An Independently Folding Domain of RNA Tertiary Structure within the *Tetrahymena* Ribozyme†

Felicia L. Murphy and Thomas R. Cech*

Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215

Received January 7, 1993

ABSTRACT: The *Tetrahymena thermophila* pre-rRNA contains a 413-nucleotide self-splicing group I intron. This intron has been converted into a sequence-specific endonuclease or ribozyme. A 160-nucleotide portion of the ribozyme consisting of both highly conserved sequence elements (P4 and P6) and nonconserved peripheral extensions (P5abc and P6ab) was synthesized as a separate molecule. Solvent-based Fe(II)–EDTA, a probe that monitors higher-order RNA structure, revealed a protection pattern that was a large subset of that observed in the whole ribozyme. Data from dimethyl sulfate modification and partial digestion with nucleases were also consistent with maintenance of the proper secondary and tertiary structure in the shortened RNA molecule. Thus, this 160-nucleotide molecule (P4–P6 RNA) is an independently folding domain of RNA tertiary structure. A series of mutations and deletions were made within the P4–P6 domain to further dissect its tertiary structure. Fe(II)–EDTA and dimethyl sulfate analysis of these mutants revealed that the domain consists of two substructures, a localized subdomain involving the characteristic adenosine-rich bulge in P5a, and a subdomain-stabilized structure involving long-range interactions. Therefore, like some proteins, the intron RNA is modular, containing a separable domain and subdomain of tertiary structure.

RNA molecules were once thought to function only in facilitating the information transfer between DNA and protein or in providing a structural scaffold for protein. There are now many examples of RNA catalysts (Michel et al., 1989; Altman, 1990; Cech, 1990; Pace & Smith, 1990; Symons, 1992). RNA catalysis generally involves nucleophilic attack at a phosphorus atom in RNA resulting in transesterification or hydrolysis (Cech, 1987). However, RNA can also catalyze hydrolysis of an aminoacyl linkage (Piccirilli et al., 1992), and evidence is mounting that it may be the catalytic component of ribosomes (Noller et al., 1992, and references therein). The prevalence of RNA-like cofactors suggests that RNAs once catalyzed a complex array of chemical reactions (Visser & Kellogg, 1978; Benner et al., 1989). Since RNA is a significant and potentially manipulatable source of catalysts, it is important to elucidate principles of RNA structural organization.

Free radicals produced by solvent-based Fe(II)–EDTA have proven useful for RNA tertiary structure analysis. On the basis of model studies with DNA, the probe mediates a free radical reaction with the sugar moiety of the backbone that ultimately results in strand scission. The reactive species has been proposed to be hydroxyl radical (Hertzberg & Dervan, 1984; Tullius, 1985, 1986) or a ferryl ion species such as (FeO_2^+)–EDTA (Rush & Koppenol, 1986; Rahhal & Richter, 1988). The Fe(II)–EDTA probe does not exhibit significant sensitivity to nucleotide sequence or secondary structure (helices versus single-stranded regions), but instead monitors higher-order structure (Latham & Cech, 1989; Celander & Cech, 1990). On the basis of model studies with yeast $tRNA^{Phe}$, protection from cleavage with the probe correlates with decreased solvent accessibility (Latham & Cech, 1989). Other structural features such as the proposed minor groove narrowing in bent DNA can also result in decreased cleavage by the probe (Burkhoff & Tullius, 1987). Since the probe reacts with ribose moieties, every residue is tested for higher-order structure. We describe the higher-order structure visualized by this probe as tertiary structure.

Using the solvent-based Fe(II)–EDTA probe, it has been shown that several ribozymes derived from self-splicing group I introns can adopt a complex tertiary structure with an inside and an outside (Latham & Cech, 1989; Heuer et al., 1991). On the basis of the extent of Fe(II)–EDTA protection (Figure 1) and direct visualization by electron microscopy (Y.-H. Wang, J. Griffith, F. L. Murphy, and T. R. Cech, unpublished results), the *Tetrahymena* ribozyme appears to have a compact structure, a feature characteristic of protein enzymes. Tertiary structure formation by the *Tetrahymena* ribozyme has a cation requirement that can be met by any of several divalent cations (Grosshans & Cech, 1989; Latham & Cech, 1989; Celander & Cech, 1991), while ribozyme-mediated catalysis specifically requires Mg^{2+} or Mn^{2+} (Grosshans & Cech, 1989; Piccirilli et al., 1993). As $MgCl_2$ concentration is increased, the acquisition of catalytic activity correlates with structure formation as monitored by the Fe(II)–EDTA cleavage pattern (Celander & Cech, 1991).

We reasoned that the *Tetrahymena* ribozyme may, like protein, have a modular organization. If domains of higher-order structure occur, it should be possible to divide the RNA into pieces that maintain tertiary structure independent of the remainder of the molecule. The solvent-based Fe(II)–EDTA probe should reveal a pattern of protection in a tertiary domain that is a subset of that observed in the intact ribozyme.

Group I introns are composed of "core" elements that are highly conserved in secondary structure and to some extent in nucleotide sequence among all group I introns, and peripheral extensions that vary among introns (Figure 2a). The core elements are directly implicated as the catalytic and structural center of the ribozyme. Although the core might seem to be a good candidate for a structural domain, data

† This work was supported by the National Institutes of Health (Grant GM28039). T.R.C. is an Investigator, Howard Hughes Medical Institute, and an American Cancer Society Professor.

* Author to whom correspondence should be addressed.

0006-2960/93/0432-5291$04.00/0

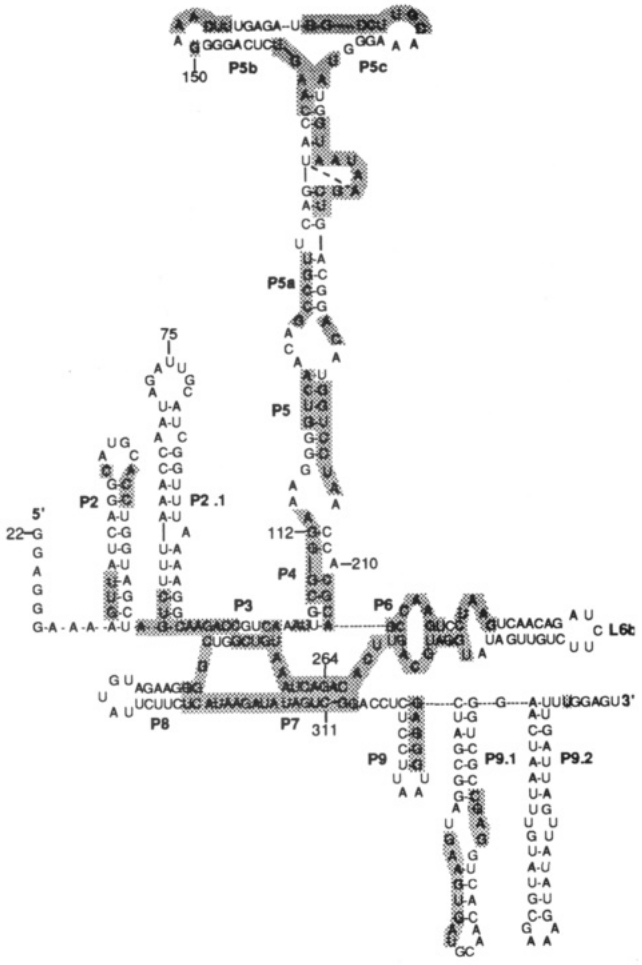

FIGURE 1: Secondary structure of the L-21 *Sca*I *Tetrahymena* ribozyme. Ribose moieties protected from cleavage with Fe(II)–EDTA are shaded. The ribozyme, which has truncated 5′ and 3′ ends relative to the full-length intron, is a sequence-specific endoribonuclease (Zaug et al., 1986). The original Fe(II)–EDTA map (Latham & Cech, 1989) was revised previously (Heuer et al., 1991) and herein (B. Laggerbauer, F. L. Murphy, & T. R. Cech, unpublished results). Secondary structure is from Michel and Dujon (1983), Waring et al. (1983), and Michel and Westhof (1990), with helices named according to Burke et al. (1987).

suggest that in the *Tetrahymena* and phage T4 sunY introns the core alone does not effectively maintain structure. For example, in these introns deletion of one or a combination of peripheral extensions can destroy activity or result in an increased Mg^{2+} requirement for activity (Price et al., 1985; Doudna & Szostak, 1989; Joyce et al., 1989; Michel et al., 1992). Increased Mg^{2+} is thought to stabilize the active form of the intron. Other introns require protein splicing factors (Lambowitz & Perlman, 1990), and in one case there is evidence consistent with the protein stabilizing the active structure (Guo & Lambowitz, 1992). It remains to be seen whether some naturally occurring small introns like that of *Anabaena* will have an independently folding core domain (Xu et al., 1990).

Peripheral extensions such as the P5 extension (Figure 2a) in the *Tetrahymena* intron and P7.1, 7.2 in the T4 introns are also candidates for domains of higher-order structure. While peripheral extensions are not conserved among all introns, there is substantial conservation within intron subgroups. We considered the extension of the P5 stem a good candidate for a tertiary domain. This extension is common to many group IB and IC introns including the *Tetrahymena* intron (Michel & Westhof, 1990). It contains a characteristic adenosine-rich (A-rich) bulge located a conserved distance from the core (Michel & Cummings, 1985; Collins, 1988). It has been shown that the P5 extension, as an independent molecule, can activate the catalytic activity of an intron in which the P5 extension has been deleted (van der Horst et al., 1991). Furthermore, structural analysis of a deletion derivative of the *Tetrahymena* ribozyme indicated that P4 and elements above it (P5abc) retained structure even when the core was disrupted (Latham & Cech, 1989).

As a candidate for a tertiary domain, a molecule containing P4 and the P5 and P6 extensions was designed. The P6 extension was included since it has been postulated that the P4 and P6 stems are coaxially stacked (Kim & Cech, 1987; Michel et al., 1990). Also included were short 5′ and 3′ extensions that have been postulated to form a triple-helical scaffold with the P4 and P6 duplexes (Michel et al., 1990; Michel & Westhof, 1990). This RNA molecule (P4–P6 RNA; Figure 2b) contains elements of the highly conserved core in addition to the less conserved P5 and P6 extensions. Notably, the internal loop between P4 and P5 (J4/5) contains highly conserved adenosines proposed to be close to the site of catalysis (Michel & Westhof, 1990; Wang & Cech, 1992). Using this RNA, we demonstrate that catalytic RNA can be composed of domains and subdomains that maintain tertiary structure independently.

MATERIALS AND METHODS

Materials. Moloney murine leukemia virus RNase H$^-$ reverse transcriptase (Superscript) was purchased from GIBCO BRL. Dithiothreitol (DTT) was purchased from ICN. T7 RNA polymerase was provided by Arthur J. Zaug.

Plasmid Construction. Plasmid pP4P6 which encodes the P4–P6 RNA was constructed from pTZL-21 (Grosshans & Cech, 1991) using the phagemid mutagenesis technique of Kunkel et al. (1987). Nucleotides 24–103 of the *Tetrahymena* L-21 *Sca*I ribozyme sequence were deleted. The sequence GAAG was inserted after nucleotide C 262 of the *Tetrahymena* sequence producing an *Ear*I site. Plasmid p5abc S encoding the ΔP5 RNA was a gift from Tan Inoue. The plasmids encoding J5/5a RNA and ΔA-rich bulge RNA, pBP55a and pDel2, respectively, were constructed using pP4P6 as the template for phagemid mutagenesis. Miniprep DNAs were prepared using the boiling method of Holmes and Quigley (1981). Mutations were confirmed by dideoxy sequencing.

RNA Preparation. Plasmids were digested with the appropriate restriction enzyme and transcribed using T7 RNA polymerase. Plasmid pP4P6 was digested with *Ear*I or *Fnu4H*I to generate P4–P6 RNA and ΔP6 RNA, respectively. Plasmid p5abc S was digested with *Sac*I to generate ΔP5 RNA. *Ear*I digestion was used for both pBP55a and pDel2. Transcription and gel purification were performed as described (Latham et al., 1990) except that 5% polyacrylamide/7 M urea gels were used. End-labeling of RNAs at their 5′ end with ^{32}P was performed as described (Latham et al., 1990). Following labeling, RNAs were again gel purified on 5% polyacrylamide/7 M urea gels.

Fe(II)–EDTA Reactions. Fe(II)–EDTA reactions were performed as described (Latham & Cech, 1989) with minor modifications. End-labeled RNA (400 000–500 000 cpm) was incubated at 50 °C for 10 min in 8 μL of 44 mM Tris-HCl, pH 7.5, in the presence or absence of $MgCl_2$. Preincubation in the presence of Mg^{2+} allows the ribozyme to adopt the active conformation (Herschlag & Cech, 1990). Following transfer to 42 °C for 5 min, 1 μL of Fe(II)–EDTA [10 mM $(NH_4)_2Fe(SO_4)_2$, 20 mM Na_2EDTA, pH 8] and then 1 μL of 50 mM DTT were added to a final sample volume of 10 μL. Thus, in addition to the RNA, the final solution contained 35 mM Tris-HCl, pH 7.5, 1 mM $(NH_4)_2Fe(SO_4)_2$, 2 mM

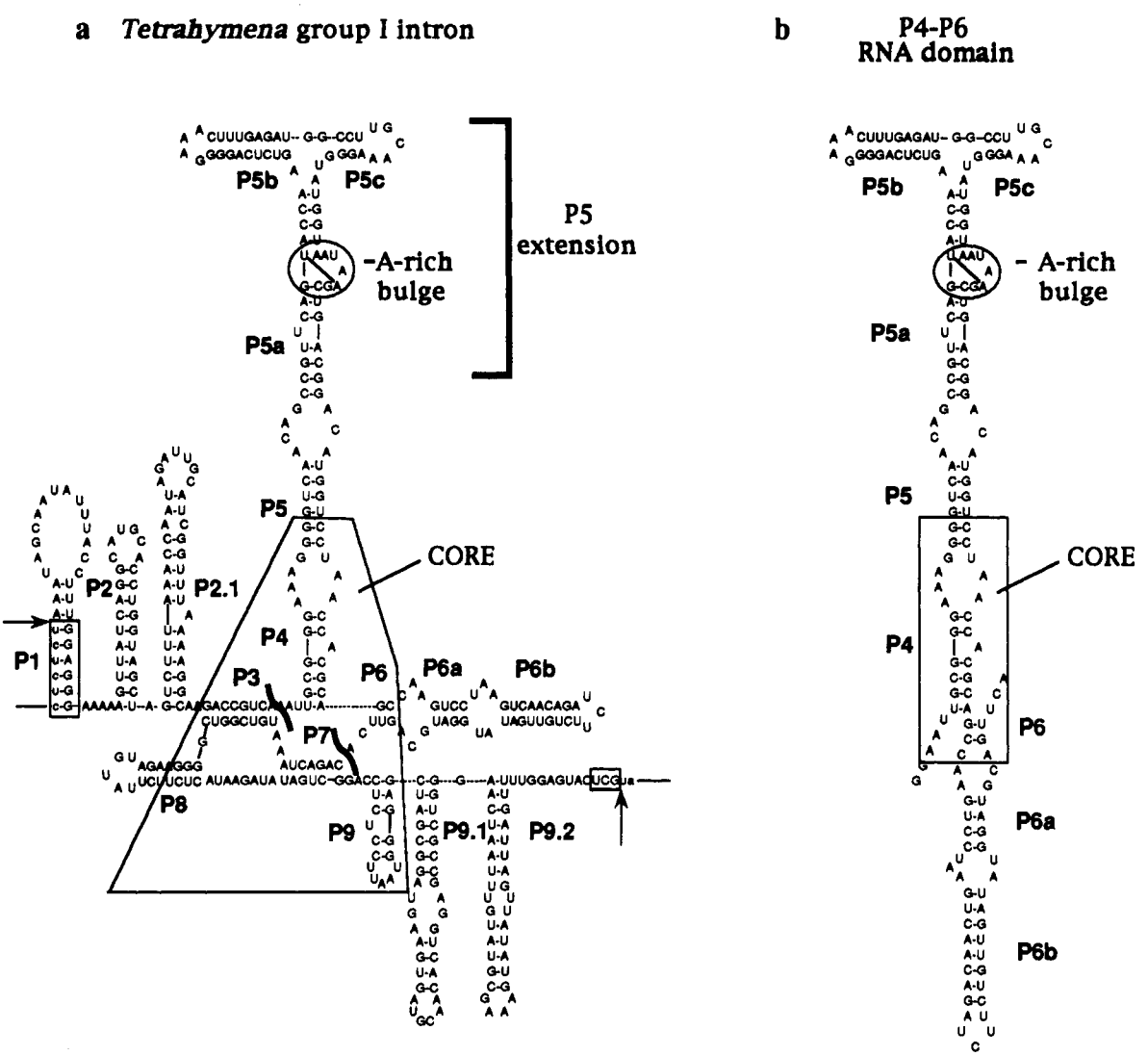

FIGURE 2: Location of the P4–P6 domain within the *Tetrahymena thermophila* group I intron. (a) Secondary structure of the intron. The P1 helix, enclosed in the rectangle, contains both exon (lowercase letters) and intron sequences. This helix is positioned in the core for the first step of self-splicing. Sequences contained within the highly conserved intron core are boxed. The remaining sequences comprise peripheral extensions that are not present in all introns or have variable structures among introns. The semiconserved A-rich bulge in the P5 extension is circled. The diagonal line connecting the U and the A represents a base pair (Michel & Westhof, 1990). Arrows indicate the 5′ and 3′ splice sites. Thick wavy lines highlight boundaries of domain sequence. (b) The RNA domain containing sequence elements P4–P6 (P4–P6 RNA). This sequence contains elements of the conserved core (boxed) and the semiconserved A-rich bulge (circled).

Na_2EDTA, pH 8, and 5 mM DTT with 13 mM $MgCl_2$ included where appropriate. Reactions were incubated at 42 °C for 80 min and then stopped with 10 mM thiourea. All reactants were prepared just before use. An equal volume of loading buffer (10 M urea, 10 mM Tris, 8.3 mM boric acid, 0.1 mM Na_2EDTA, 0.1% bromophenol blue, 0.1% xylene cyanole) was added and reactions were subjected to electrophoresis. For the Mg^{2+} titration, reaction conditions were altered to those of Celander and Cech (1991). Final concentrations of Tris-HCl, pH 7.5, and Na_2EDTA were decreased to 20 mM and 1.1 mM, respectively. Variable $MgCl_2$ was included in the reaction. The degree of protection from cleavage with the Fe(II)–EDTA probe as a function of Mg^{2+} concentration was quantitated using a PhosphorImager (Molecular Dynamics). Enzymatic sequencing lanes were prepared as described (Donis-Keller et al., 1977).

Dimethyl Sulfate Modification. DMS reactions were performed as described (Peattie, 1979; Inoue & Cech, 1985) with some modifications. A 200-μL volume containing 0.5 μg of RNA, 35 mM Tris-HCl, pH 7.5, and 13 mM $MgCl_2$ was preincubated as described above. After the 42 °C preincubation, DMS was diluted 1:4 in ethanol and added where appropriate to a final dilution of 1:800. Time points were taken between 5 and 30 min. Aliquots were stopped in $^1/_4$ volume 1 M β-mercaptoethanol and 1.5 M sodium acetate, pH 7. Carrier tRNA (10 μg) was added prior to precipitation with 3 volumes of ethanol. Pellets were resuspended in 200 μL of 0.3 M sodium acetate, pH 7, and reprecipitated with ethanol followed by a 70% ethanol wash. Dried samples were resuspended prior to reverse transcription. DMS was handled in the hood. Pipet tips and tubes were treated with 5 M NaOH prior to disposal.

Reverse Transcription. A 10-μL volume containing 0.2 pmol of DMS-treated RNA and 0.6 pmol of 5′-^{32}P-labeled complementary DNA primer was heated at 90 °C for 1.25 min in 1× annealing buffer (50 mM Tris-HCl, pH 8.3, 60 mM NaCl, 10 mM DTT; or 50 mM Tris-HCl alone can be used). The samples were slowly cooled over approximately 20 min to a temperature at or below 37 °C. Aliquots (2 μL) of the annealed mix were then combined at room temperature with 1 μL 1× annealing buffer; 2 μL of a solution containing 50 mM Tris-HCl, pH 8.3, 60 mM NaCl, 9.5 mM $MgCl_2$, 2.5 mM DTT, and 100 units/μL reverse transcriptase; and 1 μL of dNTP mix (2 mM dGTP, 2 mM dCTP, 2 mM dTTP, 2 mM dATP). Samples were incubated at 50 °C for 45 min at which time an equal volume of stop solution (95%

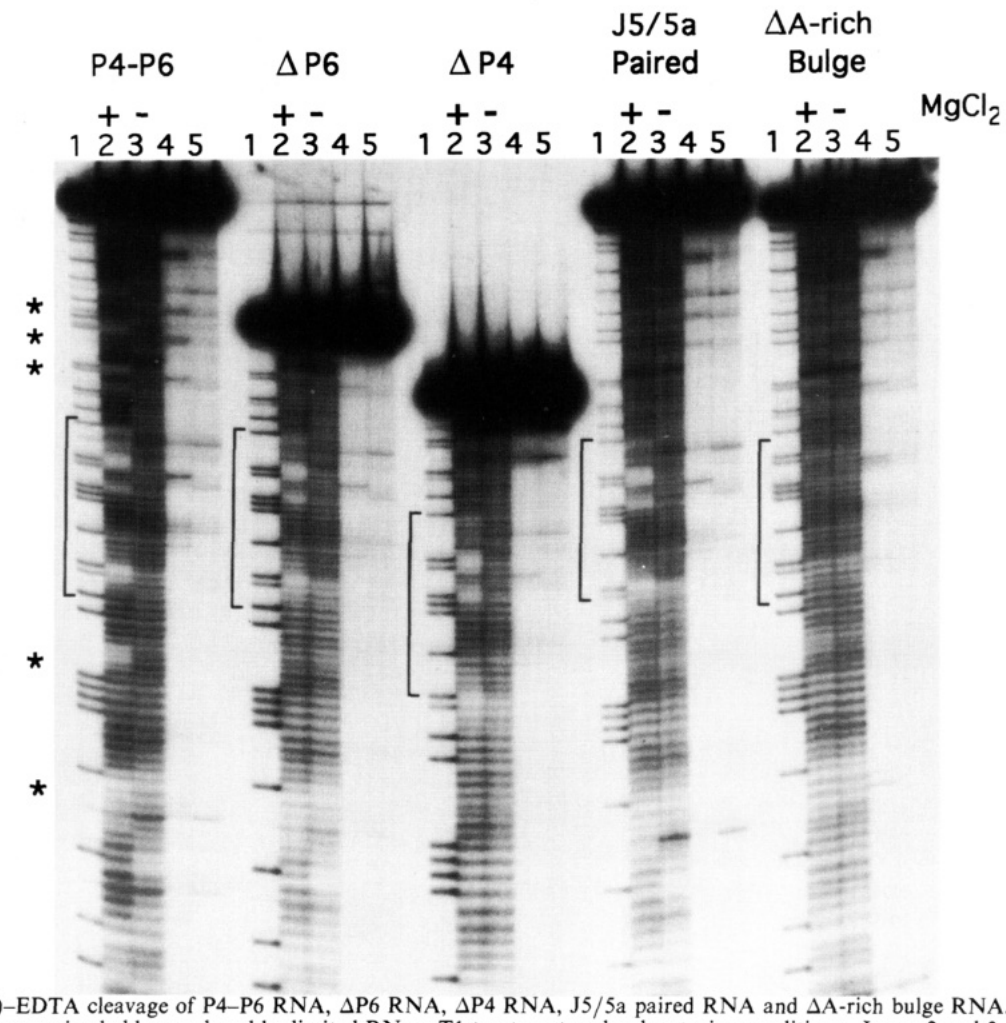

FIGURE 3: Fe(II)–EDTA cleavage of P4–P6 RNA, ΔP6 RNA, ΔP4 RNA, J5/5a paired RNA and ΔA-rich bulge RNA. RNA is 5′-^{32}P-labeled. Lane 1, guanosine ladder produced by limited RNase T1 treatment under denaturing conditions. Lanes 2 and 3, RNAs subjected to limited treatment with solvent-based Fe(II)–EDTA in the presence (+) or absence (–) of 13 mM Mg^{2+}. Lanes 4 and 5, mock-treated RNAs in the presence or absence of Mg^{2+}, respectively. P4–P6 RNA displays complete domain structure. The ΔP6, ΔP4, and J5/5a paired RNAs display subdomain structure. The ΔA-rich bulge RNA displays essentially no higher-order structure with Fe(II)–EDTA. Sites of protection corresponding to the subdomain structure are bracketed. Some sites of protection corresponding to the major substructure are denoted by an asterisk. An 8% 7 M urea gel is shown. Positions near the 5′ end were visualized using 20% gels.

formamide, 10 mM Tris, 8.3 mM boric acid, 0.1 mM Na_2-EDTA, 0.1% bromophenol blue, 0.1% xylene cyanole) was added. Samples were heated to 95 °C for 2 min prior to loading a 6% polyacrylamide/7 M urea gel.

RESULTS

Tertiary Structure of P4–P6 RNA Revealed Using Fe(II)–EDTA. The P4–P6 RNA molecule (Figure 2b) was 5′-end-labeled and subjected to limited reaction with the solvent-based Fe(II)–EDTA probe. This RNA exhibited an Fe(II)–EDTA protection pattern (Figure 3). As was the case for the intact ribozyme, the Fe(II)–EDTA pattern was dependent on Mg^{2+} (Figure 3) or Ca^{2+} (data not shown). In the absence of Mg^{2+}, a fairly uniform pattern of cleavage at every residue was observed. In the presence of Mg^{2+}, regions of protection from cleavage were observed. Ribose moieties that were inaccessible to the probe are highlighted on the secondary structure in Figure 4a. Protection can be thought of as an RNA footprint caused by intramolecular RNA–RNA interactions, analogous to a DNA footprint caused by DNA–protein interactions. If the P4–P6 RNA formed only the secondary structure shown in Figure 4a, no protection from the Fe(II)–EDTA probe would be expected.

The protected nucleotides observed in the P4–P6 RNA molecule are a large subset of those protected in the ribozyme (compare Figures 1 and 4a). In general, the pattern of protection at a given site covers fewer residues in the P4–P6 RNA than in the full-length ribozyme. Because the P4–P6 RNA maintains its native tertiary structure as an independent molecule, it constitutes an RNA structural domain.

The extent of protection of individual positions of the RNA was quantitated and is graphed as a function of Mg^{2+} concentration in Figure 5a. Tertiary structure formation by the P4–P6 RNA molecule, as assayed by a gain of protection from Fe(II)–EDTA cleavage, exhibits a Mg^{2+} concentration dependence similar to that reported previously for the intact ribozyme (Celander & Cech, 1991).

The Domain Is Composed of Two Substructures. What is required for domain structure? To address this question, the effect of 3′ deletion on domain structure was examined. When the P4–P6 domain molecule was truncated after nucleotide 213 (ΔP6 RNA), protection from Fe(II)–EDTA cleavage was lost at positions close to P6 in secondary structure, such as P4 and P5, as well as positions far from P6 in secondary structure, such as P5b (Figures 3 and 4b). The ΔP6 RNA retained protection in P5c and surrounding the A-rich bulge. This residual structure is a subset of the structure observed in the complete P4–P6 RNA domain molecule, suggesting that it is relevant to the structure of the intron. We refer to it as "subdomain" structure. Upon deletion of the P4 stem and alteration of the P5 stem sequence (ΔP4 RNA), protection

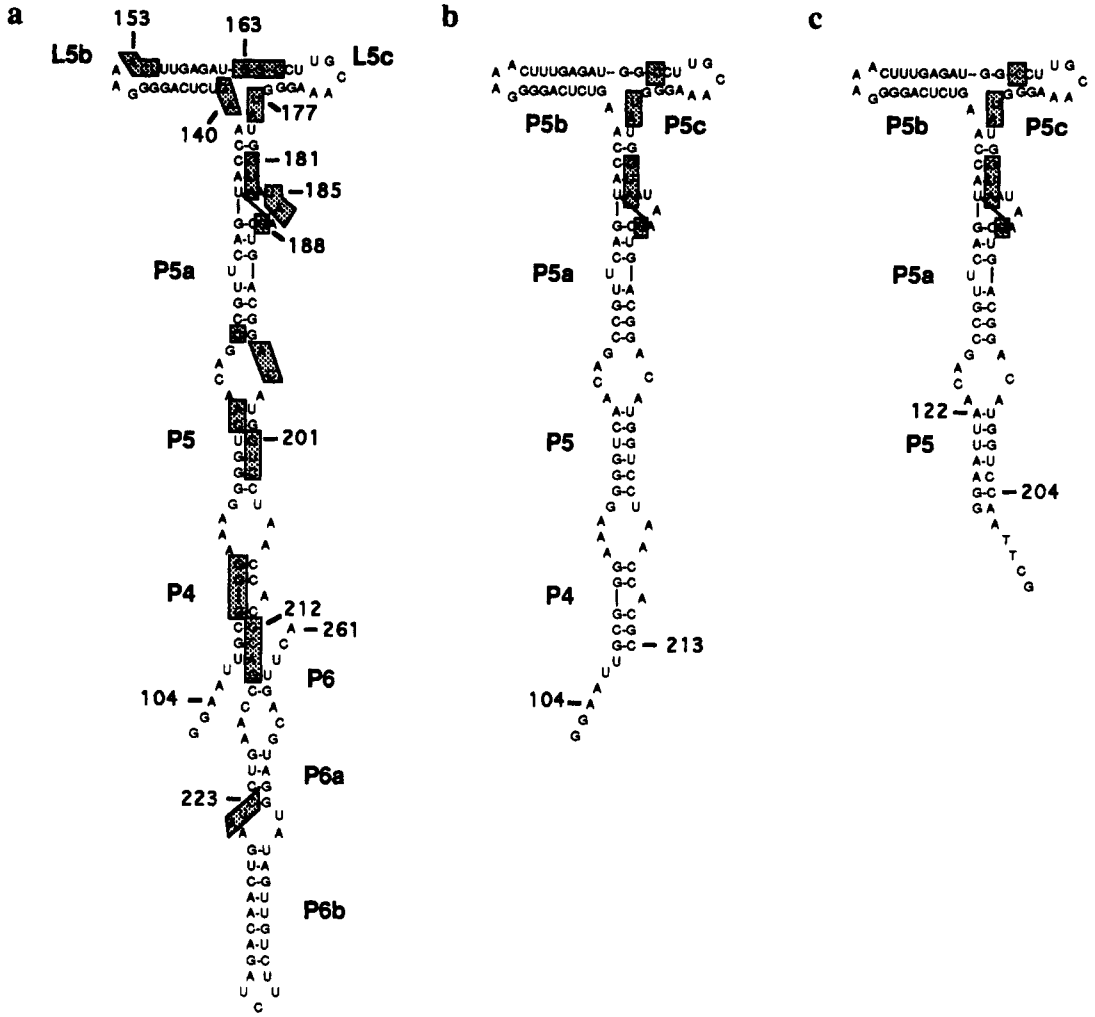

FIGURE 4: Secondary structures: Ribose moieties protected from cleavage with Fe(II)–EDTA are shaded. (a) P4–P6 RNA. (b) ΔP6 RNA. (c) ΔP4 RNA. The *Tetrahymena* intron numbering system is used for the shortened RNAs. Each numbered nucleotide within a shaded region denotes the 5′ nucleotide of a position quantitated in Figure 5. These figures were mapped by visual inspection and confirmed by PhosphorImager quantitation. They are estimated as accurate to plus or minus one nucleotide.

from Fe(II)–EDTA was lost at the same sites and again only the subdomain structure persisted (Figures 3 and 4c).

Therefore, the P4–P6 RNA domain can be divided into two substructures, correlating with a minor and a major component of the Fe(II)–EDTA pattern. The minor component correlates with the subdomain structure and does not require the P6 extension. The major component of the Fe(II)–EDTA pattern correlates with a second structural unit which is stabilized by the P6 extension. We refer to it as the major substructure because it involves the majority of the Fe(II)–EDTA pattern. Note that the major substructure includes Fe(II)–EDTA protection of core elements as well as protection in the nonconserved peripheral extensions. Since the major substructure includes protected regions at sites distant in the secondary structure, it appears to involve long-range interactions. While all sites of protection are Mg^{2+} dependent, the positions corresponding to the subdomain become protected at lower Mg^{2+} concentration than the remaining domain positions (Figure 5b,c).

Nuclease Cleavage and DMS Modification Confirm Structure Maintenance in the Shortened RNA Molecules. Nucleases were used as a second type of probe to assess the structural relevance of the P4–P6 RNA molecule and the deletion derivatives to the ribozyme structure. RNAs were subjected to mild digestion with the single-strand specific ribonucleases T1 and T2 as well as the double-strand-specific V1 nuclease. From comparison of the pattern of nuclease digestion observed in the ribozyme to that observed in the P4–P6 RNA, sites accessible to V1 and T2 nucleases remained unchanged. RNase T1 cleaved the ribozyme but outside of the portion comprising the domain; P4–P6 RNA was cleaved at one major site, following G169 (F. L. Murphy, Ph.D. Thesis). From these nuclease data, we conclude that the P4–P6 region has similar structure separately and within the ribozyme, with additional tertiary structure involving the G169 region lost in the P4–P6 RNA. In going from the P4–P6 RNA to shorter molecules, light cleavage was observed in the L5b loop with RNases T1 and T2. In ΔP4 RNA, the P5 stem appeared to be unstable since some cleavage was observed with all three nucleases (F. L. Murphy, Ph.D. Thesis). Thus, the deletion derivatives, ΔP6 and ΔP4, maintain a secondary structure similar to that of P4–P6 RNA and show some loss of higher-order structure involving the L5b loop.

Dimethyl sulfate (DMS) was used as a third type of structural probe to compare the P4–P6 RNA and the various deletion RNAs to the intact ribozyme. DMS modification followed by reverse transcription reveals the accessibility of the base-pairing faces of adenine (N_1) and cytosine (N_3) to methylation (Inoue & Cech, 1985; Moazed et al., 1986). Sites that were methylated in the ribozyme were also methylated in each of the shortened RNAs (Figure 6). The data are consistent with secondary structure maintenance in all of the molecules. With each deletion, new sites of modification appeared. These new modifications all occurred in single-stranded regions, consistent with a progressive loss of tertiary structure (Figure 6). Comparing the methylation pattern of

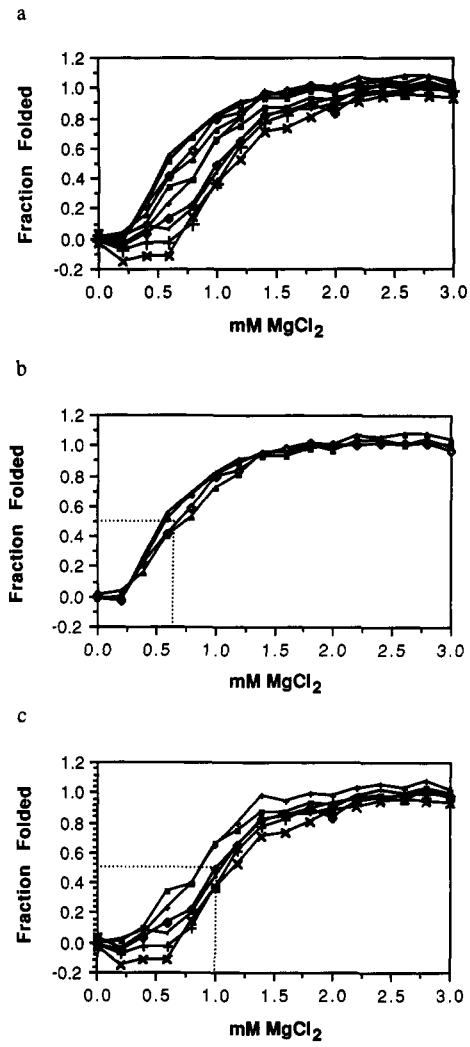

FIGURE 5: Mg^{2+} dependence of folding of P4–P6 RNA. (a) The relative extents of protection from Fe(II)–EDTA cleavage at 10 individual positions of protection (see Materials and Methods; Figure 4) were quantitated. (b) Only positions corresponding to the subdomain structure are shown (165, 177–178, 181–183, and 188). (c) Only positions corresponding to the major substructure are shown (140–141, 153–155, 185–186, 201–203, 212–215, 223–225). The two folding transitions at lower Mg^{2+} represent positions 185–186 and 140–141. 5′-^{32}P-Labeled RNAs were subjected to a limited treatment with Fe(II)–EDTA (see Materials and Methods) in the presence of various concentrations of $MgCl_2$ in the range 0–4 mM. After polyacrylamide gel electrophoresis, the dried gel was quantified using a Molecular Dynamics PhosphorImager. PhosphorImager counts at each position were quantitated (by volume integration) at each $MgCl_2$ concentration. Mock-treated RNA was used to assess background values at each position, and these were subtracted. Lanes were normalized using nonchanging regions; 0 mM $MgCl_2$ was used as the fully unfolded reference, and 4 mM $MgCl_2$ was used as the fully folded reference. No further increase in the extent of protection was observed with increased $MgCl_2$ up to 50 mM. Fraction folded for a given position equals $(V_0 - V_n)/(V_0 - V_4)$ where V_0 is the volume integral at 0 mM $MgCl_2$, V_4 is the volume integral at 4 mM $MgCl_2$, and V_n is the volume integral at each $MgCl_2$ concentration n. Variation in values obtained for fraction folded in independent experiments was less than 10%. Extent of protection, V_0/V_4, for positions 140–141, 153–155, 165, 177–178, 181–183, 185–186, 188, 201–203, 212–215, and 223–225 was 2.5, 2.6, 2.6, 2.5, 5.2, 3.1, 2.6, 2.3, 2.6, and 1.8, respectively. Variation in values for extent of protection in two independent experiments was less than 30%.

the ribozyme to P4–P6 RNA, the A-rich bulge and the GAAA loop of L5b were protected from methylation in both molecules. This suggests that the A-rich bulge and the GAAA loop are involved in tertiary structure in both the ribozyme and the P4–P6 RNA molecule. Protection from methylation may be due to a direct interaction such as a hydrogen bond or an indirect steric occlusion. In contrast, the L5c loop, which was protected from methylation in the ribozyme, became accessible to methylation in the P4–P6 RNA. This is consistent with the gain of an RNase T1 cut in the L5c loop of P4–P6 RNA and indicates that structure has been lost in this loop. The L5c loop may interact with a part of the ribozyme that has been deleted in the P4–P6 RNA. Alternatively, L5c could interact within the 160 nucleotides comprising the P4–P6 region but no longer be stable in the isolated P4–P6 RNA molecule.

Comparing the methylation pattern of the P4–P6 RNA to ΔP6 and ΔP4, all sites of methylation in the P4–P6 RNA were maintained in the smaller RNAs. Notably, upon deletion of P6 the distant GAAA loop in L5b became accessible to methylation. This indicates that the GAAA loop may be involved in a tertiary interaction in the major substructure of the domain.

Hypothesis of a Bend in J5/5a. Extensive loss of tertiary structure upon deletion of the P6 extension, indicated by both the Fe(II)–EDTA and DMS assays, suggested a long-range interaction involving the P6 extension and the P5 extension. In order for these distant sites to interact, there would have to be one or more sites of bending in the secondary structure in Figure 4a. The bend could be rigid, or there could be a flexible hinge which is locked into place by tertiary interactions removed from the site of bending. Where is such a bend likely to occur? The junction between P4 and P6 is part of the conserved core and has been proposed to coaxially stack (Kim & Cech, 1987; Michel et al., 1990). It is therefore unlikely to mediate bending. Two components outside of the core in good position to facilitate bending are the conserved A-rich bulge and the internal loop between P5 and P5a (J5/5a).

As a test of the model that bending occurs in J5/5a, this internal loop was converted into a duplex by changing bases on one side to be complementary to those of the other (Figure 7a). If this internal loop were involved in a bend required for interaction between the P6 extension and the P5 extension, the major substructure should be lost in the mutant. As shown in Figure 3, this was the case; only subdomain structure was retained in the J5/5a mutant as assessed by Fe(II)–EDTA cleavage. Furthermore, DMS treatment of the J5/5a mutant showed a dramatic increase in the accessibility of the GAAA loop (Figure 7b), also indicative of loss of the major substructure. Several additional changes in the methylation pattern were observed in the J5/5a mutant. The internal loop between P5 and P5a (J5/5a) was no longer accessible to DMS, confirming its helical conformation. In addition, new modifications were observed in J6a/6b and in the A-rich bulge. These could not be monitored in the previous experiments because J6a/6b is not present in the deletion RNAs, while the A-rich bulge is present but is not observed due to the placement of the primer. Bases that became accessible to DMS are implicated in hydrogen-bonding interactions in the major substructure. It should be emphasized that only a single functional group on A and C bases is observable with this method, and it is likely that other structural changes have also occurred. No additional protection is recovered in the mutant with increased Mg^{2+} concentration up to 50 mM (data not shown).

The phylogenetic conservation of the A-rich bulge suggests an important role in the structure and function of the P5 extension. The asymmetric bulge is likely to contribute to bending (Bhattacharyya et al., 1990). Furthermore, it has been suggested that the A-rich bulge may contact the core in P4 (Flor et al., 1989). In order to assess the role of the A-rich

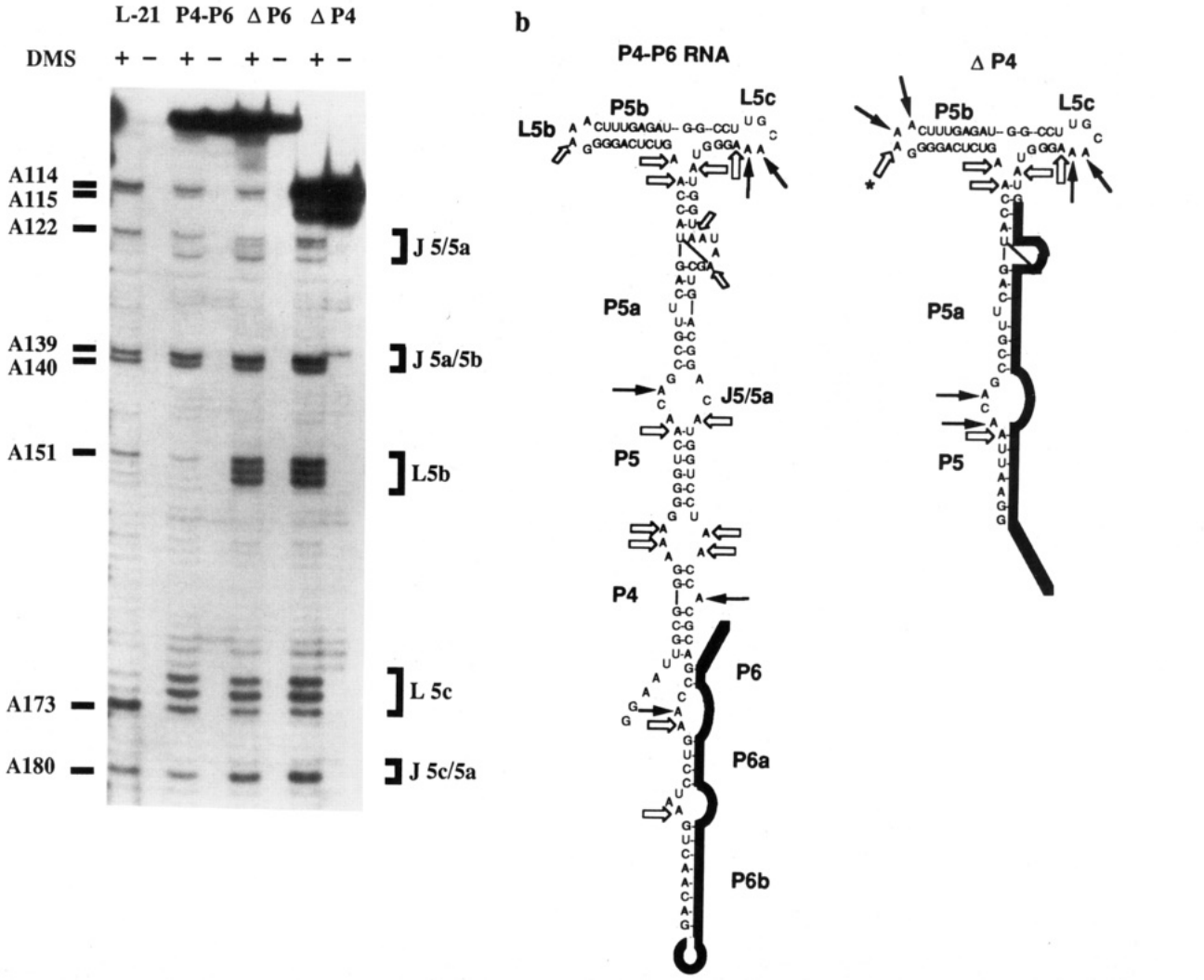

FIGURE 6: DMS modification of P4–P6 RNA and deletion RNAs. (a, left) The RNAs were incubated with (+) or without (–) DMS and then reversed transcribed using a 5′-^{32}P-labeled primer (see Materials and Methods). L-21 denotes the L-21 *Sca*I ribozyme. Samples were run next to a dideoxy sequencing ladder (not shown). Primer extension stops one nucleotide prior to modified base (Inoue & Cech, 1985). (b, right) Sites of modification are mapped onto secondary structures. Open arrows denote modifications observed in the ribozyme and in the shortened RNAs. Solid arrows denote modifications that were not present in the ribozyme but were observed in the respective shortened molecule. The same pattern of modifications was observed for ΔP4 and ΔP5. The asterisk (*) denotes a site of enhanced modification in the shortened RNAs. Notably, a reverse transcriptase stop was seen for C170 with moloney murine leukemia virus and avian myeloblastosis virus reverse transcriptases but not with Superscript. No others cytosines were detected under our conditions. The solid line represents sequences for which no modification data were obtained due to placement of the primer. Mapping was based on consistent results from three independent experiments.

bulge in domain structure, a mutant domain molecule was constructed in which the A-rich bulge was deleted (ΔA-rich bulge; Figure 7a). The mutant showed a complete loss of structure by the criterion of Fe(II)–EDTA protection (Figure 3). DMS modification showed that the GAAA loop and J6a/6b were methylated and that the secondary structure was intact (Figure 7b). Therefore, the A-rich bulge is required for subdomain structure, and loss of subdomain structure results in loss of the remainder of the higher-order domain structure.

DISCUSSION

We demonstrate that the *Tetrahymena* ribozyme contains an independently folding domain of RNA tertiary structure (P4–P6 RNA; Figure 2b). The pattern of protection from Fe(II)–EDTA cleavage observed in the domain is a subset of that observed in the catalytically active ribozyme (Figures 1, 3, and 4). Corroborating evidence for maintenance of tertiary structure in the P4–P6 domain molecule comes from chemical and nuclease probing experiments (Figure 6 and F. L. Murphy, Ph.D. Thesis). Because deletion of the P6 extension (P6, P6a, and P6b) or conversion of the J5/5a internal loop to an RNA duplex caused loss of protection at distant sites in the secondary structure, the tertiary structure of the domain appears to involve long-range contacts.

It is well-known that pieces of larger RNA molecules can maintain their secondary structures [for example, Kean and Draper (1985), Mougel et al. (1988), Leffers et al. (1988), and Weeks et al. (1990)] or assemble using secondary structure interactions [for example, Chambers (1971), Szostak (1986), Uhlenbeck (1987), and Goldschmidt-Clermont et al. (1991)]. RNA fragments can maintain local base stacking or noncanonical hydrogen bonding as demonstrated in isolated RNA terminal loops (Cheong et al., 1990; Heus & Pardi, 1991) and internal loops (Brunel et al., 1990; Chow et al., 1992a). The extent to which a structured RNA can be divided into pieces which maintain long-range *tertiary* structure is less well-known. Ribosomal RNAs are interesting in that the secondary structure domains defined phylogenetically (Noller & Woese, 1981) are represented as domains of higher-order ribonucleoprotein in models of the small subunit rRNA (Stern et al., 1988; Brimacombe et al., 1988). Fragments of RNA

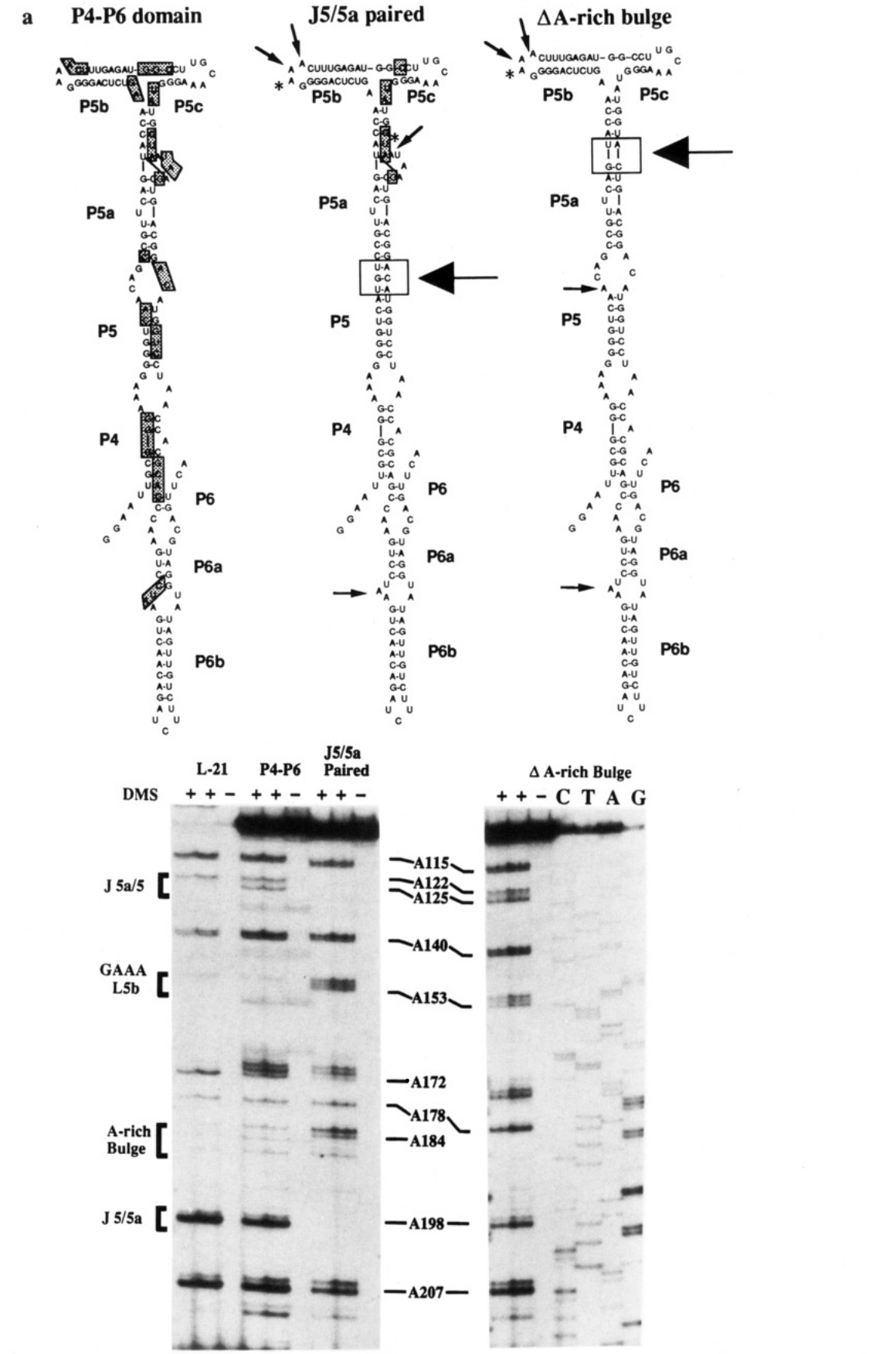

FIGURE 7: Analysis of mutations supporting the model of a bend in J5/5a and a hierarchical relationship between the subdomain structure and the major substructure. (a, top) Summary of chemical probing data of J5/5a and ΔA-rich bulge RNAs relative to P4–P6 RNA is displayed on secondary structures. Sites of mutation are boxed and denoted by large arrow. Shading represents sites of protection from Fe(II)–EDTA cleavage. Small arrows denote new DMS modifications observed in mutants but not in P4–P6 RNA. Asterisks denote enhanced modification relative to the P4–P6 RNA. All modifications observed in P4–P6 RNA are maintained in the mutants with the following exceptions. No methylation occurs in J5/5a when this has been converted to a duplex in the J5/5a paired mutant. No modification occurs in the A-rich bulge since this sequence has been deleted. In addition, A173 is protected from methylation in the ΔA-rich bulge mutant, suggesting that this base pair is destabilized in the subdomain structure. (b, bottom) DMS modification data for the intact ribozyme (L-21), P4–P6 RNA, J5/5a paired, and the ΔA-rich bulge as described in Figure 6.

corresponding to ribosomal domains can bind ribosomal proteins, and data such as limited protection from chemical modification or resistance to nucleases suggest that a degree of higher-order structure is observed in some of these naked RNAs (Garrett et al., 1984; Kean & Draper, 1985; Noller, 1991).

The power of the Fe(II)–EDTA probe used here is that it monitors every ribose moiety for teritary structure, apparently on the basis of solvent accessibility. This gives a more global view of the molecule's higher-order structure than that of functional group probes (Ehresmann et al., 1987) or shape-specific probes (Chow et al., 1992b). The Mg^{2+} dependence of folding, as visualized by gain of protection from Fe(II)–EDTA cleavage, is essentially the same for the P4–P6 domain molecule (Figure 5) and the active form of the ribozyme (Celander & Cech, 1991). This is perhaps surprising since increased Mg^{2+} concentrations are frequently required to overcome structural defects in mutant introns (Burke et al., 1986; Flor et al., 1989; Michel et al., 1992). The lack of an increased Mg^{2+} requirement reinforces that the domain is not defective but is an autonomous structure. Regions of protection from Fe(II)–EDTA tend to span more nucleotides in the intact ribozyme than in the isolated domain (compare Figures 1 and 4a). This could be due to increased conformational flexibility of the domain molecule relative to the ribozyme or to increased domain surface area due to loss of interactions between the domain and the rest of the ribozyme.

Site-directed mutagenesis of the domain molecule in combination with chemical probing has led to a model for the domain's structural organization, as follows. (1) There is a major substructure that involves a direct long-range interaction between the P6 extension and L5b. The GAAA tetraloop of L5b becomes accessible to DMS (Figure 6) upon loss of the major substructure. This suggests that the substructure involves a tertiary interaction to the GAAA loop. More recently, site-directed mutagenesis of this loop has shown that it makes a direct rather than an indirect contribution to domain structure (F. L. Murphy and T. R. Cech, unpublished results). (2) The J5/5a internal loop is one site of bending that facilitates formation of the domain structure. When the J5/5a internal loop is replaced with a Watson–Crick duplex, the major substructure is lost. In contrast to other mutants that disrupt structure, this mutant phenotype is not suppressed by increased Mg^{2+} concentration, consistent with loss of ability to bend in J5/5a. An ultraviolet light induced cross-link between A122 and A196 in the intact *Tetrahymena* intron is also consistent with the hypothesis of a bend at this site (W. Downs and T. R. Cech, unpublished results).

Within the domain is a subdomain, the folding of which requires the A-rich bulge. The number of protected nucleotides in the subdomain is comparable to that in the T7 transcript of yeast $tRNA^{Phe}$ (Latham & Cech, 1989). In this tRNA, about half of the residues protected from Fe(II)–EDTA cleavage are directly involved in a tertiary interaction as determined by comparison with the known X-ray structure. If the rules governing this tRNA are generalizable, then the subdomain may have an amount of higher-order structure similar to that of tRNA. The Fe(II)–EDTA protection associated with the subdomain is observed at a slightly lower Mg^{2+} concentration than the remaining sites of domain protection (Figure 5). This further emphasizes the autonomy of the subdomain structure.

There is a hierarchical relationship between the subdomain and major substructure. Disruption of the major substructure of the domain does not affect the subdomain structure, whereas disruption of the subdomain structure disrupts the major substructure. Although our data do not measure the kinetics of folding, this hierarchical relationship suggsts that the subdomain structure must be in place for the remainder of the domain structure to form. These results emphasize the modular organization of this RNA structure and suggest the possibility that highly structured RNAs may, in general, be assembled beginning with smaller tertiary units.

We propose that the subdomain has been maintained phylogenetically in the *Tetrahymena* intron and related introns because it nucleates folding of the larger domain. It could contribute to a bend that helps to mediate the major substructure or otherwise anchor the remainder of the domain structure. In the context of the full-length intron, the domain may act to stabilize the intron core. This is consistent with the observation that deletion of the P5 extension from the *Tetrahymena* intron increases the Mg^{2+} concentration requirement for activity (van der Horst et al., 1991; Joyce et al., 1989).

Division of RNA into domains not only gives us insight into the structural organization of RNA, but on a practical level it provides simplified units for structural studies. The pattern of protection within the domain can now be assigned to a tertiary structure occurring within 160 nucleotides instead of 395. The solvent-based Fe(II)–EDTA probe, in combination with manipulation of primary and secondary structure, should be useful for identifying domains and subdomains in other RNAs or RNA fragments.

ACKNOWLEDGMENT

We thank Tan Inoue for providing the p5abc S construct and Arthur J. Zaug for preparing the pP4P6. We thank Alexandra Gampel for useful discussions.

REFERENCES

Altman, S. (1990) *J. Biol. Chem. 265*, 20053–20056.
Benner, S. A., Ellington, A. D., & Tauer, A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7054–7058.
Bhattacharyya, A., Murchie, A. I. H., & Lilley, D. M. (1990) *Nature 343*, 484–487.
Brimacombe, R., Atmadja, J., Stiege, W., & Schuler, D. (1988) *J. Mol. Biol. 199*, 115–136.
Brunel, C., Romby, P., Westhof, E., Romaniuk, P. J., Ehresmann, B., & Ehresmann, C. (1990) *J. Mol. Biol. 215*, 103–111.
Burke, J. M., Irvine, K. D., Kaneko, K. J., Kerker, B. J., Oettgen, A. B., Tierney, W. M., Williamson, C. L., Zaug, A. J., & Cech, T. R. (1986) *Cell 45*, 167–176.
Burke, J. M., Belfort, M., Cech, T. R., Davies, R. W., Schweyen, R. J., Shub, D. A., Szostak, J. W., & Tabak, H. F. (1987) *Nucleic Acids Res. 15*, 7217–7221.
Brukhoff, A. M., & Tullius, T. D. (1987) *Cell 48*, 935–943.
Cech, T. R. (1987) *Science 236*, 1532–1539.
Cech, T. R. (1990) *Annu. Rev. Biochem. 59*, 543–568.
Celander, D. W., & Cech, T. R. (1990) *Biochemistry 29*, 1355–1361.
Celander, D. W., & Cech, T. R. (1991) *Science 251*, 401–407.
Chambers, R. W. (1971) in *Progress in Nucleic Acid Research and Molecular Biology* (Davidson, J. N., & Cohn, W. E., Eds.) pp 489–525, Academic Press, New York and London.
Cheong, C., Varani, G., & Tinoco, I., Jr. (1990) Nature *346*, 680–682.
Chow, C. S., Hartmann, K. M., Rawlings, S. L., Huber, P. W., & Barton, J. K. (1992a) *Biochemistry 31*, 3534–3542.
Chow, C. S., Behlen, L. S., Uhlenbeck, O. C., & Barton, J. K. (1992b) *Biochemistry 31*, 972–982.
Collins, R. A. (1988) *Nucleic Acids Res. 16*, 2705–2715.
Donis-Keller, H., Maxam, A. M., & Gilbert, W. (1977) *Nucleic Acids Res. 4*, 2527–2538.
Doudna, J. A., & Szostak, J. W. (1989) *Mol. Cell Biol. 9*, 5480–5483.

Ehresmann, C., Baudin, F., Mougel, M., Romby, P., Ebel, J.-P., & Ehresmann, B. (1987) *Nucleic Acids Res. 15*, 9109–9129.
Flor, P. J., Flanegan, J. B., & Cech, T. R. (1989) *EMBO J. 8*, 3391–3399.
Garrett, R. A., Christensen, A., & Douthwaite, S. (1984) *J. Mol. Biol. 179*, 689–712.
Goldschmidt-Clermont, M., Choquet, Y., Girard-Bascou, J., Michel, F., Schirmer-Rahire, M., & Rochaix, J.-D. (1991) *Cell 65*, 135–143.
Grosshans, C. A., & Cech, T. R. (1989) *Biochemistry 28*, 6888–6894.
Grosshans, C. A., & Cech, T. R. (1991) *Nucleic Acids Res. 19*, 3875–3880.
Guo, Q., & Lambowitz, A. M. (1992) *Genes Dev. 6*, 1357–1372.
Herschlag, D., & Cech, T. R. (1990) *Biochemistry 29*, 10159.
Hertzberg, R., & Dervan, P. B. (1984) *Biochemistry 23*, 3934–3945.
Heuer, T. S., Chandry, P. S., Belfort, M., Celander, D. W., & Cech, T. R. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 11105–11109.
Heus, H. A., & Pardi, A. (1991) *Science 253*, 191–193.
Holmes, D. S., & Quigley, M. (1981) *Anal. Biochem. 114*, 193–197.
Inoue, T., & Cech, T. R. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 648–652.
Joyce, G. F., van der Horst, G., & Inoue, T. (1989) *Nucleic Acids Res. 17*, 7879–7889.
Kean, J. M., & Draper, D. E. (1985) *Biochemistry 24*, 5052–5061.
Kim, S.-H., & Cech, T. R. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8788–8792.
Kunkel, T. A., Roberts, J. D., & Zakour, R. A. (1987) *Methods Enzymol. 154*, 367–382.
Lambowitz, A. M., & Perlman, P. S. (1990) *Trends Biochem. Sci. 15*, 440–444.
Latham, J. A., & Cech, T. R. (1989) *Science 245*, 276–282.
Latham, J. A., Zaug, A. J., & Cech, T. R. (1990) *Methods Enzymol. 181*, 558–569.
Leffers, H., Egebjerg, J., Andersen, A., Christensen, T., & Garrett, R. A. (1988) *J. Mol. Biol. 204*, 507–522.
Michel, F., & Dujon, B. (1983) *EMBO J. 2*, 33–38.
Michel, F., & Cummings, D. J. (1985) *Curr. Genet. 10*, 69–79.
Michel, F., & Westhof, E. (1990) *J. Mol. Biol. 216*, 585–610.
Michel, F., Umesono, K., & Ozeki, H. (1989) *Gene 82*, 5–30.
Michel, F., Ellington, A. F., Couture, S., & Szostak, J. W. (1990) *Nature 347*, 578–580.
Michel, F., Jaeger, L., Westhof, E., Kuras, R., Tihy, F., Xu, M.-Q., & Shub, D. A. (1992) *Genes Dev. 6*, 1373–1385.
Moazed, D., Stern, S., & Noller, H. F. (1986) *J. Mol. Biol. 187*, 399–416.
Mougel, M., Philippe, C., Ebel, J.-P., Ehresmann, B., & Ehresmann, C. (1988) *Nucleic Acids Res. 16*, 2825–2839.
Noller, H. F. (1991) *Annu. Rev. Biochem. 60*, 191–227.
Noller, H. F., & Woese, C. R. (1981) *Science 212*, 403–411.
Noller, H. F., Hoffarth, V., & Zimniak, L. (1992) *Science 256*, 1416–1419.
Pace, N. R., & Smith, D. (1990) *J. Biol. Chem. 265*, 3587.
Peattie, D. A. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1760–1764.
Piccirilli, J. A., McConnell, T. S., Zaug, A. J., Noller, H. F., & Cech, T. R. (1992) *Science 256*, 1420–1424.
Piccirilli, J. A., Vyle, J. S., Caruthers, M. H., & Cech, T. R. (1993) *Nature 361*, 85–88.
Price, J. V., Kieft, G. L., Kent, J. R., Sievers, E. L., & Cech, T. R. (1985) *Nucleic Acids Res. 13*, 1871–1889.
Rahhal, S., & Richter, H. W. (1988) *J. Am. Chem. Soc. 110*, 3126–3133.
Rush, J. D., & Koppenol, W. H. (1986) *J. Biol. Chem. 261*, 6730–6733.
Stern, S., Weiser, B., & Noller, H. F. (1988) *J. Mol. Biol. 204*, 447–481.
Symons, R. H. (1992) *Annu. Rev. Biochem. 55*, 641–671.
Szostak, J. W. (1986) *Nature 322*, 83–86.
Tullius, T. D., & Dombroski, B. A. (1985) *Science 230*, 679–681.
Tullius, T. D., & Dombroski, B. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5469–5473.
Uhlenbeck, O. C. (1987) *Nature 328*, 596–600.
Van der Horst, G., Christian, A., & Inoue, T. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 184–188.
Visser, C. M., & Kellogg, R. M. (1978) *J. Mol. Evol. 11*, 171.
Wang, J.-F., & Cech, T. R. (1992) *Science 256*, 526–529.
Waring, R. B., Scazzocchio, C., Brown, T. A., & Davies, R. W. (1983) *J. Mol. Biol. 167*, 595–605.
Weeks, K. M., Ampe, C., Schultz, S. C., Steitz, T. A., & Crothers, D. M. (1990) *Science 249*, 1281–1285.
Xu, M. Q., Kathe, S. D., Goodrich-Blair, H., Nierzwicki-Bauer, S. A., & Shub, D. A. (1990) *Science 250*, 1566–1570.
Zaug, A. J., Been, M. D., & Cech, T. R. (1986) *Nature 324*, 429–433.