important for RNA folding. Using phosphorothioate interference with RNA activity and structure in the small (249 nt) group I intron from *Anabaena*, we used two independent assays to detect backbone phosphates important for catalysis and those involved in intron folding. Comparison of the interference sites identified in each assay shows that positions affecting catalysis cluster primarily in the conserved core of the intron, consistent with conservation of functionally important phosphates, many of which are magnesium ion binding sites, in diverse group I introns, including those from *Azoarcus* and *Tetrahymena*. However, unique sites of folding interference located outside the catalytic core imply that different group I introns, even within the same subclass, use distinct sets of tertiary interactions to stabilize the structure of the catalytic core.**

## INTRODUCTION

Group I self-splicing introns comprise a diverse family of RNA sequences that fold to create a similar active site for RNA transesterification. Splicing is achieved by two consecutive phosphodiester exchange reactions in which the 2′-OH of a bound guanosine cofactor attacks the 5′ splice site phosphate, releasing the 5′ exon for nucleophilic attack at the 3′ splice junction. The conserved catalytic core consists of base paired regions P3, P4, P6 and P7 and the five joining segments J3/4, J4/5, J5/4, J6/7 and J8/7, which fold to create binding sites for the guanosine cofactor and the helical splice sites (Fig. 1). In some group I introns, large peripheral domains with low primary or secondary structural conservation stabilize this active site architecture by buttressing its outer surfaces, while smaller introns utilize loop–helix interactions to maintain the active site conformation. In all cases, specific and non-specific magnesium ion binding is essential for both intron catalysis and folding.

Sequence and secondary structural comparisons do not readily reveal tertiary contacts and chemical features, including ion binding sites, that may be shared by diverse introns. Nucleotide analog interference mapping (NAIM) and footprinting experiments revealed similarities in the functional groups and possible metal ion binding sites required for catalysis by the *Tetrahymena* intron and the more compact *Azoarcus* group I intron (1,2). Such interferences are likely to include both those sites that disrupt structural interactions as well as those involved directly in the catalytic mechanism. To distinguish between structural versus catalytic effects, independent assays are needed for RNA folding and ribozyme reactivity.

We investigated contributions of backbone phosphate oxygens to structure and catalysis in the small (249 nt) group I intron found between the second and wobble anticodon positions in the pre-tRNA$^{Leu}$ of the cyanobacterium *Anabaena* PCC7120 (3). While its short substrate helix (P1) may be held in place by the tRNA secondary or tertiary structure during splicing (4), the intron itself lacks peripheral domains known to stabilize larger group I introns. These include the P5abc domain and the P2.1 and P9.1 domains that nucleate folding and stabilize the catalytic core, respectively, of the *Tetrahymena* intron (5,6), as well as P7.1 and P7.2, two domains important for stability of the T4 *sunY* intron (7). Despite its small size, however, at optimal GTP and magnesium ion concentrations the *Anabaena* intron exhibits the

*To whom correspondence should be addressed at present address: Department of Molecular and Cell Biology, 305 Hildebrandt Hall, Mail Stop 3206, University of California, Berkeley, CA 94720-3206, USA. Tel: +1 510 643 0225; Fax: +1 510 643 0080; Email: doudna@uclink.berkeley.edu
Present address:
Andrej Lupták, Department of Molecular Biology, Massachusetts General Hospital, Boston, MA, USA

 

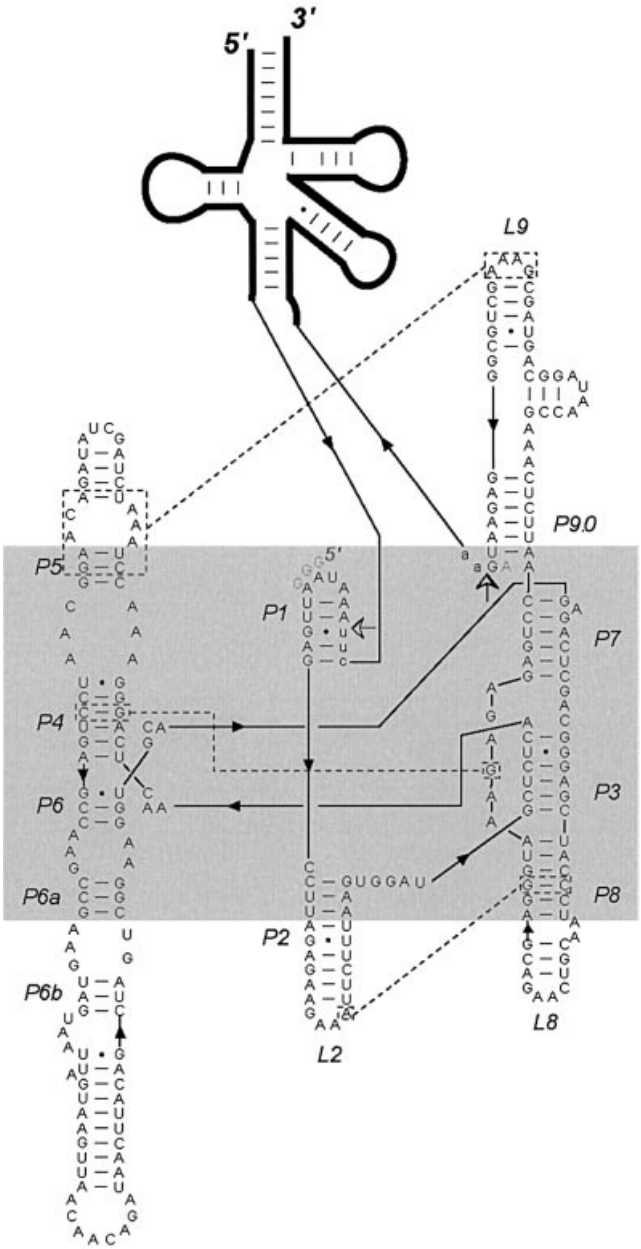


**Figure 1.** Secondary structure of *Anabaena* PCC7120 pre-tRNA$^{leu}$. Upper and lower case letters indicate the group I intron and exon (tRNA) sequences, respectively, with empty arrows showing the splice sites. Grey letters near the splice sites indicate parts where the ribozyme construct differs from the intron. Thick black lines represent parts of the sequence not shown explicitly; thin lines show covalent linkages with filled arrows pointing in the 5′→3′ direction. Tertiary interactions are indicated with dashed lines. Regions of the intron core that show high evolutionary conservation appear in the grey rectangle. The P8 helix, which has traditionally not been considered a core domain, is included in the core because it is highly conserved among group I introns (18).

fastest 5′ splice site cleavage rate measured for any group I intron (~14 min$^{-1}$ at 32°C) (4). Furthermore, the ribozyme version of this intron, lacking 5′ and 3′ splice sites, has been shown to behave similarly to the *Tetrahymena* ribozyme; apparent kinetic differences are a consequence of the short P1 helix rather than the functional organization of the core (8). The ribozyme binds its substrate specifically, cleaves it with high fidelity and rapidly releases the products; its only enzymatic 'imperfection' comes from the rapid release of most bound substrates, thus lowering its catalytic efficiency (8).

The absence of large peripheral domains in the *Anabaena* intron suggests that different RNA–RNA and metal ion-mediated interactions stabilize the active site and contribute to catalysis compared to those identified in the larger *Tetrahymena* intron. To determine the locations of such interactions, we used phosphorothioate interference to perturb *Anabaena* intron folding and catalytic activity, respectively, in two independent assays. Comparison of the interference sites identified in each assay shows that positions affecting catalysis cluster primarily in the conserved core of the intron, consistent with conservation of magnesium ion binding sites and functionally important RNA–RNA interactions in diverse group I introns, including those from *Azoarcus* and *Tetrahymena* (1). However, unique sites of folding interference located outside the catalytic core imply that different group I introns, even within the same phylogenetic subclass (as defined by secondary structure), use distinct sets of interactions to stabilize the intron structure.

## MATERIALS AND METHODS

The *Anabaena* sp. *tRNALeu* pre-tRNA gene, the ribozyme version of the *Anabaena* intron (L-4 248A) subcloned from the *tRNA$^{Leu}$* gene and the *Azoarcus* ribozyme (L-3 206G) subcloned from the *tRNA$^{Ile}$* gene were transcribed from linearized plasmids in the presence of 5′-*O*-(1-thiophosphate) NTPs and purified as described (9). Ribozymes used in the folding assays were dephosphorylated with alkaline phosphatase (NEB) and 5′-$^{32}$P-labeled using T4 polynucleotide kinase. The *Anabaena* pre-tRNA was transcribed using a limiting amount of $Mg^{2+}$ to prevent self-cleavage, purified by denaturing PAGE and 3′ end-labeled using T4 RNA ligase (Boehringer) and [$^{32}$P]pCp (10). The spliced intron was 5′ end-labeled using α-[$^{32}$P]GTP as the substrate in the 5′-splicing reaction (3 mM $MgCl_2$, 33 mM Tris, 66 mM HEPES, pH 7.1, and 400 nM [α-$^{32}$P]GTP for 20 min at 27°C). In subsequent experiments the unspliced and spliced *Anabaena* introns, $^{32}$P-labeled on opposite ends, were analyzed on separate sequencing gels.

Phosphorothioate interference with folding was determined as follows. The purified *Anabaena* ribozyme was resuspended in a minimum amount of TE buffer (2–5 µl) and 10% glycerol, applied without annealing to a 10% PAGE gel containing 2 mM $MgCl_2$ in TH buffer (33 mM Tris, 66 mM HEPES, pH 7.1) and run at room temperature at 8–10 W with identical $Mg^{2+}$-containing TH buffers as running buffers. After 3 h electrophoresis, the gels were exposed wet to a photographic film and the bands corresponding to the folded and unfolded RNA species were excised and eluted into TE buffer. Extracted and precipitated RNA was resuspended in DEPC-treated water and mixed with an equal volume of loading buffer (8 M urea with xylene cyanol and bromophenol blue loading dyes). Each sample was then divided and half was reacted with a 0.1× volume of 10 mM iodine dissolved in ethanol (–20°C) in order to cleave the RNA backbone at the sites of phosphorothioate incorporation (11). Introducing urea

before the iodine cleavage helps to denature the RNA, resulting in more uniform cleavage of the backbone. The cleaved RNA and its uncleaved control were loaded onto sequencing PAGE gels: a 20% gel to map the labeled end of the molecule and a 10% gel to map the opposite end. This way the unfolded and folded populations and their control reactions (showing the sites of background RNA degradation) were compared in terms of the phosphorothioate interference with $Mg^{2+}$-dependent folding. The gels were dried and exposed to image plates (Molecular Dynamics), scanned using a Molecular Dynamics Storm 820 scanner and analyzed in the peak mode of the ImageQuant software.

Integrated peaks from the folded and unfolded subpopulations were compared directly in each experiment: the integrated peaks were first normalized to the total signal in all resolved peaks in a given lane and then folding interference, the ratio of unfolded to folded values, was calculated for each position. Folding interference from at least three experiments was then averaged and evaluated as follows: weak interferences were assigned to positions with average folding interference between 1.5 and 2 and strong interference was assigned to values >2. The standard deviation for all folding interference was 0.29, thus, strong interference corresponds to those that are at least 3 SD above the mean. Positions with significant background degradation were excluded from the analysis. Folding interference was determined for positions starting with A10 and ending with C240; positions U221, U228, U238 and G239 were not resolved and were excluded from the analysis.

Phosphorothioate interference with splicing was determined as follows. All peaks from iodine cleavage reactions were integrated and normalized with a trace line fitted to the data in each lane. This scaling procedure was necessary because the integrated values were found to be slightly biased towards the positions closer to the labeled end of the molecule, i.e. the pre-tRNA, which was labeled at its 3′ end, showed higher integrated values in the 3′ end of the molecule and the spliced intron, which was labeled at its 5′ end, had stronger peaks at its 5′ end. This bias may originate from better electrophoretic resolution of bands that are closer to the labeled end of the molecule or from slight iodine over-digestion of the incorporated phosphorothioates, if their incorporation rate was higher than one per molecule. The scaling procedure helped to reduce the standard deviation of phosphorothioate incorporation at each position, without any apparent effect on the relative values between positions. The phosphorothioate splicing interference values for each position were calculated differently from those for folding in that the integrated and scaled peaks were first averaged for each position in the unspliced and spliced populations separately and the interference ratio for each position in the sequence was calculated by dividing the average unspliced value by the average spliced value. The average standard error for all splicing interference was 0.29. As with the folding interference, weak interference was assigned to positions with values between 1.5 and 2 and strong interference was assigned to values >2.

Size exclusion chromatography was performed with unlabeled phosphorothioate-doped ribozymes using an analytical Superdex 75 column (Pharmacia) at 0.5 ml $min^{-1}$ flow rate. The buffer used consisted of TH, pH 7.1, 50 mM KCl and 0–50 mM $MgCl_2$. KCl was added to the native buffer to prevent electrostatic interaction with the gel matrix. Typically, 10 µl of RNA was injected in 10 mM Tris–HCl buffer, pH 8, containing 0.1 mM EDTA. A sample loaded in formamide had the same retention time as a sample loaded in TE buffer. The samples used in the chromatographic analysis were derived from those used in the folding phosphorothioate assay and thus contained low level phosphorothioate substitutions.

## RESULTS AND DISCUSSION

### Two-state folding of the *Anabaena* group I intron ribozyme

The first aim of this study was to establish whether the *Anabaena* ribozyme undergoes two-state folding, which would facilitate analysis of folding by phosphorothioate interference. The *Anabaena* ribozyme construct used in the folding assay is a truncated form of the intron (Ana L-4 ωA; Fig. 1, grey letters) that starts 4 nt downstream from the 5′ splice site; the first 3 nt are mutated from UAA to GGG for efficient transcription initiation. The construct ends at the 3′ splice site where the terminal guanosine (ωG) is mutated to A (ωA) (Fig. 1, grey letters).

To measure its magnesium-dependent compaction, the *Anabaena* intron was transcribed and analyzed using size exclusion chromatography. For comparison, we tested a related ribozyme, a truncated form of the *Azoarcus* pre-$tRNA^{Ile}$ group I intron (Azo L-3 ωG), under the same conditions. The *Azoarcus* ribozyme exhibits remarkable stability and very fast two-state folding kinetics (12,13) and is the smallest known self-splicing group I intron (14). Magnesium-dependent compaction of the RNA manifests itself as a decrease in average hydrodynamic radius, resulting in retarded mobility on a size exclusion column (Fig. 2a). An analogous assay was used by Weeks and co-workers to study collapsed non-native states of the larger bI5 group I intron (15). In that case, the collapsed non-native state of bI5 as measured by size exclusion chromatography and native PAGE was observed at well below the half-maximum magnesium ion concentration required for full catalytic activity or full protection from hydroxyl radicals ($[Mg^{2+}]_{1/2} \approx 3$ versus 17 mM). In our study, the Ana L-4 ωA and Azo L-3 ωG constructs were found to have a $[Mg^{2+}]_{1/2}$ of collapse of 3 and 1 mM, respectively, when measured in the presence of 50 mM monovalent salt (KCl) and Tris–HEPES buffer (Fig. 2b). In contrast to the bI5 intron, compaction of the bacterial group I introns correlates with folding into the active structure: the *Anabaena* splicing reaction saturates at 4 mM $Mg^{2+}$ (4) and the *Azoarcus* intron was shown by Fe(II)–EDTA footprinting to have an average midpoint of $Mg^{2+}$-dependent folding at 1 mM while its $[Mg^{2+}]_{1/2}$ of splicing is 2.4 mM under similar conditions (13). These data are consistent with a two-state model for intron folding in which molecules fold cooperatively into the active structure without forming stable inactive folding intermediates.

### Independent assays to detect folding versus splicing interference

Phosphorothioate interference was used to identify and compare phosphates in the *Anabaena* intron backbone that contribute to catalytic activity and RNA folding. Two forms of

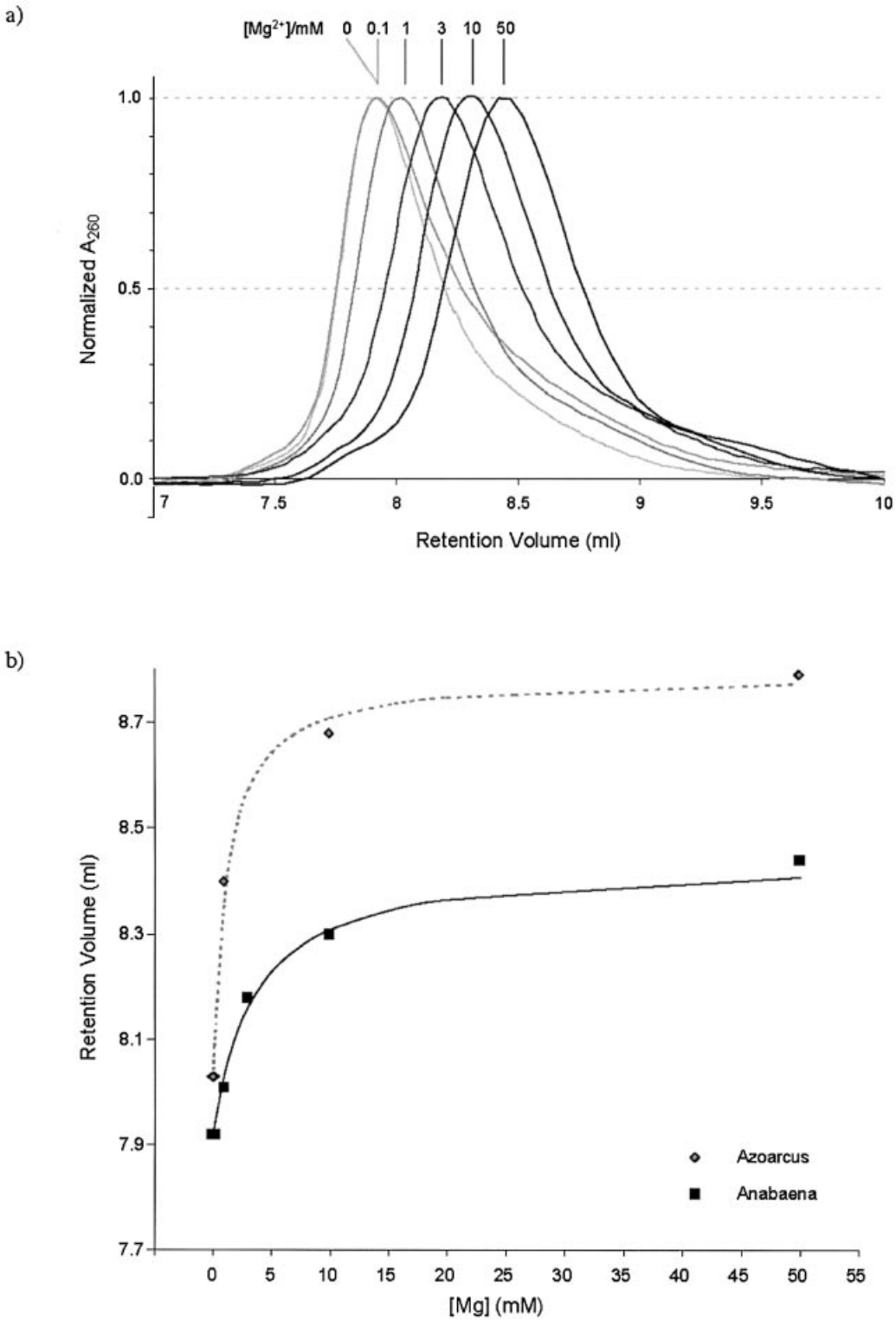


**Figure 2.** (**a**) Size exclusion chromatography of the *Anabaena* intron. Chromatography was performed in Tris–HEPES buffer, pH 7.1, 50 mM KCl and varying concentrations of $MgCl_2$ (indicated above the normalized chromatograms). (**b**) Compaction of *Azoarcus* (diamonds) and *Anabaena* (squares) ribozymes. Peak maxima of the chromatograms are plotted as a function of $[Mg^{2+}]$. Data points were fitted to a two-state function assuming no $Mg^{2+}$ binding cooperativity (15). The resulting $[Mg^{2+}]_{1/2}$ for *Azoarcus* and *Anabaena* group I ribozymes are 1 and 3 mM, respectively.

the intron, a splicing-competent construct flanked by the pre-tRNA and the Ana L-4 ωA ribozyme construct, were transcribed *in vitro* using T7 RNA polymerase and a ratio of nucleoside triphosphates (NTPs) to 5′-*O*-(1-thiophosphate) NTPs sufficient to produce one or two random $R_p$ phosphorothioate substitutions per intron molecule (9). After purifica-

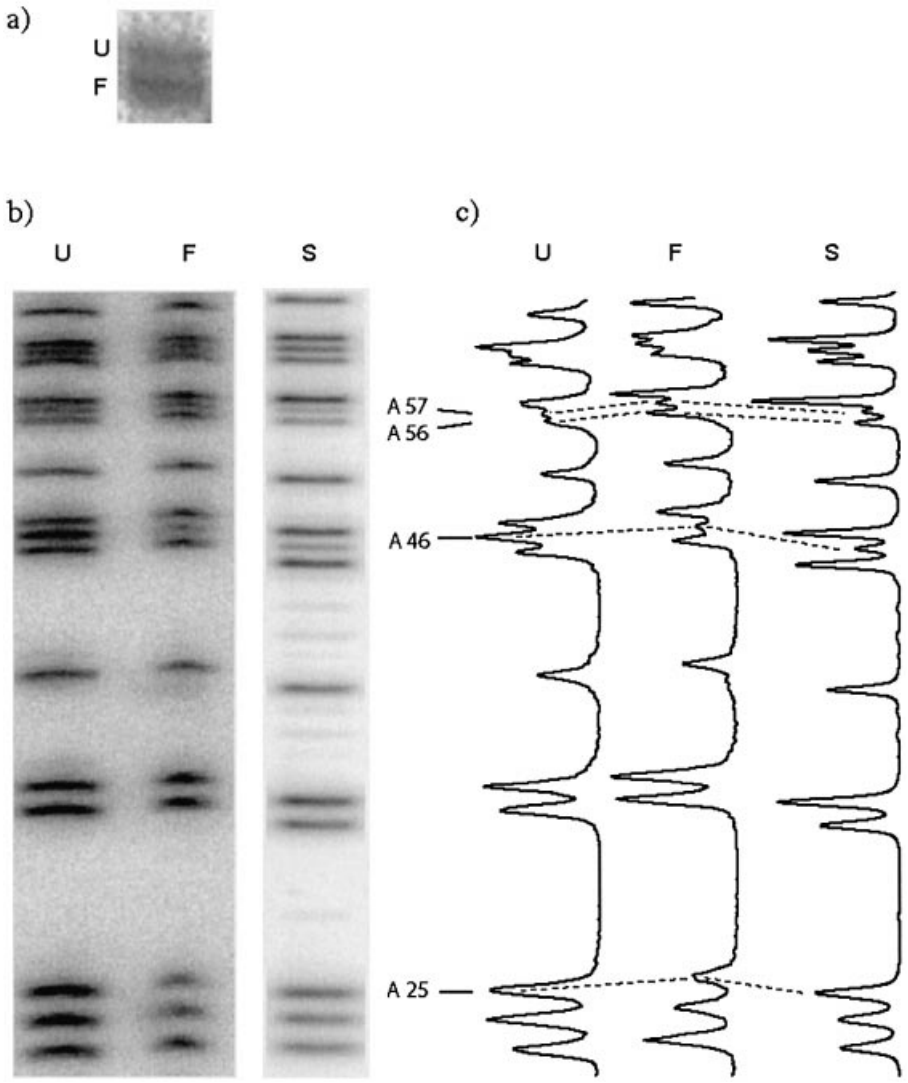


**Figure 3.** (**a**) Separation of unfolded (U) and folded (F) subpopulations of the *Anabaena* ribozyme on native PAGE (2 mM $MgCl_2$, TH buffer, 10% polyacrylamide). (**b**) Sequencing PAGE of iodoethanol-cleaved *Anabaena* group I introns doped with α-phosphorothioate adenosines. Unfolded (U), folded (F) and spliced (S) subpopulations are shown. All molecules are labeled on the 5′ end; ribozymes used in the folding assay were $^{32}$P-phosphorylated using T4 polynucleotide kinase and the spliced intron was labeled with [α-$^{32}$P]GTP used as a substrate in the splicing reaction and analyzed on a separate gel. (**c**) Trace analysis of the sequencing gels. Positions which show significant folding (A25 and A46) and splicing (A46, A56 and A57) interferences are indicated.

tion, phosphorothioate-containing RNA was assayed in two ways. To analyze phosphorothioate effects on RNA folding, the ribozyme construct was incubated in buffer containing magnesium ions at the $[Mg^{2+}]_{1/2}$ concentration, followed by separation of folded and unfolded species on a native polyacrylamide gel (Fig. 3a) (16). To assess the effects on ribozyme catalysis, the self-splicing construct was reacted with exogenous [α-$^{32}$P]GTP substrate at subsaturating concentration (400 nM, $K_{m,G} \approx 240$ μM at 32°C and in 15 mM $MgCl_2$) (4). This allowed for efficient labeling of the spliced intron as well as for some retardation of splicing with respect to folding of the intron. Spliced products were separated from unreacted RNAs by denaturing PAGE. In each case, RNAs were excised and eluted from the gel matrix, cleaved with iodoethanol and analyzed on a sequencing gel to identify positions of phosphorothioate incorporation (11).

The sequencing PAGE showed relative levels of phosphorothioate enrichment in given subpopulations (Fig. 3b). The traces of the PAGE lanes were analyzed by integrating the peaks corresponding to individual positions in the sequence (Fig. 3c). Sites of phosphorothioate interference at positions that affect folding (i.e. A25), splicing (i.e. A56 and 57) and both folding and splicing (i.e. A46) were observed (Fig. 3b and c). Average interferences were mapped to positions along the *Anabaena* intron primary and secondary structure (Fig. 4; positions 101–149 are omitted because no interference was detected in the P6b region).

### Phosphorothioate interference with intron splicing is localized within the catalytic core

Positions of splicing interference in the *Anabaena* intron (Fig. 5) are similar to those in the *Azoarcus* and *Tetrahymena*

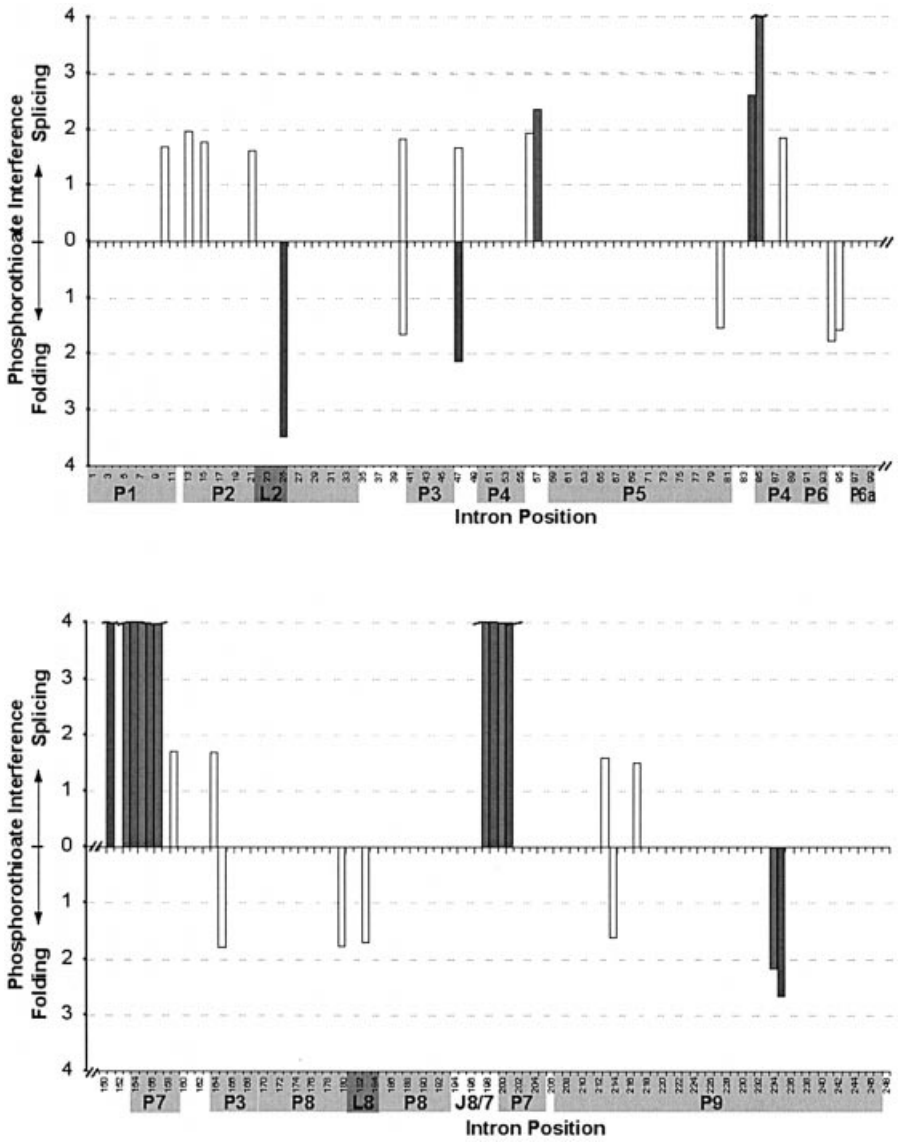


**Figure 4.** Comparison of splicing (up bars) and folding (down bars) phosphorothioate interference in the *Anabaena* group I intron. Interference values are plotted along the primary structure of the intron; secondary structure elements are shaded and labeled bellow the sequence positions. Strong interference is represented by filled bars and weak interference by empty bars. Very strong splicing interference (>4) was truncated at 4. The sequence between position 100 and 150 (the P6b region of the intron) was omitted because neither splicing nor folding interference was found in this region.

introns (1). Each position corresponding to a site detected in the *Azoarcus* and *Tetrahymena* introns gives a strong (>4) splicing interference in *Anabaena* as well (Fig. 4). Strong interferences cluster within the evolutionarily conserved core of the intron in the P4, J6/7, P7 and J8/7 regions. In addition, two strong interferences are found in J4/5 (A57 and A84) and several weak interferences are found in the core of the intron in P1 (A10), P2 (C13 and U15), P3 (G40 and G164), J3/4 (A47), J4/5 (A56), P4 (U88) and P7 (U159). Weak interferences are also found in peripheral parts of the molecule: A21 in P2, and A213 and C217 in P9. Interestingly, only one splicing interference occurs at a position thought to be involved in a tertiary interaction: at G198, which forms a base triple with G54–C86 (17).

## Phosphorothioate interference with intron folding occurs primarily outside the intron catalytic core

Phosphorothioate substitutions that interfere with global folding of the *Anabaena* intron are dispersed throughout the core and the peripheral domains of the molecule (Figs 4 and 5). Only four positions show strong (≥2-fold) phosphorothioate effects: A25 (L2 loop), A47 (J3/4) and A234 and A235 (L9 loop); all of these nucleotides are conserved among the IC3 introns present in the CRW database (18). Weak folding interferences are found in P3 (G40 and G165), P5 (G80), J6/6a (A95 and A96), L8 (U180 and A183) and P9 (A214). Interestingly, only phosphorothioate substitution at G40 and A47 inhibits both splicing and folding. G40 has been proposed to bind magnesium in the related *Azoarcus* intron, where this

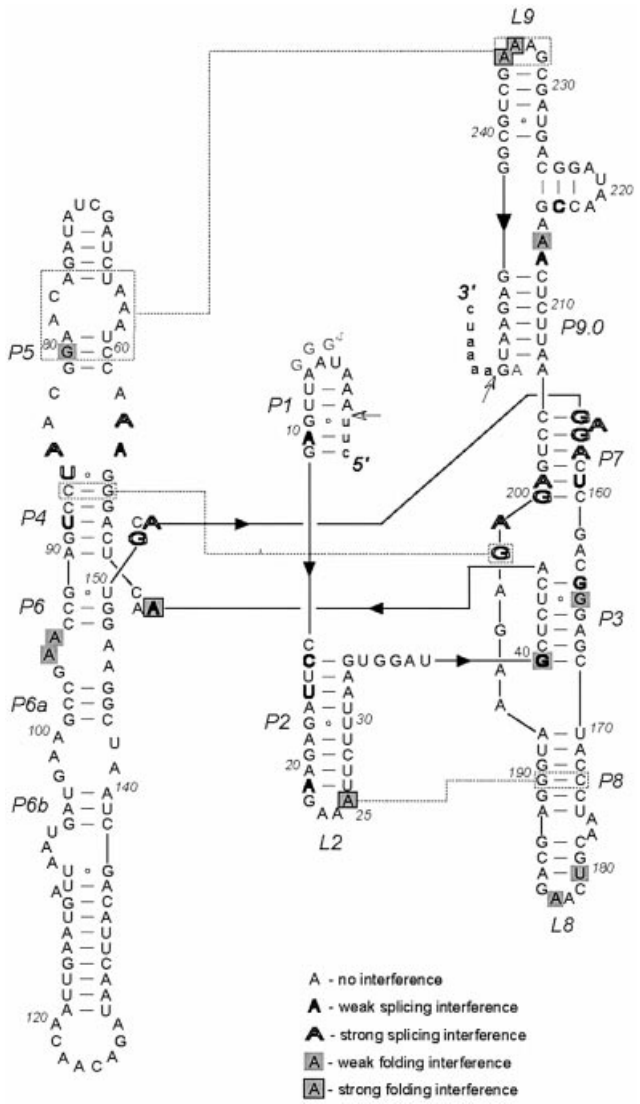


**Figure 5.** Phosphorothioate interference mapped to the secondary structure of the *Anabaena* PCC7120 group I intron. Interference is indicated as follows: grey boxed letters, weak folding interference; grey boxes with black outlines, strong folding interference; bold letters, weak splicing interference; outlined letters, strong splicing interference. Other labeling is as in Figure 1.

site was found to be susceptible to $Tb^{3+}$-dependent hydrolysis in both pre-tRNA and ribozyme constructs (2). Thus it is possible that G40 is a metal ion binding site in *Anabaena* as well. The position corresponding to *Anabaena* A47, commonly placed in the J3/4 strand in secondary structure diagrams of group I introns, was instead modeled as part of the P3 helix in the medium resolution crystal structure of the *Tetrahymena* group I intron core (19). Models of the td and *Azoarcus* introns show the four-helical junction of the core stabilized through interaction of the J6/6a internal loop with P3 (13,20). Thus folding of the backbone in this tightly packed region may be more sensitive to phosphorothioate substitution and may affect the global folding of the intron more than other parts of the core, which show only splicing interference.

Other folding interferences are located in peripheral parts of the intron proposed to be involved in tertiary interactions stabilizing small group I introns (21). These interactions are between the L2 tetraloop and the minor groove of P8 and between the L9 tetraloop and the tetraloop receptor of the P5 domain (Fig. 5). The interference in the L2 GAAA tetraloop is unusual in that the interference arises at the phosphate preceding the last adenine of the tetraloop (A25). Previous studies of metal ion interactions with GAAA tetraloops have shown that metal ions bind the 5′ phosphates of the first and second adenosines. For example, in one GAAA tetraloop solution structure, a cobalt(III) hexammine mimics a fully solvated magnesium ion that contacts the phosphates of the first two adenosines (22) and in a $^{31}P$ NMR study of the tetraloop, divalent metal ions coordinate the pro-$R_p$ and pro-$S_p$ oxygens of the first and second adenosines through outer and inner sphere coordination, respectively (23). Interestingly, in a GAAA model hairpin, an $R_p$ thiophosphate at the second adenosine significantly stabilizes the whole stem–loop, possibly through an alternative structure (23,24). However, no folding interference is observed at the second adenosine position in the L2 loop (A24). The strong folding interference at A25 suggests that this phosphate participates in a magnesium-dependent interaction not observed previously in a GAAA tetraloop.

The interaction between L9 and P5 is thought to be mediated through an 11 nt tetraloop receptor motif, which has been implicated in many tertiary interactions in RNA (25,26). In the L9 GAAA tetraloop, both the second and third adenosines show strong folding interference. Interference is observed in the tetraloop receptor as well (Fig. 5), suggesting that disruption of this interaction affects the stability of the entire intron more than disruption of the L2–P8 interaction. In the *Azoarcus* intron, where both L2–P8 and L9–P5 interactions are mediated through tetraloop receptors, substitution of one of the tetraloop receptors (in P8) with a base paired helix corresponding to the *Anabaena* P8 was shown to destabilize the entire intron (12).

The most striking phosphorothioate interferences with *Anabaena* intron folding are in the L8 region of the molecule. This loop belongs to a family of stable YNMG tetraloops (27) but it is not evolutionarily conserved. Although it has never been implicated in a tertiary interaction in any group I intron, it seems to play an important role in folding of the *Anabaena* intron. Interferences are observed in the third position of the loop (A183) and in the closing U–A base pair (U180). The loop may be involved in a novel $Mg^{2+}$-dependent interaction, possibly with the P6 region of the molecule. Based on the crystal structures of the 30S ribosomal subunit (28) and phylogenetic conservation, Bevilacqua and co-workers have recently suggested that YNMG tetraloops can form stable interactions with two extruded adenosines of other tetraloops (27). Thus, the *Anabaena* L8 loop may interact with extrahelical adenosines of the lower part of the P6b region, although no folding interference has been detected in this region of the molecule.

## Conclusions

The observed phosphorothioate interference may result from several possible effects of sulfur substitution with respect to the phosphate group: (i) a change in the propensity of the phosphate to coordinate metal ions from binding ‘hard’ metals (e.g. calcium or magnesium) to binding more polarizable

'soft' metals (e.g. manganese and cadmium) (29); (ii) a change in the charge distribution of the phosphate (30), possibly affecting the hydrogen bonding network around the group; (iii) the increased atomic radius of sulfur relative to oxygen, introducing steric hinderances at tightly packed sites and affecting the structure of the sites (24,31–33). Attempts to distinguish magnesium ion binding sites from direct RNA–RNA interactions using manganese 'rescue' (1,34) were unsuccessful due to weak (statistically insignificant) changes in the observed phosphorothioate interference when low concentrations of manganese were included in the assays.

It is instructive to compare the results obtained here with those for the *Tetrahymena* intron, where experiments seeking to characterize metal binding sites necessary for folding identified both monovalent and divalent metal binding sites in the P4–P6 domain that were subsequently observed in the crystal structure (16,34–36). Surprisingly, these sites were not detected in phosphorothioate interference studies of splicing of the whole intron (1,37,38). Although splicing interference occurs in the core near the active site of the intron, it was not observed in the P5abc region, where most of the folding interferences in the P4–P6 domain are clustered. In this case, folding of the P4–P6 domain occurs much faster than that of the entire intron (<1 $s^{-1}$ versus ~1 $min^{-1}$), preventing detection, in a splicing assay, of interferences that strictly affect the folding pathway (39–41). While the smaller *Anabaena* group I intron undergoes apparent two-state folding, the lack of overlap in phosphorothioate interferences detected in the folding versus splicing assays likewise suggests that catalysis is rate limiting. It is also possible that stabilization of this intron by its flanking tRNA exons (4) and by the GTP substrate decreases the relative importance of intron tertiary contacts, reducing detection in a splicing phosphorothioate interference study. In addition, interference with splicing may be stronger than that with folding because perturbation of a critical phosphate in the catalytic center affects activity more than disruption of a single phosphate-mediated tertiary interaction affects global folding. Because the intron folds cooperatively, phosphorothioate substitutions that slightly impede folding may not trap the RNA in an inactive conformation, further lessening the observed interference at those sites.

In summary, the results for phosphorothioate interference with splicing indicate that the tertiary interactions within the *Anabaena* intron core are similar to those of the *Azoarcus* and *Tetrahymena* introns (1). In contrast, sites of phosphorothioate interference with folding indicate that while some tertiary interactions are conserved among different introns, a unique site of possible tertiary interaction in the *Anabaena* intron involves the L8 loop. Thus, different group I introns, even within the same phylogenetic subclass, use distinct sets of tertiary interactions to stabilize the structure of the catalytic core.

## ACKNOWLEDGEMENTS

We thank Liz Doherty for help during the early parts of this project, Scott Strobel and his laboratory for fruitful discussions and J.Carothers, M.Hanczyc and R.Larralde for comments on the manuscript. This work was supported in part by the NIH (GM22778).

## REFERENCES

1. Strauss-Soukup,J.K. and Strobel,S.A. (2000) A chemical phylogeny of group I introns based upon interference mapping of a bacterial ribozyme. *J. Mol. Biol.*, **302**, 339–358.
2. Rangan,P. and Woodson,S.A. (2003) Structural requirement for $Mg^{2+}$ binding in the group I intron core. *J. Mol. Biol.*, **329**, 229–238.
3. Xu,M.Q., Kathe,S.D., Goodrich-Blair,H., Nierzwicki-Bauer,S.A. and Shub,D.A. (1990) Bacterial origin of a chloroplast intron: conserved self-splicing group I introns in cyanobacteria. *Science*, **250**, 1566–1570.
4. Zaug,A.J., McEnvoy,M.M. and Cech,T.R. (1993) Self-splicing of the group I intron from *Anabaena* pre-tRNA: requirement for base-pairing of the exons in the anticodon stem. *Biochemistry*, **32**, 7946–7953.
5. Engelhardt,M.A., Doherty,E.A., Knitt,D.S., Doudna,J.A. and Herschlag,D. (2000) The P5abc peripheral element facilitates preorganization of the *Tetrahymena* group I ribozyme for catalysis. *Biochemistry*, **39**, 2639–2651.
6. Lehnert,V., Jaeger,L., Michel,F. and Westhof,E. (1996) New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the *Tetrahymena thermophila* ribozyme. *Chem. Biol.*, **3**, 993–1009.
7. Doudna,J.A. and Szostak,J.W. (1989) Miniribozymes, small derivatives of the sunY intron, are catalytically active. *Mol. Cell. Biol.*, **9**, 5480–5483.
8. Zaug,A.J., Dávila-Aponte,J.A. and Cech,T.R. (1994) Catalysis of RNA cleavage by a ribozyme derived from the group I intron of *Anabaena* pre-$tRNA^{leu}$. *Biochemistry*, **33**, 14935–14947.
9. Ryder,S.P. and Strobel,S.A. (1999) Nucleotide analog interference mapping. *Methods*, **18**, 38–50.
10. Bruce,A.G. and Uhlenbeck,O.C. (1978) Reactions at the termini of tRNA with T4 RNA ligase. *Nucleic Acids Res.*, **5**, 3665–3677.
11. Gish,G. and Eckstein,F. (1988) DNA and RNA sequence determination based on phophorothioate chemistry. *Science*, **240**, 1520–1522.
12. Tanner,M.A. and Cech,T.R. (1996) Activity and thermostability of the small self-splicing group I intron in the pre-$tRNA^{ile}$ of the purple bacterium *Azoarcus*. *RNA*, **2**, 74–83.
13. Rangan,P., Masquida,B., Westhof,E. and Woodson,S.A. (2003) Assembly of core helices and rapid tertiary folding of a small bacterial group I ribozyme. *Proc. Natl Acad. Sci. USA*, **100**, 1574–1579.
14. Reinhold-Hurek,B. and Shub,D.A. (1992) Self-splicing introns in tRNA genes of widely divergent bacteria. *Nature*, **357**, 173–176.
15. Buchemueller,K.L., Webb,A.E., Richardson,D.A. and Weeks,K.M. (2000) A collapsed non-native RNA folding state. *Nature Struct. Biol.*, **7**, 362–366.
16. Cate,J.H., Hanna,R.L. and Doudna,J.A. (1997) A magnesium ion core at the heart of a ribozyme domain. *Nature Struct. Biol.*, **4**, 553–558.
17. Tanner,M.A. and Cech,T.R. (1997) Joining the two domains of a group I ribozyme to form the catalytic core. *Science*, **275**, 847–849.
18. Cannone,J.J., Subramanian,S., Schnare,M.N., Collett,J.R., D'Souza,L.M., Du,Y., Feng,B., Lin,N., Madabusi,L.V., Muller,K.M. *et al.* (2002) The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron and other RNAs. *BMC Bioinformatics*, **3**, 2.
19. Golden,B.L., Gooding,A.R., Podell,E.R. and Cech,T.R. (1998) A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. *Science*, **282**, 259–264.
20. Waldsich,C., Masquida,B., Westhof,E. and Schroeder,R. (2002) Monitoring intermediate folding states of the *td* group I intron *in vivo*. *EMBO J.*, **21**, 5281–5291.
21. Jaeger,L., Michel,F. and Westhof,E. (1994) Involvement of a GNRA tetraloop in long-range RNA tertiary interactions. *J. Mol. Biol.*, **236**, 1271–1276.
22. Rüdisser,S. and Tinoco,I.J. (2000) Solution structure of cobalt(III) hexammine complexed to the GAAA tetraloop and metal-ion binding to G.A mismatches. *J. Mol. Biol.*, **295**, 1211–1223.
23. Maderia,M., Horton,T.E. and DeRose,V.J. (2000) Metal interactions with a GAAA RNA tetraloop characterized by $^{31}P$ NMR and phosphorothioate substitutions. *Biochemistry*, **39**, 8193–8200.
24. Horton,T.E., Maderia,M. and DeRose,V.J. (2000) Impact of phosphorothioate substitutions on the thermodynamic stability of an RNA GAAA tertraloop: an unexpected stabilization. *Biochemistry*, **39**, 8201–8207.
25. Costa,M. and Michel,F. (1995) Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.*, **14**, 1276–1285.

26. Tanner,M.A. and Cech,T.R. (1995) An important RNA tertiary interaction of group I and group II introns is implicated in gram-positive RNase P RNAs. *RNA*, **1**, 349–350.
27. Proctor,D.J., Schaak,J.E., Bevilacqua,J.M., Falzone,C.J. and Bevilacqua,P.C. (2002) Isolation and characterization of a family of stable RNA tetraloops with the motif YNMG that participate in tertiary interactions. *Biochemistry*, **41**, 12062–12075.
28. Wimberly,B.T., Brodersen,D.E., Clemons,W.M., Jr., Morgan-Warren,R.J., Carter,A.P., Vonrhein,C., Hartsch,T. and Ramakrishnan,V. (2000) Structure of the 30S ribosomal subunit. *Nature*, **407**, 327–339.
29. Lippard,S.J. and Berg,J.M. (1994) *Principles of Bioinorganic Chemistry*. University Science Books, Mill Valley, CA.
30. Frey,P.A. and Sammons,R.D. (1985) Bond order and charge localization in nucleoside phosphorothioates. *Science*, **228**, 541–545.
31. Brautigam,C.A. and Steitz,T.A. (1998) Structural principles for the inhibition of the 3′-5′ exonuclease activity of *Escherichia coli* DNA polymerase I by phosphorothioates. *J. Mol. Biol.*, **277**, 363–377.
32. Smith,J.S. and Nikonowicz,E.P. (2000) Phosphorothioate substitution can substantially alter RNA conformation. *Biochemistry*, **39**, 5642–5652.
33. Schwans,J.P., Cortez,C.N., Olvera,J.M. and Piccirilli,J.A. (2003) 2′-Mercaptonucleotide interference reveals regions of close packing within folded RNA molecules. *J. Am. Chem. Soc.*, **125**, 10012–10018.
34. Basu,S. and Strobel,S.A. (1999) Thiophilic metal ion rescue of phosphorothioate interference within the *Tetrahymena* ribozyme P4-P6 domain. *RNA*, **5**, 1399–1407.
35. Cate,J.H., Gooding,A.R., Podell,E., Zhou,K., Golden,B.L., Kundrot,C.E., Cech,T.R. and Doudna,J.A. (1996) Crystal structure of a group I intron ribozyme domain: principles of RNA packing. *Science*, **273**, 1678–1685.
36. Basu,S., Rambo,R.P., Strauss-Soukup,J., Cate,J.H., Ferré D'Amaré,A.R., Strobel,S.A. and Doudna,J.A. (1998) A specific monovalent metal ion integral to the AA platform of the RNA tetraloop receptor. *Nature Struct. Biol.*, **5**, 986–992.
37. Christian,E.L. and Yarus,M. (1992) Analysis of the role of phosphate oxygens in the group I intron from *Tetrahymena*. *J. Mol. Biol.*, **228**, 743–758.
38. Christian,E.L. and Yarus,M. (1993) Metal coordination sites that contribute to structure and catalysis in the group I intron from *Tetrahymena*. *Biochemistry*, **32**, 4475–4480.
39. Zarrinkar,P.P. and Williamson,J.R. (1994) Kinetic intermediates in RNA folding. *Science*, **265**, 918–924.
40. Sclavi,B., Sullivan,M., Chance,M.R., Brenowitz,M. and Woodson,S. (1998) RNA folding at milisecond intervals by synchrotron hydroxyl radical footprinting. *Science*, **279**, 1940–1943.
41. Doherty,E.A., Herschlag,D. and Doudna,J.A. (1999) Assembly of an exceptionally stable RNA tertiary interface in a group I ribozyme. *Biochemistry*, **38**, 2982–2990.