Assembly of an Exceptionally Stable RNA Tertiary Interface in a Group I Ribozyme†

Elizabeth A. Doherty,‡ Daniel Herschlag,|| and Jennifer A. Doudna*,‡,§

Department of Molecular Biophysics and Biochemistry, 260/266 Whitney Avenue, P.O. Box 208114, and Howard Hughes Medical Institute, Yale University, New Haven, Connecticut 06520, and Department of Biochemistry, B400 Beckman Center, Stanford University, Stanford, California 94305

Received August 31, 1998; Revised Manuscript Received December 17, 1998

ABSTRACT: Group I intron RNAs contain a core of highly conserved helices flanked by peripheral domains that stabilize the core structure. In the *Tetrahymena* group I ribozyme, the P4, P5, and P6 helices of the core pack tightly against a three-helix subdomain called P5abc. Chemical footprinting and the crystal structure of the *Tetrahymena* intron P4−P6 domain revealed that tertiary interactions between these two parts of the domain create an extensive solvent-inaccessible interface. We have examined the formation and stability of this tertiary interface by providing the P5abc segment in trans to a *Tetrahymena* ribozyme construct that lacks P5abc ($E^{\Delta P5abc}$). Equilibrium gel shift experiments show that the affinity of the P5abc and $E^{\Delta P5abc}$ RNAs is exceptionally strong, with a K_d of ∼100 pM at 10 mM $MgCl_2$ (at 37 °C). Chemical and enzymatic footprinting shows that the RNAs are substantially folded prior to assembly of the complex. Solvent accessibility mapping reveals that, in the absence of P5abc, the intron RNA maintains a nativelike fold but its active-site helices are not tightly packed. Upon binding of P5abc, the catalytic core becomes more tightly packed through indirect effects of the tertiary interface formation. This two-component system facilitates quantitative examination of individual tertiary contacts that stabilize the folded intron.

Large RNA molecules, like proteins, readily form specific molecular shapes adapted for ligand binding and catalysis. Catalytic RNAs including group I and group II introns and RNase P have compact interiors, involving close association of several segments of secondary structure. Close packing of RNA helices requires extensive screening of the phosphodiester backbone by cations in solution (*1*, *2*). Furthermore, noncanonical base pairs and hairpin loops provide distinct arrays of hydrogen bond donors and acceptors and irregular surface features that serve as recognition sites for RNA helices in the formation of higher order structures (*3*−*8*). How tertiary contacts, charge screening, and other factors stabilize RNA helix packing is not well understood.

The X-ray crystal structure of the P4−P6 domain from the *Tetrahymena thermophila* group I intron provided the first detailed view of an extended tertiary interface in RNA (*9*). In the 160 nucleotide domain, a flexible internal loop, J5/5a, allows the backbone to make a ∼180° bend, enabling parallel association of two coaxially stacked helical segments (Figure 1). One-half of the P4−P6 domain, a three-helix subdomain called P5abc, binds at two distinct locations to the other half of the domain, constituted by the coaxially stacked P5, P4, and P6 helices (Figure 1). The A-rich bulge of P5abc contacts the minor groove of helix P4 by backbone and base-mediated hydrogen bonds. Additionally, a GAAA tetraloop in P5abc docks into its tetraloop receptor by both cross-helical base stacking and extensive hydrogen bonding. Similar GAAA tetraloop/tetraloop receptor interactions are thought to occur in other large RNAs including group I and group II introns and ribosomal RNA (*5*). Thus, the P4−P6 domain crystal structure provides a basis for examining the formation and stability of tertiary interactions that characterize helix packing in RNA.

While the P5abc subdomain is not universally conserved in group I introns (*10*), its deletion results in a large decrease in self-splicing, suggesting that P5abc plays an important functional role (*11*, *12*). Chemical footprinting studies showed that P5abc independently forms a highly stable native structure (*13*). The crystal structure of the P4−P6 domain further revealed that this subdomain is organized around a core of divalent cations (*9*, *14*). Kinetic folding experiments using time-resolved hydroxyl radical cleavage showed that P5abc is the first part of the ribozyme to fold (*15*). The importance of P5abc is further highlighted by the fact that in some introns it is replaced by a protein cofactor (*16*).

To investigate helix packing in the *Tetrahymena* ribozyme, we have deleted P5abc from the ribozyme ($E^{\Delta P5abc}$) and provided it in trans, following the approach of Inoue and co-workers (Figures 1 and 2; refs *11* and *12*). This system allows investigation of the strength and characteristics of tertiary interactions that cannot be easily examined in the context of the intact ribozyme. Association of P5abc and $E^{\Delta P5abc}$ recreates the helix−helix interface of the P4−P6 domain through tertiary contacts between the GAAA tetra-

† This work was supported by grants from the Searle Scholars Program, the David and Lucile Packard Foundation, and NIH Grant GM56361 to J.A.D. and NIH Grant GM49243 to D.H. E.A.D. was supported in part by NIH training Grant 5T32 GM07223.

* To whom correspondence should be addressed. Phone: (203) 432-3108. Fax: (203) 432-3104. E-mail: doudna@csb.yale.edu.

‡ Department of Molecular Biophysics and Biochemistry, Yale Univ.

|| Department of Biochemistry, Stanford Univ.

§ Howard Hughes Medical Institute.

10.1021/bi982113p CCC: $18.00
Published on Web 02/17/1999

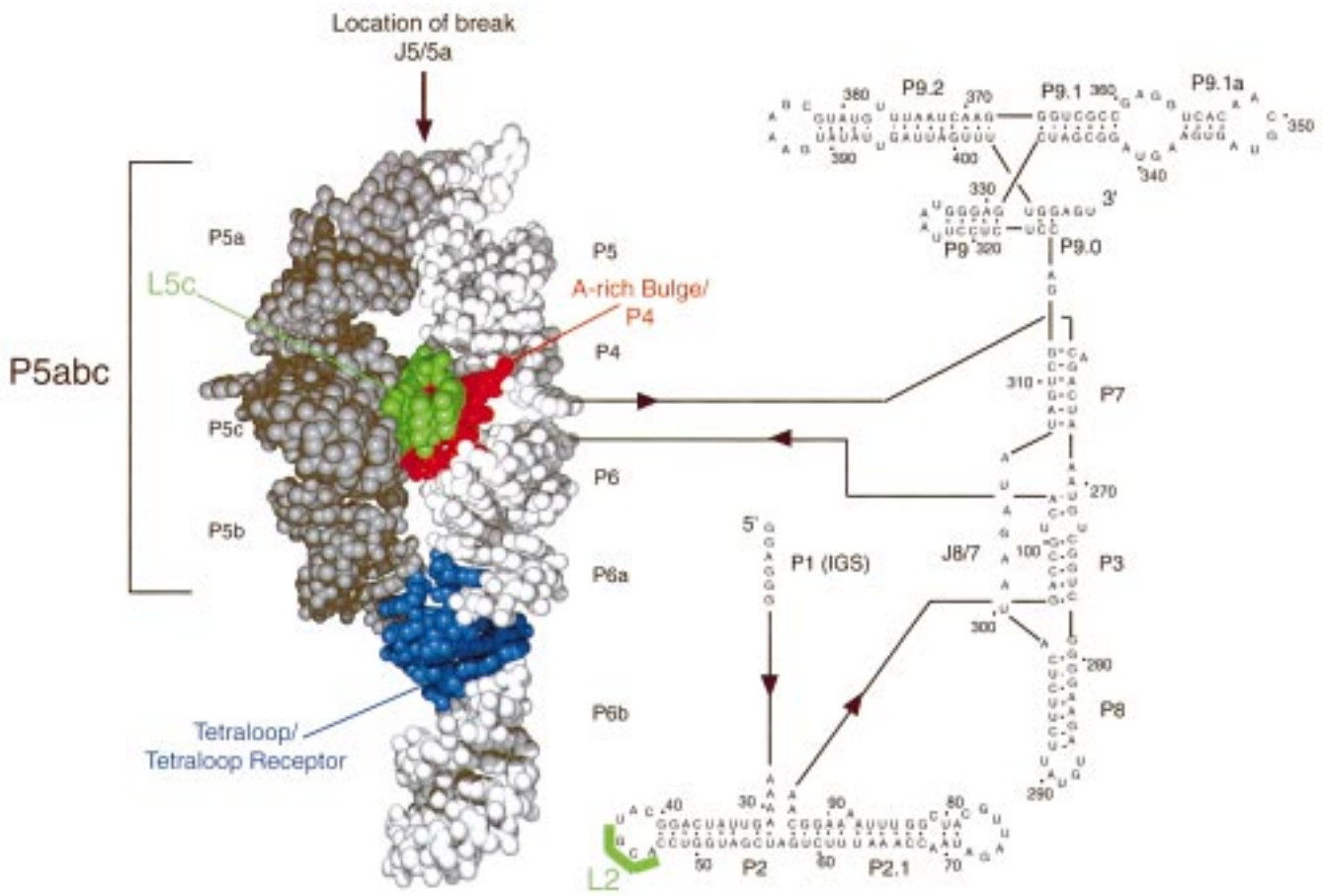

FIGURE 1: Diagram of the L-21 *Sca I Tetrahymena* ribozyme showing a space-filling representation of the 2.8 Å crystal structure of the P4−P6 domain (*9*) and a secondary structure representation of the remainder of the RNA. P stands for a base-paired region, J for a joining region between two duplexes, and L for a terminal loop. The P5abc portion of the P4−P6 domain is colored in dark gray and the location of the break in the intact ribozyme to create separate P5abc and $E^{\Delta P5abc}$ RNAs is marked with an arrow. The two tertiary contact sites within the P4−P6 domain are shown in red and blue. The loops of L5c and L2, thought to form a pseudoknot, are highlighted in green.

loop and its receptor and between the A-rich bulge and P4 (Figures 1 and 2; ref *6*). A third tertiary contact, identified by phylogenetic covariation analysis and mutagenesis, occurs between the L5c loop on P5abc and the L2 loop near the 5′-end of $E^{\Delta P5abc}$, and probably involves Watson−Crick pairing of the self-complementary UGCA sequence in each loop (Figures 1 and 2; ref *17*).

We have conducted a structural and thermodynamic analysis of the assembly of P5abc and $E^{\Delta P5abc}$ to understand how the P5abc/$E^{\Delta P5abc}$ complex is stabilized and how the P5abc subdomain influences ribozyme structure. Chemical and enzymatic footprinting of the RNAs suggests that the complex assembles into the native structure from substantially prefolded P5abc and $E^{\Delta P5abc}$. These experiments further suggest that P5abc subtly influences the catalytic core structure rather than induces large rearrangements within $E^{\Delta P5abc}$. Strikingly, the affinity of the RNAs is exceedingly strong, with a K_d[1] of ~100 pM at 10 mM $MgCl_2$ determined by equilibrium gel mobility shift experiments at 37 °C. The strength of the interaction is likely due to prestructuring of both P5abc and $E^{\Delta P5abc}$ by magnesium ions and to extensive tertiary interactions along the RNA−RNA interface. This two-component system facilitates quantitative examination of specific tertiary interactions important in stabilizing the *Tetrahymena* group I intron RNA.

MATERIALS AND METHODS

RNA Preparation. Plasmids containing DNA sequences for P5abc and $E^{\Delta P5abc}$ were prepared by PCR amplification of sequences from pTZIVS. A T7 promoter sequence was included in the 5′ primer. The P5abc deletion was created using overlapping PCR primers. PCR products were ligated into pUC19 and transformed into JM109 cells. Plasmid DNA was isolated and cleaved with either *Bsa*I or *Sca*I prior to transcription to define the 3′-end of the RNAs. Transcription with T7 RNA polymerase was performed as described (*18*). To determine the concentration of the RNAs, each construct was hydrolyzed by incubation for several hours at room temperature in a solution of NaOH at pH 11 and the absorbance at 260 nm was measured. Concentration was calculated using extinction coefficients derived from those of the individual ribonucleotides at 260 nm and pH 11 (P-L Biochemicals, Circular No. OR-10).

Thermal Denaturation. Thermal denaturation was performed on a Cary UV spectrophotometer in a buffer containing 30 mM HEPES−KOH, pH 7.5 (25 °C), and 0.5 mM Na_2EDTA, pH 8.0, at several $MgCl_2$ concentrations. Immediately before performing thermal denaturation experiments, RNAs were prefolded by slow cooling from 80 to 5

[1] Abbreviations: K_d, equilibrium dissociation constant; K_a, equilibrium association constant; PCR, polymerase chain reaction; EDTA, ethylenediaminetetraacetic acid.

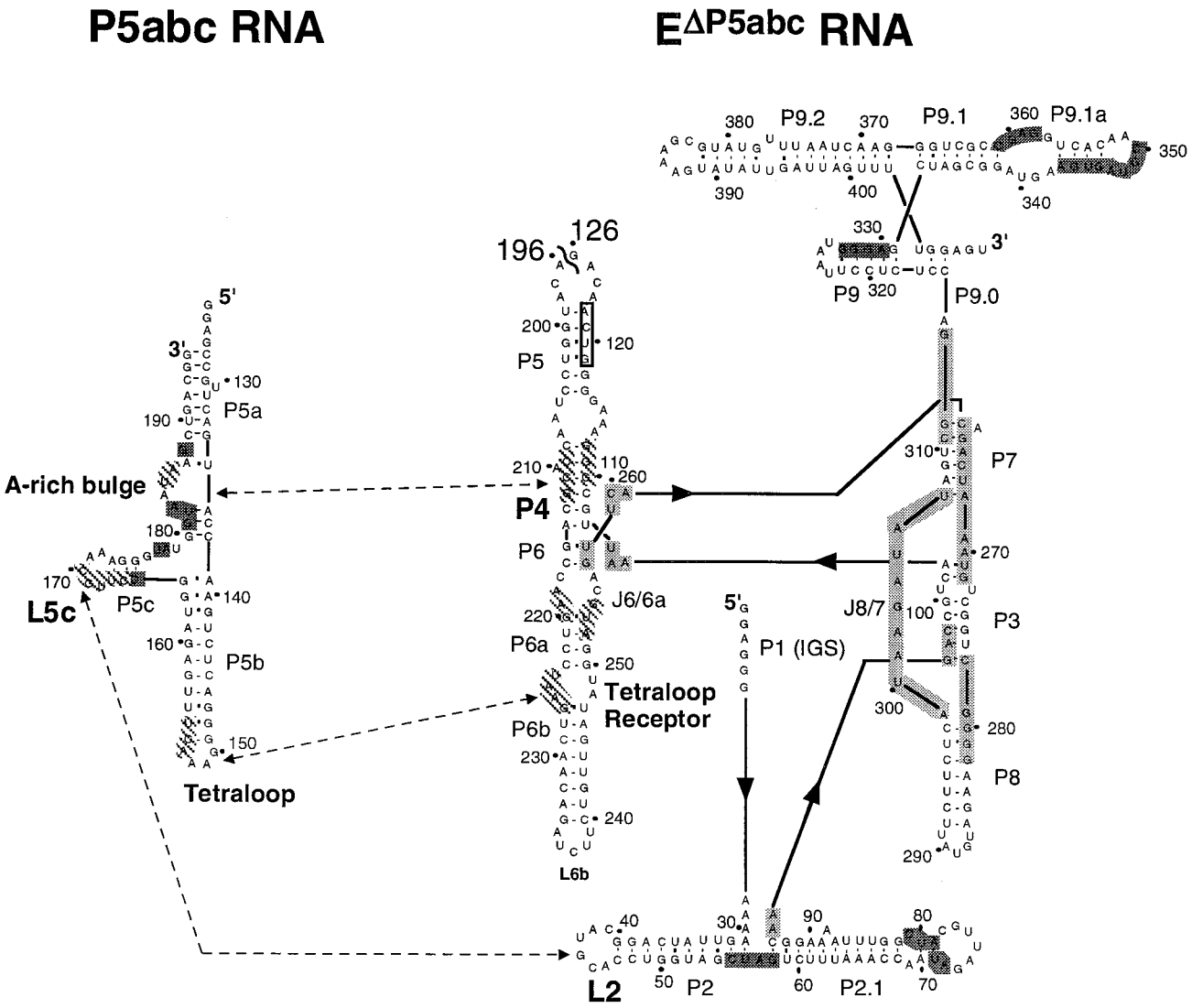

FIGURE 2: Secondary structures of P5abc and $E^{\Delta P5abc}$. Numbering is as for the intact *Tetrahymena* ribozyme. Dashed lines with double arrows show the location of tertiary contacts between P5abc and $E^{\Delta P5abc}$ (*9*, *17*). To construct $E^{\Delta P5abc}$, the sequence between positions 126 and 196 was deleted and G126 and A196 were covalently joined to create a loop above helix P5 in place of the bent internal bulge, J5/5a (*11*). Four extra nucleotides were added to the 5′-end of the P5abc RNA to ensure efficient transcription and end-labeling. Shaded and hatched regions show results of Fe(II)-EDTA protection. Positions 22−50, 127−145, 190−195, and 370−409 were not mapped. The extent of protection at each region was calculated by dividing the intensity of the region in the absence of $MgCl_2$ by the intensity in the presence of $MgCl_2$ after correcting for differences in lane loading and background cleavage in the starting material. Positions are considered to be protected only if the maximum extent of protection is 1.5 or higher in all experiments at 10−100 mM $MgCl_2$. Darkly shaded areas show approximately the same extent of protection from Fe(II)-EDTA cleavage in $E^{\Delta P5abc}$ or P5abc alone as in the P5abc/$E^{\Delta P5abc}$ complex. The extent of protection at positions 165, 177, 181−182, and 188 in P5abc ranges 2−10-fold and is not significantly changed upon addition of $E^{\Delta P5abc}$ while the extent of protection in the P9−P9.1−P9.1a region in $E^{\Delta P5abc}$ is about 1.5−2.5-fold in either the presence or the absence of P5abc. Lightly shaded areas show enhanced protection in the P5abc/$E^{\Delta P5abc}$ complex compared with the free RNAs. The extent of protection in P3, P7, J6/7, J8/7, P8, and J7/3 in $E^{\Delta P5abc}$ is 1.5−3-fold without P5abc but increases by an additional 2 to 40-fold in the presence of P5abc. Hatched areas are not protected in $E^{\Delta P5abc}$ or P5abc alone but become protected in the complex. The boxed area in P5 is protected within the intact *Tetrahymena* ribozyme but not protected in the P5abc/$E^{\Delta P5abc}$ complex.

°C in the same buffer in which thermal denaturation was performed. RNA solutions were heated from 5 to 95 °C with a heating rate of 0.5 °C/min. The spectra obtained were independent of the RNA concentration over a range of 0.1−1 μM, indicating the absence of contributions from multimeric species. The reversibility of folding at 0−1 mM $MgCl_2$ was confirmed by cooling RNAs from 80 to 5 °C at 0.5 °C/min in the UV apparatus.

Enzymatic and Chemical Footprinting. RNase T1 and V1 digests of 5′-end-labeled RNAs (*18*) were conducted in a buffer containing 30 mM Tris-HCl, pH 7.5 (25 °C), 10 mM NaCl, 10 mM $MgCl_2$, and 0.1 μg/μL $tRNA^{Phe}$ using 0.01−0.1 units T1 (Boehringer Mannheim) and 0.002−0.02 units V1 (Amersham Pharmacia Biotech). Fe(II)-EDTA footprinting was performed as described (*19*). Fe(II)-EDTA experiments were conducted on 5′-end-labeled RNA in buffers containing either 30 mM Tris-HCl, pH 7.5 (25 °C), and 10 mM NaCl or 20−50 mM HEPES−KOH, pH 6.8 (25 °C). The results were the same in each buffer. Attempts to perform Fe(II)-EDTA footprinting in the higher ionic strength gel mobility shift buffer below were unsuccessful as the high salt affected the resolution of the gels.

Gel Mobility Shift Experiments. ^{32}P-Labeled P5abc RNA (=10 pM) was incubated with increasing concentrations of $E^{\Delta P5abc}$ in 10 μL of 1 × THE buffer, (33 mM Tris, 66 mM HEPES, and 0.1 mM Na_2EDTA, pH 7.5, at 25 °C) or 1 ×

TME buffer, (33 mM Tris, 66 mM MES, and 0.1 mM Na_2-EDTA, pH 6.0 at 25 °C), 30 mM KCl, 4% glycerol, and 0.025% xylene cyanol in the presence of varying concentrations of $MgCl_2$. Prior to loading the gels, P5abc/$E^{\Delta P5abc}$ complexes were incubated for 10 min at 50 °C then for 1−72 h at 37 °C to allow equilibration (see Results). Gels were loaded with current running and were run for 1−5 h in 1 × THE or 1 × TME with 30 mM KCl and [$MgCl_2$] identical to that of the reaction buffer. To maintain temperature during gel loading, the gel boxes were attached to a water bath and water at 37 °C was circulated continuously around the gel.

Data Analysis. Gels were imaged using a Fuji phosphorimager, and band intensities were determined using the program MacBAS 2.0 (Fuji). UV spectra were analyzed with the program Igor Pro 3.0 (WaveMetrics). Fe(II)-EDTA experiments were analyzed as described (*20*). For gel mobility shift experiments, least-squares fits to a single-site binding equation (eq 1) were performed using Kaleidagraph 3.0 (Synergy Software).

$$f = (a - b)[[E^{\Delta P5abc}]/([E^{\Delta P5abc}] + K_d)] + b \qquad (1)$$

In eq 1, a is the maximum fraction of P5abc bound extrapolated to saturation while b is the amount of P5abc apparently bound in the absence of $E^{\Delta P5abc}$, accounting for small variations in background intensity (see below). K_d is the apparent equilibrium dissociation constant. After subtracting background intensity (typically 10−20% of the total intensity in each lane due to the extremely low concentration of [^{32}P]P5abc), the fraction of [^{32}P]P5abc shifted, f, was determined by dividing the intensity of the unshifted RNA by the total intensity of the lane. Typically, more than 95% of [^{32}P]P5abc was bound to $E^{\Delta P5abc}$ at saturation. Gel mobility shifts performed with concentrations of each RNA above the K_d to determine the stoichiometry of the complex were fit using eq 2,

$$f = \{[1/(2[P5abc])]([P5abc] + [E^{\Delta P5abc}] + K_d) - (a - b)\sqrt{([[P5abc] + [E^{\Delta P5abc}] + K_d)^2 - 4[P5abc][E^{\Delta P5abc}]]}\} + b$$

in which a is the maximum fraction of P5abc bound extrapolated to saturation and b is the fraction of P5abc apparently bound in the absence of $E^{\Delta P5abc}$.

RESULTS

Previous experiments showed that P5abc RNA supplied in trans greatly enhances the activity of a mutant *Tetrahymena* intron lacking P5abc (*11, 12*). The RNAs used here, P5abc and $E^{\Delta P5abc}$, are similar to those used in these previous studies. Their secondary structures are shown in Figure 2.

Substantially Prefolded P5abc and $E^{\Delta P5abc}$ RNAs Assemble to Form a Native Complex. To examine the assembly of the P5abc/$E^{\Delta P5abc}$ complex, we began by investigating the structure of the free and complexed RNAs as a function of magnesium concentration and temperature. Previous chemical footprinting studies of P5abc RNA constructs similar to the one used here showed that the P5abc subdomain folds independently into a stable native structure (*13, 14*). We used thermal denaturation monitored by UV absorbance spectroscopy as well as chemical and enzymatic footprinting to determine conditions that support folding of the P5abc RNA construct used here. First derivative plots of UV spectra show that at 2 or 10 mM $MgCl_2$, P5abc unfolds in a single sharp transition with a T_m of 75−80 °C, while at 0.1 or 0.5 mM $MgCl_2$, P5abc unfolds with a T_m of 40−45 °C (Figure 3a). Thus, 2 mM $MgCl_2$ is sufficient for P5abc to fold into a highly stable architecture.

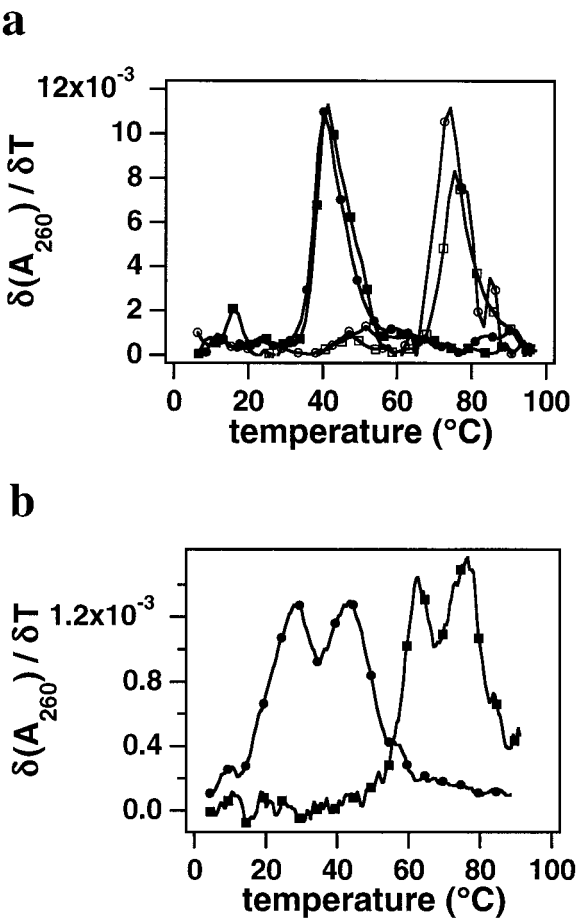

FIGURE 3: (a) Thermal denaturation of 1 μM P5abc monitored by UV absorbance spectroscopy. Derivative plots of raw data. Curves are marked with symbols for clarity. (●) 0.1 mM $MgCl_2$; (■) 0.5 mM $MgCl_2$; (○) 2 mM $MgCl_2$; (□) 10 mM $MgCl_2$. (b) Thermal denaturation of 100 nM $E^{\Delta P5abc}$. Derivative plot of raw data smoothed with a 5-point smoothing function. (●) No $MgCl_2$; (■) 5 mM $MgCl_2$.

We examined the secondary structure of P5abc by digestion with RNases T1 and V1 (at 10 mM $MgCl_2$, 42 °C), which cleave the RNA backbone in single-stranded and double-stranded regions, respectively. RNase T1 cuts P5abc strongly only in the tetraloop and L5c while RNase V1 cuts the RNA only within each duplex region, suggesting that P5abc is folded into a secondary structure consistent with that shown in Figure 2 (data not shown). Fe(II)-EDTA footprinting, which monitors the solvent accessibility of the ribose moieties in the backbone (*1*) shows that the tertiary structure of P5abc is essentially the same as that of other P5abc constructs (Figure 4a and Figure 2, shaded areas on P5abc; refs *13* and *14*). This analysis revealed that 1.5 mM $MgCl_2$ is sufficient to give half-maximal protection of P5abc, while complete protection is achieved at 4−6 mM $MgCl_2$ (Figure 4c). The higher $MgCl_2$ concentration requirement for complete Fe(II)-EDTA protection of P5abc at 37 °C compared with that for maximum thermal stability (4−6 mM

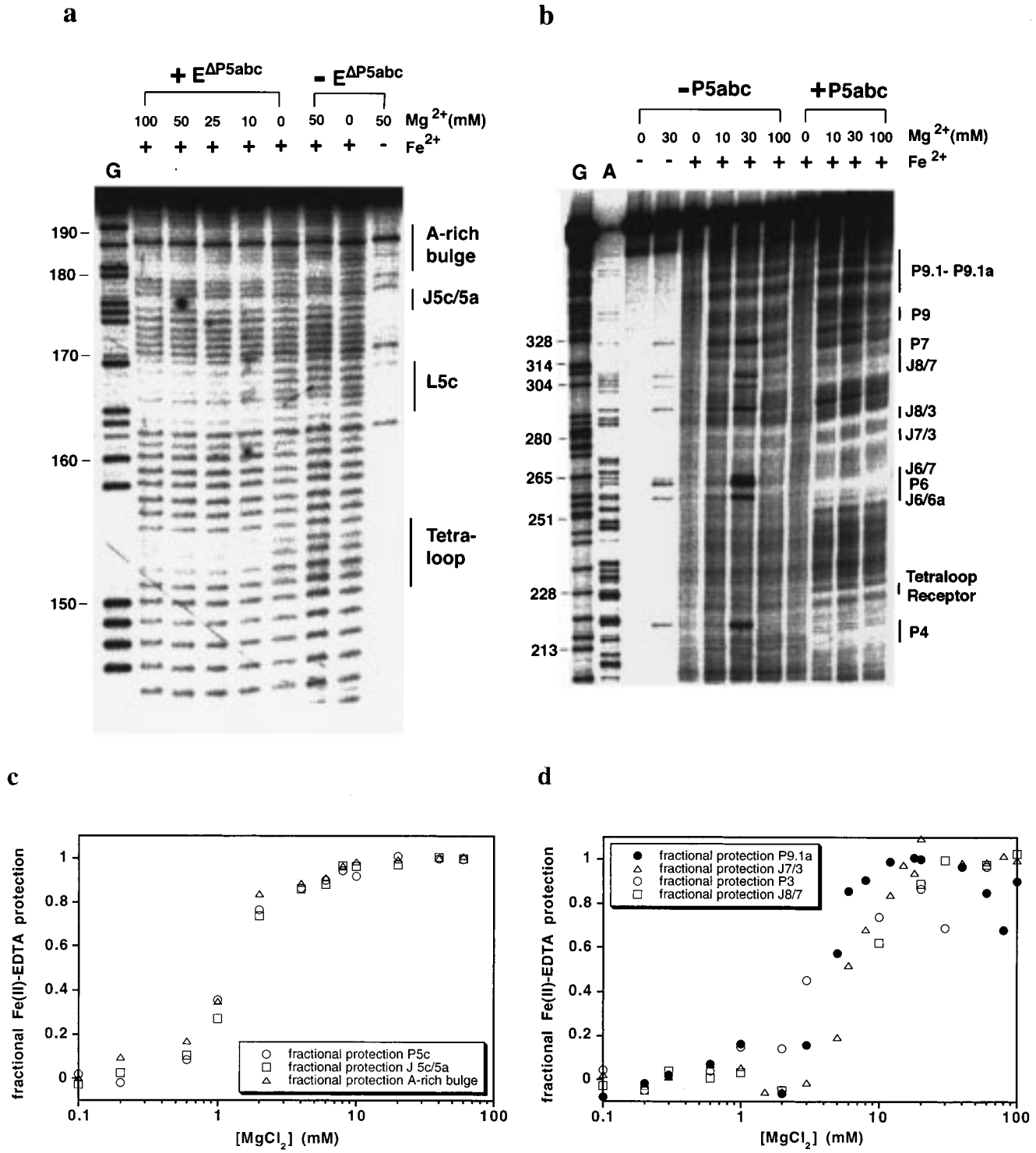

FIGURE 4: Fe(II)-EDTA footprinting of 5′-end-labeled P5abc and $E^{\Delta P5abc}$ RNAs. (a) Footprint of P5abc RNA at 42 °C in 30 mM Tris-HCl pH 7.5 (25 °C) buffer with [$MgCl_2$] indicated above the lanes. All reaction lanes contain 2 mM Na_2EDTA, pH 8.0, 1 mM sodium ascorbate, and 1 mM $Fe(NH_4)_2(SO_4)_2$. Indicated lanes also contain 1 μM $E^{\Delta P5abc}$. The far-right lane is a control for background cleavage in the RNA and was treated as for the others except that the $Fe(NH_4)_2(SO_4)_2$ cleavage reagent was omitted. Numbers on the left side of the gel refer to backbone sequence positions mapped with an RNase T1 ladder (G). Fe(II)-EDTA bands are shifted up one nucleotide from bands in the T1 ladder and the number labels. (b) Footprint of the $E^{\Delta P5abc}$ RNA at 37 °C in 20 mM HEPES−KOH, pH 6.8 (25 °C), and [$MgCl_2$] as indicated. Reactions also contain 2 mM Na_2EDTA pH 8.0, 1 mM sodium ascorbate, and 1 mM $Fe(NH_4)_2(SO_4)_2$ except the two far-left control reactions in which $Fe(NH_4)_2(SO_4)_2$ cleavage reagent was omitted. Indicated lanes contain 1 μM P5abc. Backbone sequence positions (numbers at left) are mapped based on RNase T1 (G) and RNase U2 (A) cleavage. (c) Fractional protection from Fe(II)-EDTA cleavage in P5abc at 37 °C in the same buffer as in panel a as a function of [$MgCl_2$]. (○) Protection at position 165 in P5c; (□) protection at position 177 between P5a and P5c; (△) protection at position 181 near the A-rich bulge. To calculate the fractional Fe(II)-EDTA protection, background intensity was subtracted and data were corrected for differences in lane loading. Then band intensity at a given [$MgCl_2$] was divided by band intensity at 0 mM $MgCl_2$. The resulting fractional intensities were normalized from 0 to 1 over the range of [$MgCl_2$]. (d) Fractional protection from Fe(II)-EDTA cleavage in $E^{\Delta P5abc}$ under the same conditions as in panel b as a function of [$MgCl_2$]. (●) Protection at positions 345−350 in P9.1a; (△) protection at positions 269−270 in J7/3; (○) protection at positions 95−96 in P3; (□) protection at positions 300−306 in J8/7.

vs 2 mM $MgCl_2$; Figures 3a and 4c) might reflect Mg^{2+} binding that increases the rigidity of folded P5abc. When a saturating concentration of $E^{\Delta P5abc}$ is added to P5abc, additional protection from Fe(II)-EDTA cleavage is observed in the A-rich bulge, the tetraloop, and L5c such that the protection pattern is nearly identical to that of the P5abc region within the intact ribozyme (Figure 4a and Figure 2, hatched areas on P5abc; refs *1* and *13*). These results strongly suggest that P5abc folds into a stable, native tertiary structure alone and forms the expected tertiary contacts in its complex with $E^{\Delta P5abc}$.

Previous experiments showed that a *Tetrahymena* intron RNA lacking P5abc does not self-splice efficiently at moderate concentrations of Mg^{2+} ions (*11*, *12*). This suggests two possibilities: $E^{\Delta P5abc}$ is folded in a stable conformation that has low activity, or the $E^{\Delta P5abc}$ structure, though largely correct, is unstable under these conditions.

To determine conditions for $E^{\Delta P5abc}$ folding, we first used thermal denaturation monitored by UV spectroscopy. A first derivative plot of a UV spectrum at 5 mM $MgCl_2$ shows two broad transitions, one beginning at 50−60 °C with a T_m of about 65 °C and one with a T_m of about 80 °C (Figure 3b). Without $MgCl_2$, $E^{\Delta P5abc}$ unfolds at a much lower temperature. A UV spectrum at 10 mM $MgCl_2$ is very similar to that at 5 mM $MgCl_2$ over the same temperature range (data not shown). Thus, $E^{\Delta P5abc}$ folds into a stable structure at 5 mM $MgCl_2$ below 50 °C. However, the unfolding transitions of $E^{\Delta P5abc}$ occur over a significantly broader temperature range than those of the intact ribozyme (*21*), indicating that $E^{\Delta P5abc}$ unfolding is less cooperative.

To probe the average structure of $E^{\Delta P5abc}$ in solution, we used RNase and Fe(II)-EDTA footprinting. RNase V1 cleaves $E^{\Delta P5abc}$ (at 10 mM $MgCl_2$, 37 °C) strongly in each of the duplex regions, including P3 and P7, while RNase T1 cuts $E^{\Delta P5abc}$ only in exposed loops and in the P6 duplex (not shown). This indicates that the $E^{\Delta P5abc}$ secondary structure forms stably under these conditions with the exception of the two base pair P6 stem. Fe(II)-EDTA protection at 5−100 mM $MgCl_2$ and 25 °C, 37 °C and 50 °C occurs at most of the same locations as in the intact ribozyme (Figure 4b and Figure 2, shaded areas on $E^{\Delta P5abc}$; refs *1* and *13*). Protection was not observed in the P4 stem and the tetraloop receptor, which directly contact P5abc, or in the J6/6a internal bulge and the P5 helix. The lack of protection in J6/6a may be related to the instability of the P6 stem, while the lack of protection in P5 may be due to the truncation at the top of this helix (see Figures 1 and 2). These results suggest that the ribozyme structure is only subtly perturbed by the deletion of P5abc. Fe(II)-EDTA footprinting over a range of magnesium concentrations shows that the P3 and P7 stems and the J8/7 strand within the $E^{\Delta P5abc}$ catalytic core require 20−30 mM $MgCl_2$ for maximal protection from cleavage while the peripheral, nonconserved P9.1a stem is fully protected at about 10 mM $MgCl_2$ (Figure 4d). In contrast, within the intact *Tetrahymena* ribozyme, all of these regions require only 2 mM $MgCl_2$ to achieve maximal Fe(II)-EDTA protection (*19*). Additionally, the extent of Fe(II)-EDTA protection throughout $E^{\Delta P5abc}$ is low, generally not more than 2-fold, suggesting that, on average, the $E^{\Delta P5abc}$ core helices are not tightly packed and are rather solvent accessible (Figure 2, legend).

Table 1: Association of P5abc and $E^{\Delta P5abc}$ at Several [$MgCl_2$][a]

[$MgCl_2$] (mM)	K_d (nM)[b]	ΔG (kcal/mol)[b]
5	9.8	11.3 ± 0.3
6	1.5	12.5 ± 0.2
7	0.66	13.0 ± 0.1
10	0.19	13.7 ± 0.3
	0.17[c]	13.8 ± 0.3
25	≤0.08[c]	≥14.4
50	≤0.07[c]	≥14.4

[a] Conditions: 1 × THE (pH 7.5). [b] K_d and ΔG reported here are average values from two to four independent experiments. Errors are standard deviations. [c] 1 × TME (pH 6.0) buffer, 30 mM KCl, 37 °C.

When a saturating concentration of P5abc is added to $E^{\Delta P5abc}$, new Fe(II)-EDTA protection is observed in P4, the tetraloop receptor, and J6/6a (Figure 4b and Figure 2, hatched areas on $E^{\Delta P5abc}$), resulting in a protection pattern nearly identical to that of the intact *Tetrahymena* ribozyme (*1*, *13*). Association with P5abc also results in considerably stronger Fe(II)-EDTA protection within the catalytic core than was seen in $E^{\Delta P5abc}$ alone (Figure 4b and Figure 2, lightly shaded areas). The degree of protection of the nonconserved peripheral portions of $E^{\Delta P5abc}$ does not change with addition of P5abc (Figure 4b and Figure 2, darkly shaded areas). This indicates that P5abc binding significantly reduces the solvent accessibility of the $E^{\Delta P5abc}$ backbone specifically within the catalytic core, perhaps by restricting the core's conformational flexibility (*22*). In summary, chemical and enzymatic footprinting experiments suggest that $E^{\Delta P5abc}$ is folded into a nativelike structure with a loosely packed core that is significantly tightened by P5abc binding.

P5abc and $E^{\Delta P5abc}$ Form an Exceptionally Stable Complex. We used gel mobility shift assays to determine the affinity of P5abc and $E^{\Delta P5abc}$ (Figure 5). Both RNAs migrate as single bands on native gels at concentrations up to 1 μM at 5−50 mM $MgCl_2$ (not shown). To examine the role of Mg^{2+} ions in promoting complex assembly, we determined the apparent equilibrium dissociation constant, K_d, at several concentrations of $MgCl_2$ between 2 and 50 mM at 37 °C (Table 1). At 2 and 3 mM $MgCl_2$, binding is not detected even at 10 μM $E^{\Delta P5abc}$. Binding is detectable at 4 mM $MgCl_2$ and the K_d decreases to less than 70 pM at 50 mM $MgCl_2$. A sample gel mobility shift experiment at 10 mM MgCl2 is shown in Figure 5.

At each $MgCl_2$ concentration, gel shifts were run after incubation of the complex for 1−72 h in order to determine the time required to reach equilibrium (Figure 5c). The required time increased with increasing $MgCl_2$, with less than 10 h required at 5 mM $MgCl_2$ and about 24 h required at 7−10 mM $MgCl_2$. Incubation of RNAs over long time periods with divalent cations can lead to significant RNA degradation. Thus, the fraction of free $E^{\Delta P5abc}$ that remained folded during the long incubations required to achieve equilibrium was estimated by native gel analysis (see Materials and Methods). This analysis showed that about 80% of $E^{\Delta P5abc}$ migrated as folded RNA after 10−15 h of incubation of the complex at 5−10 mM $MgCl_2$ at pH 7.5 (not shown). As longer incubation times were required at higher $MgCl_2$ concentrations, gel mobility shifts at 10−50 mM $MgCl_2$ were run at pH 6.0, which left more than 90% of $E^{\Delta P5abc}$ migrating as folded RNA after 24−48 h (not shown).

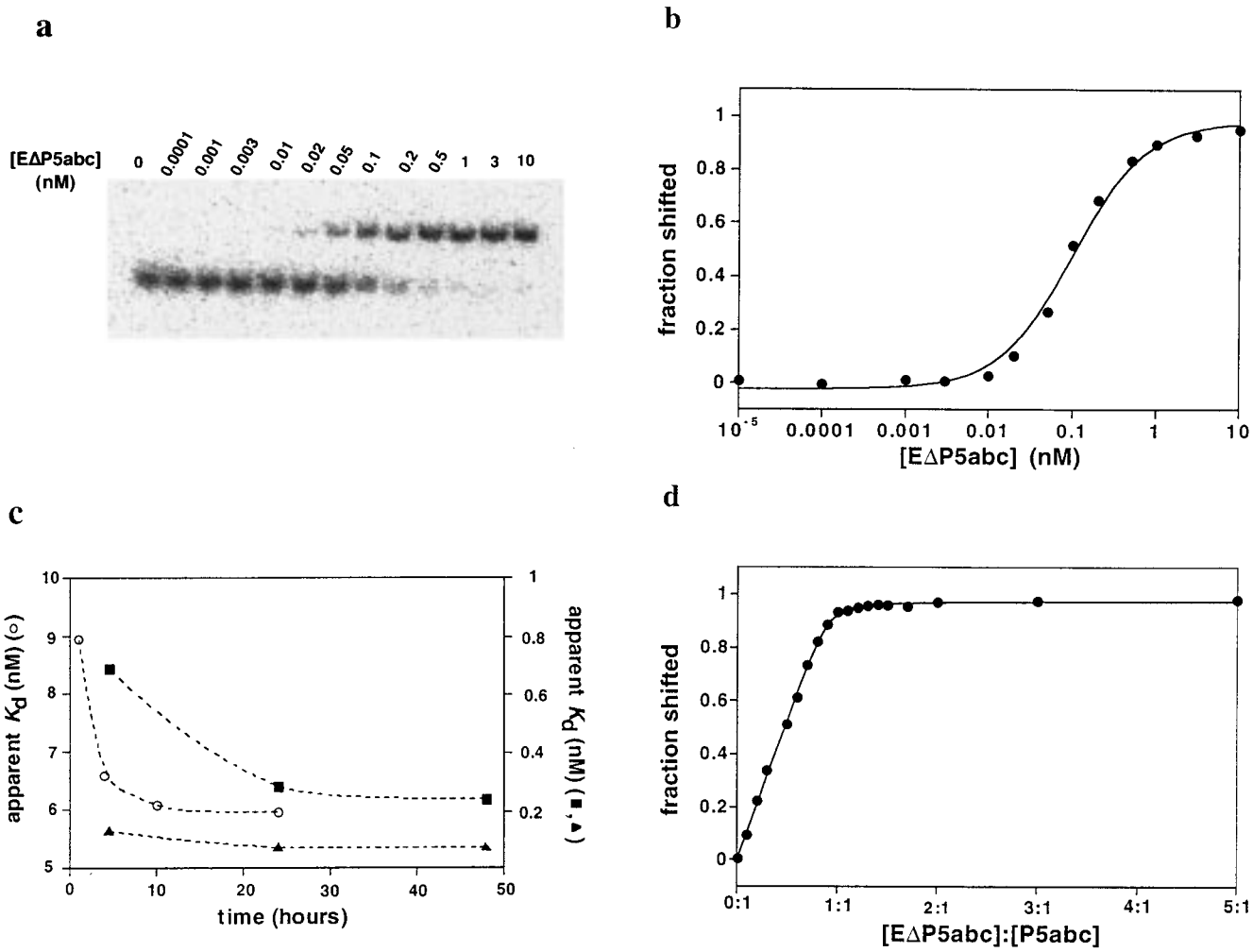

FIGURE 5: (a) Sample gel mobility shift of ^{32}P-labeled P5abc (=10 pM) with $E^{\Delta P5abc}$ in 1 × TME buffer with 30 mM KCl and 10 mM $MgCl_2$ after 24 h of incubation at 37 °C. (b) Fit of data from panel a using eq 1 (see Materials and Methods). The apparent K_d for the experiment shown is 99.5 pM. (c) Plot of apparent K_d against incubation time at 37 °C. (○) 5 mM $MgCl_2$; (■) 10 mM $MgCl_2$; (▲) 25 mM $MgCl_2$. (d) Test of the stoichiometry of the P5abc/$E^{\Delta P5abc}$ complex. Radio-labeled P5abc (75 nM) was incubated with up to 1 μM $E^{\Delta P5abc}$ at 37 °C and the same buffer conditions as in panel a. The data were fit using eq 2 (Materials and Methods). The break point in the curve occurs at 78 nM $E^{\Delta P5abc}$, suggesting that the P5abc and $E^{\Delta P5abc}$ form a 1:1 complex in the gel. Several repeats of this experiment with different preparations of each RNA and [P5abc] from 70 to 200 nM gave P5abc:$E^{\Delta P5abc}$ ratios of 1.0 ± 0.2.

The following observations suggest that apparent equilibrium dissociation constants at 5−10 mM $MgCl_2$ closely approximate the actual equilibrium constants. First, incubation of the complex under these conditions was sufficiently long to ensure that the apparent K_d value reached a limit indicating equilibrium had been reached (Figure 5c). Second, free $E^{\Delta P5abc}$ RNA remained largely folded after several hours of incubation at each experimental condition, thus minimizing artificial inflation of K_d values due to RNA degradation during equilibration of the complex. Third, the K_d values at 5 and 10 mM $MgCl_2$ are within 3-fold of those calculated from association and dissociation rate constants for assembly of the complex (E. A. Doherty, R. Russell, M. Engelhardt, J. A. Doudna, and D. Herschlag, unpublished results). Even though the apparent K_ds at 25 and 50 mM $MgCl_2$ appear to level off after about 24 h of incubation of the complex (Figure 5c and data not shown), these data represent upper limits for the actual equilibrium constants (Table 1). This is due to potential limitations from cleavages within the folded $E^{\Delta P5abc}$ RNA at the long times required for equilibration under these conditions and due to an absence of confirming kinetic data.

The data are fit reasonably well by a single-site binding equation, though the slope of the transition region is always somewhat steeper than that of the fit (Figure 5b). While this could in principle arise from multiple $E^{\Delta P5abc}$ molecules that bind to P5abc, gel shifts conducted with concentrations of each RNA well above the K_d show that the RNAs form a 1:1 complex in the gel (Figure 5d). Additionally, the mobility of the P5abc/$E^{\Delta P5abc}$ complex does not change with concentration of $E^{\Delta P5abc}$, providing no indication of higher order complexes.

Strikingly, the affinity of P5abc and $E^{\Delta P5abc}$ is more than 10-fold stronger than that for any tertiary interaction previously reported between large RNA molecules (see Discussion; refs *23*−*27*). Binding affinity of P5abc and $E^{\Delta P5abc}$ is exceptionally strong at high $MgCl_2$ concentrations at which the $E^{\Delta P5abc}$ periphery and catalytic core are completely protected from cleavage by Fe(II)-EDTA (Figure 6). This observation strongly suggests that tightly bound complexes form more readily under conditions in which P5abc and $E^{\Delta P5abc}$ molecules fold independently into substantially nativelike structures.

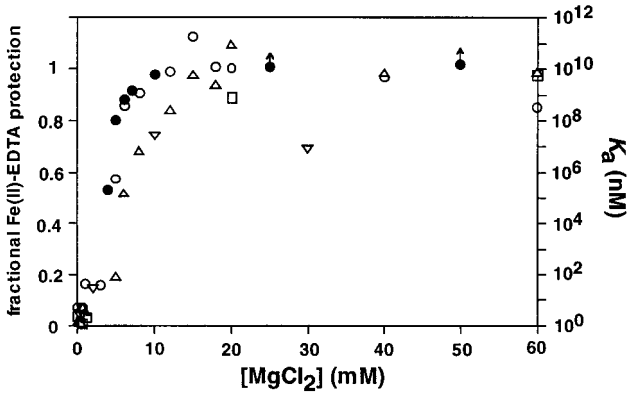

FIGURE 6: Comparison of K_a for binding of P5abc and $E^{\Delta P5abc}$ and fractional protection from Fe(II)-EDTA cleavage in $E^{\Delta P5abc}$ (from Figure 4d). (●) K_a of the P5abc/$E^{\Delta P5abc}$ complex from gel mobility shifts conducted at 4−50 mM $MgCl_2$ (see Table 1). (○) Protection at positions 345−350 in P9.1a; (△) fractional Fe(II)-EDTA protection at positions 269−270 in J7/3; (▽) fractional Fe(II)-EDTA protection at positions 95−96 in P3; (□) protection at positions 300−306 in J8/7.

DISCUSSION

We have performed a structural and thermodynamic analysis of the formation of a two-part *Tetrahymena* ribozyme (Figures 1 and 2). Compared to the intact ribozyme, this is a convenient system for quantitative examination of tertiary interactions in this large RNA. $E^{\Delta P5abc}$, containing the elements of the ribozyme catalytic core, forms an exceptionally tight complex with P5abc, reaching a K_d of less than 100 pM between 10 and 50 mM $MgCl_2$ at 37 °C (Table 1). We conclude that the stability of this complex is due to specific interactions between functional groups occurring over a large binding interface and to substantial prefolding of both RNAs by magnesium ions, as discussed below.

The P5abc/$E^{\Delta P5abc}$ Complex Assembles from Substantially Prefolded RNA Components. Chemical and enzymatic footprinting as well as thermal denaturation indicate that both the P5abc and $E^{\Delta P5abc}$ RNAs adopt substantially native folds under a wide range of temperatures and magnesium concentrations (Figures 2−4). However, the lack of strong protection from cleavage by Fe(II)-EDTA indicates that the $E^{\Delta P5abc}$ catalytic core, though folded into a nativelike structure, exhibits significant solvent penetration.

Chemical footprinting experiments provide a structural picture of P5abc/$E^{\Delta P5abc}$ association, showing that complex assembly is accompanied by both formation of the correct tertiary interactions and significant tightening of the ribozyme catalytic core (Figures 2 and 4). P5abc and $E^{\Delta P5abc}$ show high affinity only once enough $MgCl_2$ is present to completely fold both components, implying that nativelike RNAs form the most stable complexes (Figure 6). Initial kinetic experiments suggest that the bimolecular association rate of the complex is also increased over the range of $MgCl_2$ concentration required to fold the individual components, again consistent with more efficient association of preformed structural units (E. Doherty, R. Russell, M. Engelhardt, J. Doudna, and D. Herschlag, unpublished results).

The conserved group I intron catalytic core is composed of two large domains of tertiary structure, P4−P6 and P3−P9 (*10*). Previous chemical footprinting experiments showed that while P4−P6 is independently folded (*13*), the P3−P9 region, comprising stems P2, P2.1, P3, P7, P8, and P9−P9.2 and the J8/7 strand, contains its correct secondary structure but only a partially formed tertiary structure on its own (*20*). Fe(II)-EDTA protection in the P3−P9 region is only observed in the P2−P2.1 and P9.1−P9.2 nonconserved, peripheral extensions and at J7/3 and P8 in the catalytic core (*20*). Results here show that addition of the P5, P4, and P6−P6b helices to P3−P9, forming $E^{\Delta P5abc}$, is sufficient for formation of an essentially native but loosely compacted catalytic core.

Affinity of P5abc and $E^{\Delta P5abc}$ Is Exceptionally Strong. To our knowledge, the interaction between P5abc and $E^{\Delta P5abc}$ is the strongest tertiary interaction between large RNAs yet observed. Interactions between other domains of the *Tetrahymena* ribozyme range from 4 to 30 nM (80 mM $MgCl_2$ and 5 mM spermidine, 40 °C; ref *23*), while the K_d for the interaction of domain 5 with remaining portions of the ai5γ group II intron is 200−900 nM (100 mM $MgCl_2$ and 0.5−1M KCl, 42−45 °C; refs *24* and *25*). Additionally, RNase P RNA binds to a pre-$tRNA^{Asp}$ substrate with a K_d of 100 nM and to $tRNA^{Asp}$ with a K_d of 3 nM (100 mM magnesium acetate and 0.8 M NH_4Cl, 37 °C; ref *26*). If the association of P5abc and $E^{\Delta P5abc}$ occurred through duplex formation rather than tertiary contacts, an average of 8−9 base pairs would be required to obtain an equivalent ΔG (∼ 14 kcal/mol; Table 1).[2] P5abc and $E^{\Delta P5abc}$ make three tertiary contacts to form a large bimolecular interface (Figures 1 and 2). The energetic contribution provided by each of these tertiary contact sites is currently unknown. However, the L5c−L2 pseudoknot (*17*) would be expected to contribute about 3−4 kcal/mol toward the binding energy, assuming that the initiation and stacking energies of this pseudoknot are similar to those of a four base pair duplex (*28*).

In addition to specific tertiary interactions between P5abc and $E^{\Delta P5abc}$, cooperativity due to prealignment of the three contact sites on the RNAs may significantly increase the binding free energy. Coaxial stacking of P5a on P5b within the P5abc subdomain may prealign the A-rich bulge and tetraloop located at each end of the stack, while the junction of P5c with this stack may preorient L5c for contact with L2 (Figures 1 and 2). Small folding units such as P5abc may generally serve as platforms for assembly of larger units of tertiary structure (*20*, *29*). It will be interesting to examine the flexibility of P5abc and energetic effects of changes in the alignment of these tertiary contact sites.

Fe(II)-EDTA results and previous crystallographic data (*9*) suggest that formation of the P5abc/$E^{\Delta P5abc}$ complex results in significant release of solvent from the RNA−RNA interface, providing a favorable entropic contribution to the binding. The amount of solvent-inaccessible surface created by helical packing in the P4−P6 domain is roughly 2300 Å^2 (*9*). In addition, the L5c−L2 tertiary contact would be expected to bury an additional 500−1000 Å^2 of molecular surface, as estimated from the surface buried by a crystal contact between L5c and J6/6a in adjacent P4−P6 molecules in the crystals (*9*). Furthermore, the degree of Fe(II)-EDTA

[2] This calculation was performed using nearest neighbor calculations of duplex stability, where the ΔG for initiation of a duplex is 3.4 kcal/mol and the average ΔG for each additional base pair is −1.9 kcal/mol (*28*).

cleavage throughout the $E^{\Delta P5abc}$ core decreases by at least 2-fold upon P5abc binding (Figure 2, lightly shaded areas and legend), implying that assembly of the complex also results in a significant release of solvent from the $E^{\Delta P5abc}$ core. The release of solvent molecules from all of these regions into the bulk solvent should strongly favor assembly of the P5abc/$E^{\Delta P5abc}$ complex and may provide a significant contribution to the binding free energy of the RNAs.

Comparison to Interactions between Group I Introns and the Cyt 18 Protein. The *Neurospora crassa* mt LSU and *ND1* introns do not contain a peripheral segment equivalent to P5abc and are not detectably self-splicing in vitro. In these introns, the Cyt 18 protein binds tightly to the intron RNAs along the same contact interface as P5abc to promote tertiary folding and self-splicing activity (*30*). Cyt 18 also interacts strongly with a *Tetrahymena* intron mutant lacking P5abc at low Mg^{2+} concentrations where the intron is poorly folded and largely inactive (*16*). In contrast to the $E^{\Delta P5abc}$ RNA, the catalytic core regions of the *N. crassa* mt LSU and *ND1* introns show little evidence of tertiary folding without Cyt 18 even at fairly high $MgCl_2$ concentrations (25 mM $MgCl_2$, 10 mM HEPES−KOH, 100 mM KCl; ref *30*). The catalytic core of $E^{\Delta P5abc}$ folds into a nativelike structure under similar conditions. Thus, $E^{\Delta P5abc}$ may contain structural elements not found within the mt LSU and *ND1* introns of *N. crassa* that allow for folding of the catalytic core in the absence of P5abc or Cyt 18.

In summary, the P5abc/$E^{\Delta P5abc}$ system provides an excellent framework for quantitative examination of specific interactions important for the formation of tertiary contacts in this RNA, since in this two-component system the binding free energy of P5abc and $E^{\Delta P5abc}$ mutants can be easily compared to that of the wild-type RNAs.

ACKNOWLEDGMENT

We thank Kaifeng Zhou for technical assistance, members of the Crothers lab for help with UV spectroscopy, Robert Batey, Robert Rambo, and Alexander Szewczak for help with figures and calculations, and Rick Russell, Mark Engelhardt, Robert Batey, Jeff Kieft, and members of the Doudna and Herschlag labs for helpful discussions.

REFERENCES

1. Latham, J. A., and Cech, T. R. (1989) *Science 245*, 276−82.
2. Laing, L. G., Gluick, T. C., and Draper, D. E. (1994) *J. Mol. Biol. 237*, 577−87.
3. Heus, H. A., and Pardi, A. (1991) *Science 253*, 191−4.
4. Murphy, F. L., and Cech, T. R. (1994) *J. Mol. Biol. 236*, 49−63.
5. Costa, M., and Michel, F. (1995) *EMBO J. 14*, 1276−85.
6. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science 273*, 1696−9.
7. Correll, C. C., Freeborn, B., Moore, P. B., and Steitz, T. A. (1997) *Cell 91*, 705−12.
8. Strobel, S. A., Ortoleva-Donnelly, L., Ryder, S. P., Cate, J. H., and Moncoeur, E. (1998) *Nat. Struct. Biol. 5*, 60−5.
9. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science 273*, 1678−85.
10. Michel, F., and Westhof, E. (1990) *J. Mol. Biol. 216*, 585−610.
11. Joyce, G. F., van der Horst, G., and Inoue, T. (1989) *Nucleic Acids Res. 17*, 7879−89.
12. van der Horst, G., Christian, A., and Inoue, T. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 184−8.
13. Murphy, F. L., and Cech, T. R. (1993) *Biochemistry 32*, 5291−300.
14. Cate, J. H., Hanna, R. L., and Doudna, J. A. (1997) *Nat. Struct. Biol. 4*, 553−8.
15. Sclavi, B., Sullivan, M., Chance, M. R., Brenowitz, M., and Woodson, S. A. (1998) *Science 279*, 1940−3.
16. Mohr, G., Caprara, M. G., Guo, Q., and Lambowitz, A. M. (1994) *Nature 370*, 147−50.
17. Lehnert, V., Jaeger, L., Michel, F., and Westhof, E. (1996) *Chem. Biol. 3*, 993−1009.
18. Latham, J. A., Zaug, A. J., and Cech, T. R. (1990) *Methods Enzymol. 181*, 558−69.
19. Celander, D. W., and Cech, T. R. (1991) *Science 251*, 401−7.
20. Doherty, E. A., and Doudna, J. A. (1997) *Biochemistry 36*, 3159−69.
21. Banerjee, A. R., Jaeger, J. A., and Turner, D. H. (1993) *Biochemistry 32*, 153−63.
22. Cohen, S. B., and Cech, T. R. (1997) *J. Am. Chem. Soc. 119*, 6259−68.
23. Doudna, J. A., and Cech, T. R. (1995) *RNA 1*, 36−45.
24. Pyle, A. M., and Green, J. B. (1994) *Biochemistry 33*, 2716−25.
25. Michels, W. J., Jr., and Pyle, A. M. (1995) *Biochemistry 34*, 2965−77.
26. Beebe, J. A., and Fierke, C. A. (1994) *Biochemistry 33*, 10294−304.
27. Beebe, J. A., Kurz, J. C., and Fierke, C. A. (1996) *Biochemistry 35*, 10493−505.
28. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., and Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373−7.
29. Zarrinkar, P. P., and Williamson, J. R. (1994) *Science 265*, 918−24.
30. Caprara, M. G., Mohr, G., and Lambowitz, A. M. (1996) *J. Mol. Biol. 257*, 512−31.

BI982113P