The P5abc Peripheral Element Facilitates Preorganization of the Tetrahymena Group I Ribozyme for Catalysis †

ARTICLE *in* BIOCHEMISTRY · APRIL 2000

Impact Factor: 3.02 · DOI: 10.1021/bi992313g · Source: PubMed

5 AUTHORS, INCLUDING:

Daniel Herschlag
Stanford University
248 PUBLICATIONS **16,673** CITATIONS

SEE PROFILE

The P5abc Peripheral Element Facilitates Preorganization of the *Tetrahymena* Group I Ribozyme for Catalysis†

Mark A. Engelhardt,‡ Elizabeth A. Doherty,§ Deborah S. Knitt,‡ Jennifer A. Doudna,*,§,|| and Daniel Herschlag*,‡

Department of Biochemistry, Stanford University, Stanford, California 94305, and Department of Molecular Biophysics and Biochemistry, and Howard Hughes Medical Institute, Yale University, New Haven, Connecticut 06520

Received October 5, 1999; Revised Manuscript Received December 13, 1999

ABSTRACT: Phylogenetic comparisons and site-directed mutagenesis indicate that group I introns are composed of a catalytic core that is universally conserved and peripheral elements that are conserved only within intron subclasses. Despite this low overall conservation, peripheral elements are essential for efficient splicing of their parent introns. We have undertaken an in-depth structure−function analysis to investigate the role of one of these elements, P5abc, using the well-characterized ribozyme derived from the Tetrahymena group I intron. Structural comparisons using solution-based free radical cleavage revealed that a ribozyme lacking P5abc ($E^{\Delta P5abc}$) and $E^{\Delta P5abc}$ with P5abc added in trans ($E^{\Delta P5abc}$•P5abc) adopt a similar global tertiary structure at Mg^{2+} concentrations greater than 20 mM [Doherty, E. A., et al. (1999) Biochemistry *38*, 2982-90]. However, free $E^{\Delta P5abc}$ is greatly compromised in overall oligonucleotide cleavage activity, even at Mg^{2+} concentrations as high as 100 mM. Further characterization of $E^{\Delta P5abc}$ via DMS modification revealed local structural differences at several positions in the conserved core that cluster around the substrate binding sites. Kinetic and thermodynamic dissection of individual reaction steps identified defects in binding of both substrates to $E^{\Delta P5abc}$, with ≥25-fold weaker binding of a guanosine nucleophile and ≥350-fold weaker docking of the oligonucleotide substrate into its tertiary interactions with the ribozyme core. These defects in binding of the substrates account for essentially all of the 10^4-fold decrease in overall activity of the deletion mutant. Together, the structural and functional observations suggest that the P5abc peripheral element not only provides stability but also positions active site residues through indirect interactions, thereby preferentially stabilizing the active ribozyme structure relative to alternative less active states. This is consistent with the view that peripheral elements engage in a network of mutually reinforcing interactions that together ensure cooperative folding of the ribozyme to its active structure.

Catalytic RNAs can provide rate enhancements that rival those of protein enzymes (*1*−*3*). To accomplish this, an RNA must contain sufficient information in its primary sequence to meet two challenges: it must be able to form stable tertiary structure, and it must be able to stabilize the active structure relative to alternate inactive and less active conformations. Relative to proteins, RNA is expected to encounter structural difficulties arising from the charged nature and greater rotational freedom of the phosphodiester backbone, the limited number of nucleoside bases, and the sequestration of the bases by base-pairing within secondary structure (*4*, *5*). We have focused on group I introns to study the means by which RNA molecules can achieve functional structures.

The wealth of functional and structural information available for group I introns makes them a powerful system in which to examine the structural underpinnings of RNA catalytic function (for reviews see refs *5*-*7*). Phylogenetic comparison, bolstered by mutagenesis and chemical probing, revealed that the secondary structure and some specific residues of the P1, P5−P4−P6, and P3−P7 paired regions are highly conserved among all group I introns, hundreds of examples of which are currently known (Figure 1A, refs *8*-*12*). Structural models from phylogenetic comparisons, X-ray diffraction data, and chemical modification, as well as functional data from site-directed mutagenesis, suggest that this conserved core contains the groups that make direct interactions with substrates (*11*, *13*−*18*; see also ref *19* for functionally important metal sites).

Beyond the universally conserved core, group I introns have been divided into five subclasses based on their complement of less conserved sequences called peripheral elements (Figure 1B, refs *11*, *20*, and *21*). These elements appear to wrap around the conserved core to form the outer shell of a globular fold, away from the site of bond-making and bond-breaking (*15*, *17*, *22*, *23*). The conservation of peripheral elements within subclasses and the demonstration

† This work was supported by NIH grant GM 49243 to D.H., and grants from the Lucille P. Markey Charitable Trust and the Donaghue Medical Research Foundation, and NIH Grant GM 56361-01 to J.A.D. M.A.E. was supported in part by a Stanford Graduate Fellowship. D.S.K. and E.A.D. were supported by NIH training grants.

* Correspondence should be addressed to J.A.D. or D.H. J.A.D. Phone: (203) 432-3108. Fax: (203) 432-3104. E-mail: doudna@csb.yale.edu. D.H. Phone: (650) 723-9442. Fax: (650) 723-6783. E-mail: herschla@cmgm.stanford.edu.

‡ Stanford University.

§ Yale University.

|| Howard Hughes Medical Institute, Yale University, New Haven, Connecticut 06520.

10.1021/bi992313g CCC: $19.00
Published on Web 02/17/2000

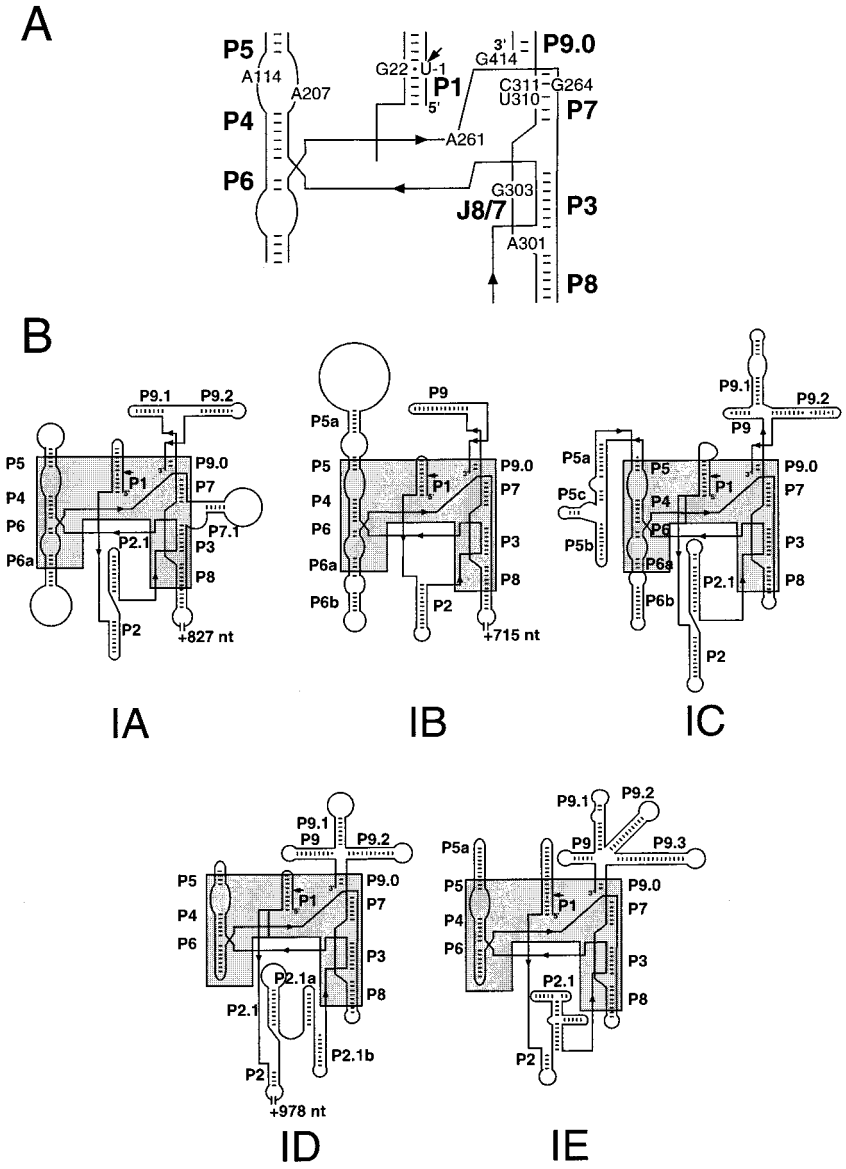

FIGURE 1: Peripheral elements decorate the conserved core of group I introns. (A) The conserved secondary structure of the core of group I introns, including paired regions P1, P3, P4, P5, P6, P7, and P9.0. Annotated positions are conserved in at least 84 of the 87 introns analyzed by Michel and Westhof (*11*). The secondary structure is labeled according to Burke et al., and numbering is for the *Tetrahymena* intron (*10*). An arrow marks the 5′ splice site. (B) Peripheral elements are shown extending from the conserved core shown in part A (in the shaded box, structures from refs *12* and *21*). Group I introns are divided into five subclasses based on which peripheral elements are attached to the ribozyme core (*11, 20, 21*). The representatives depicted are S. cerevisiae LSU rRNA (group IA), *Saccharomyces cerevisiae* OX1 [i4] (group IB), *Tetrahymena thermophila* LSU rRNA (group IC), *Podospora anserina* COB [I2] (group ID), and *Cryptendoxyla hypophloia* SSU rRNA (group IE).

that deletion of these elements results in greatly compromised self-splicing activity emphasize their importance to group I intron function (*20*, *24*-*30*). The low overall conservation of peripheral element sequence and identity, their distance from the active site, and the increased Mg^{2+} requirement for ribozyme folding and activity upon their deletion suggest these elements act indirectly on the active site by stabilizing and/or modulating the intron fold (*24*, *25*, *27*, *30*-*32*).

To examine the structural and functional contribution of a peripheral element to formation of the active group I intron structure, we have studied the P5abc element of the Tetrahymena group I intron (Figures 1B and 2). Chemical footprinting studies have shown that P5abc folds at lower Mg^{2+} concentrations than other intron elements and can fold independently to a stable, native-like structure (*33*-*35*). The 2.8 Å crystal structure of the P4−P6 domain of this intron revealed the detailed architecture of P5abc, which consists of a complex three-helix junction with a core of magnesium ions (*16*, *36*). Data from X-ray diffraction, phylogenetic covariation analysis, and site-directed mutagenesis show that P5abc makes direct contacts with P4, J6a/6b, and L2, forming an extensive interface with the P5−P4−P6 helical stack of the conserved core and further interactions with the P2 peripheral element (Figure 2, refs *15*-*17*, *34*, and *37*).

[1] Abbreviations: $E^{\Delta P5abc}$, the mutant L-21 *Sca*I *Tetrahymena* ribozyme with P5abc deleted; E^{wt}, the wild-type L-21 *Sca*I *Tetrahymena* ribozyme; E, $E^{\Delta P5abc}$, $E^{\Delta P5abc}$·P5abc, or E^{wt}; *S, 5′-^{32}P-labeled oligonucleotide substrate; IGS, the internal guide sequence of the ribozyme (see Figure 2); GMP, guanosine 5′-phosphate; DMS, dimethyl sulfate; MOPS, 3-(*N*-morpholino)propanesulfonic acid; MES, 2-(*N*-morpholino)ethanesulfonic acid; EDTA, (ethylenedinitrilo)tetraacetic acid; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid.

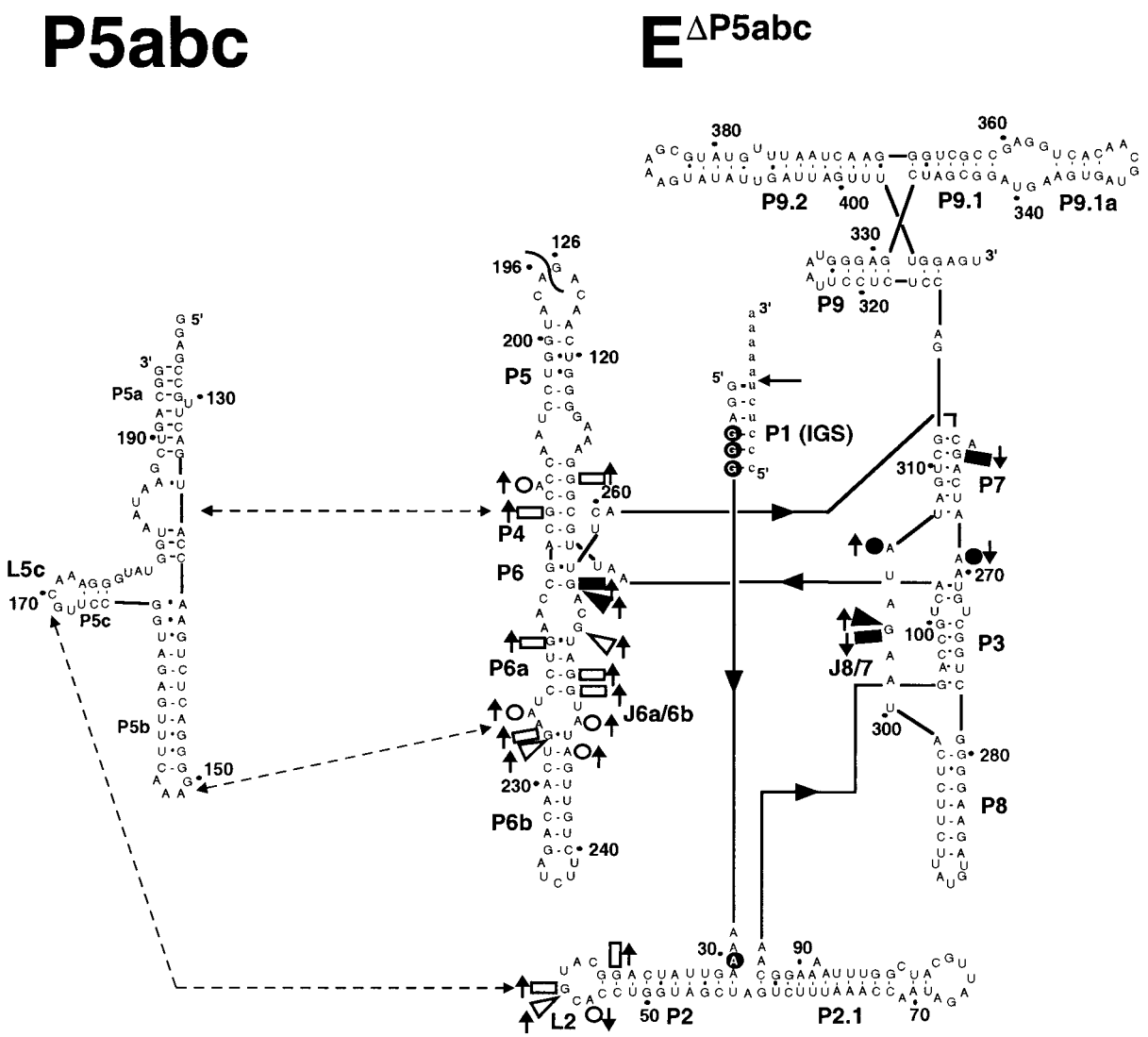

FIGURE 2: Secondary structure of the *T. thermophila* L-21 ScaI $E^{\Delta P5abc}$ ribozyme and P5abc peripheral element. The oligonucleotide substrate is shown (lowercase letters) forming interactions with the IGS of the ribozyme, and the primary site of substrate oligonucleotide cleavage is marked by an arrow. Annotated residues have ≥2-fold changes in susceptibility to DMS or kethoxal modification in $E^{\Delta P5abc}$ compared to $E^{\Delta P5abc}$•P5abc, with upward-pointing arrows indicating increased susceptibility to modification with P5abc removed and downward arrows the opposite. Of these positions, the open symbols represent residues that make direct interactions with P5abc in the crystal structure of the P4−P6 domain, or, in the case of L2, that are suggested by phylogenetic and mutational data to be making direct interactions (*15, 16*). The closed symbols represent the residues not expected to make any direct interaction with P5abc. Wedges, rectangles, and circles denote changes detected by kethoxal, DMS (aniline cleavage), and DMS (reverse transcription), respectively. Positions with white text and underlying closed circles represent residues not expected to make direct interactions with P5abc that have modification patterns that are modulated by the binding of substrates. Positions 25−370 were assayed for changes in modification, and a subset of these data is presented in Figures 6 and 10.

We have used the L-21 ScaI ribozyme lacking P5abc ($E^{\Delta P5abc}$)[1] to examine the role of the P5abc peripheral element in the structure and function of the *Tetrahymena* group I intron. Although $E^{\Delta P5abc}$ has a globally native-like fold at high Mg^{2+} concentrations (*32*), structure probing and measurement of individual reaction steps carried out herein indicate that it is structurally distinct from and functionally compromised relative to the intact ribozyme. Our results suggest that the P5abc peripheral element not only stabilizes ribozyme structure but also contributes to the positioning of groups important for catalysis as a component of the interconnected ring of elements around the catalytic core.

EXPERIMENTAL PROCEDURES

Materials. All reagents and chemicals were of the highest purity available. Ribozymes for functional studies were prepared by in vitro transcription and purified from plasmids pT7L-21 and pT7L-21ΔP5abc as described previously (*38*); the plasmid pT7L-21ΔP5abc was a generous gift from Marc Caprara and Alan Lambowitz. $E^{\Delta P5abc}$ and P5abc used for DMS mapping were prepared as described previously (*32*). Oligonucleotides were made by solid phase synthesis and were supplied by the Protein and Nucleic Acid Facility at Stanford University (Stanford, CA) or were used and characterized in previous studies. Oligonucleotides GUCG, rP, −3m,rP, and −1d,rSA5 (Chart 1) were purified by strong anion exchange HPLC (NucleoPac PA-100, Dionex, Sunnyvale, CA) using a linear gradient from 0.01 to 1 M ammonium acetate, pH 5, in 10% acetonitrile and desalted using Sep-Pak C18 columns (Waters, Franklin, MA). Oligonucleotides were 5′-labeled with approximately equimolar [γ-^{32}P]ATP using T4 polynucleotide kinase and purified by

Chart 1: Oligonucleotide Nomenclature[a]

	Position										
	-6	-5	-4	-3	-2	-1	+1	+2	+3	+4	+5
rP	C	C	C	U	C	U					
-3m,rP	C	C	C	mU	C	U					
-1d,rSA	C	C	C	U	C	dT	A				
-1d,rSA$_5$	C	C	C	U	C	dT	A	A	A	A	A
-1r,dSA$_5$	dC	dC	dC	dT	dC	U	dA	dA	dA	dA	dA

[a] m represents a 2′-methoxyribose moiety, and d represents a 2′-deoxyribose moiety. dT and dU behave identically at positions −1 and −3 (*46*; D. Herschlag, unpublished results).

20% nondenaturing polyacrylamide gel electrophoresis, as described previously (*38, 39*). $E^{\Delta P5abc}$ RNA was 3′-labeled by ligation with [5′-^{32}P]$_pC_P$ and purified by denaturing polyacrylamide gel electrophoresis as described (*40, 41*).

General Kinetic Methods. All reactions were single-turnover, with ribozyme (E) in excess of 5′-labeled oligonucleotide substrate ([*S] < 0.1 nM). Unless noted, reactions were performed in 50 mM sodium·MOPS, pH 6.8, at 37 °C. E was preincubated for 30 min in 10 mM $MgCl_2$ and buffer at 50 °C, followed by adjustment of Mg^{2+} (added as $MgCl_2$) to the desired concentration and addition of other reaction components at room temperature. Reactions were incubated for ≥3 min at the reaction temperature prior to initiation by addition of *S. At least six aliquots of 1-2 μL were removed from 10 or 20 μL reactions at specified times and quenched by the addition of 3 volumes of 100 mM EDTA, pH 8, in 75% formamide with 0.005% xylene cyanol and 0.005% bromophenol blue. Labeled substrate and product(s) were separated by 20% polyacrylamide/8 M urea denaturing gel electrophoresis, and their ratio at each time point was quantitated using a PhosphorImager (Molecular Dynamics, Sunnyvale, CA). Reactions were followed for 3 half-lives, or up to 24 h for slow reactions, over which time no significant deviation from first-order kinetics was detected. Time courses were fit to first-order exponential decay. Reactions with the $CCUCUA_5$ oligonucleotide consistently went to 85% completion, while those with other oligonucleotides went to 95-98% completion. The most likely explanation for the lower end point observed with $CCUCUA_5$ is that there is some unreactive contaminant in the preparation that comigrates with $CCUCUA_5$. The presence of such a contaminant is not expected to affect the rate constants obtained herein, as reactions were single-turnover with E in excess over oligonucleotide substrate. A Levenberg−Marquardt nonlinear least-squares fitting algorithm was used for all fits of models to the data (Kaleidagraph, Reading, PA, and Sigmaplot, Jandel Scientific, Corte Madera, CA). Errors are estimated to be <2-fold, on the basis of repeated experiments and measurements that utilized independent experimental approaches.

Determination of k_3. k_3 is the third-order rate constant with subsaturating GMP and E subsaturating with respect to *S, representing the reaction E + *S + GMP → products. k_3 was calculated as the product of the second-order rate constant $(k_{cat}/K_M)^{GMP}$, which describes reaction with subsaturating GMP and E saturating with respect to *S (E·*S + GMP → products), and $1/K_D^S$, the inverse of the equilibrium constant for dissociation of S and E [$k_3 = (k_{cat}/K_M)^{GMP}(1/K_D^S)$]. Reactions were carried out with 50 mM sodium·MES, pH 6.0, at 37 °C. The lower pH was used to slow the rate of the chemical step to render it rate-limiting for E^{wt}. A log-linear pH dependence with a slope of 1 for $(k_{cat}/K_M)^{GMP}$ at 10 and 100 mM $MgCl_2$ (data not shown) strongly suggests that the chemical step was rate-limiting for these reactions (*42, 43*). The same pH dependence was observed with E subsaturating with respect to S, suggesting that binding of S is in rapid preequilibrium prior to the chemical step so that the $K_{1/2}$ value obtained for S binding is equal to the equilibrium dissociation constant, K_D^S (data not shown).

To measure $(k_{cat}/K_M)^{GMP}$, oligonucleotide cleavage assays were carried out with E (0.5 μM E^{wt} or 3 μM $E^{\Delta P5abc}$) saturating with respect to trace 5′-labeled $CCUCUA_5$ and with four subsaturating GMP concentrations (0-6 and 0-300 μM for E^{wt} and $E^{\Delta P5abc}$, respectively). The reaction rate increased linearly with GMP at all Mg^{2+} concentrations, indicating that GMP was subsaturating. To measure K_D^S, oligonucleotide cleavage assays were carried out via G-independent hydrolysis (*44*) or with subsaturating GMP for E^{wt} (0-1 μM GMP) and $E^{\Delta P5abc}$ (100 μM GMP), and with six concentrations of E that varied from at least 5-fold below to at least 5-fold above the measured K_D^S.

Measurement of K_D^{GUCG}. K_D^{GUCG} is the equilibrium constant for dissociation of GUCG from E. Oligonucleotide cleavage assays were carried out with E^{wt} (1-10 nM) or $E^{\Delta P5abc}$ (10-80 nM) subsaturating with respect to trace 5′-labeled −1r,dSA_5, and at least seven concentrations of GUCG. Reaction rate increased linearly with E concentration, indicating that E was subsaturating with respect to −1r,dSA_5. A log-linear pH dependence with a slope of 1 was observed (data not shown), suggesting that the chemical step was rate-limiting for these reactions (*42, 43*). This suggests that binding of GUCG is in rapid preequilibrium prior to the chemical step so that the $K_{1/2}$ value obtained is equal to the equilibrium dissociation constant, K_D^{GUCG}.

The $K_{1/2}$ analysis was complicated by inhibition at high concentrations of GUCG. This inhibition was constant between purified preparations, and could be due to GUCG binding at inhibitory sites within the ribozyme or a persistent contaminant. The data in Figure 7 are fit to a three-site GUCG binding model, in which binding to one site increases the rate of reaction, and cooperative binding to two additional sites is inhibitory. The $E^{\Delta P5abc}$ data were not fit well by models for single-site inhibition or multiple noncooperative inhibitory sites, and models incorporating more than two cooperative inhibitory sites did not significantly improve the fit. The $K_{1/2}$ for E^{wt} was model-independent within 20%, and that for $E^{\Delta P5abc}$ is reported as a lower limit in the text, which is also model-independent. Because the GUCG-dependent and inhibitory phases are not well resolved for $E^{\Delta P5abc}$, only a limiting $K_{1/2}$ for GUCG binding to $E^{\Delta P5abc}$ could be determined.

Measurement of K_D^{rP} and $K_D^{-3m,rP}$. K_D^{rP} and $K_D^{-3m,rP}$ are the equilibrium constants for dissociation of rP and −3m,rP, respectively (Chart 1), from E, and were measured via inhibition. Single-turnover reactions were performed with 1 nM E^{wt} and 10 μM GMP or 3 nM $E^{\Delta P5abc}$ and 100 μM GMP, and with trace 5′-labeled $CCUCUA_5$ (*S). GMP is subsaturating at these concentrations. rP or −3m,rP (0-200 nM) was

incubated with E for 10 min prior to initiation of the reactions by addition of *S. The reactions displayed good first-order kinetics, and varying the preincubation time from 6 to 30 min had no significant effect on the observed inhibition, suggesting that binding of the inhibitors had achieved equilibrium. The dependence of the observed rate of reaction on the inhibitor concentration fit well to a competitive inhibition model (eq 1, in which k_{norm} is the observed first-order rate constant for substrate cleavage normalized to the rate constant observed in the absence of inhibitor). Because the inhibition experiments were performed with E subsaturating with respect to *S, the observed K_I values are expected to represent equilibrium dissociation constants for rP and −3m,rP from E. A control using labeled $CUCUA_5$, which binds ∼200-fold more weakly than $CCUCUA_5$ under these conditions (unpublished results), gave the same $K_{1/2}$ values within 40%.

$$k_{norm} = \frac{1}{\left(1 + \frac{[rP] \text{ or } [-3m,rP]}{K_I}\right)} \quad (1)$$

Measurement of k_c^{app}. k_c^{app} is the apparent first-order rate constant with a GUCG concentration that gives maximal activity and E saturating with respect to *S, representing the reaction E·*S·GUCG → products. k_c^{app} is an apparent rate constant because there is some inhibition at the concentrations of GUCG used; therefore, saturation at the reactive site may not be complete (see the Results). Reactions containing E (100 nM) saturating with respect to *S and 2 mM GUCG were initiated with trace 5′-labeled −1d,rSA (Chart 1). A single A instead of an A_5 tail was used to avoid the destabilizing interactions between GUCG and the second A of the A_5 tail of the typical oligonucleotide substrate (*45*). This interaction is not expected to affect the values of K_D^{GUCG} obtained in Figure 7, as the −1r,dSA_5 substrate was subsaturating in those experiments.

Increasing the concentration of E by 3-fold had no effect on the observed rates of reaction, and decreasing the GUCG concentration by 4-fold had less than a 2-fold effect on the reaction rates, suggesting that ribozyme was essentially saturating and that the reaction was at a maximum in its GUCG dependence. The use of a deoxyribose at the U(-1) position (see Chart 1 for numbering) decreases the rate of the chemical step by ∼1000-fold, but has no effect on tertiary binding interactions of the oligonucleotide substrate with E (*46*). This substitution thus allows the rate constant for the chemical step to be measured by manual mixing.

DMS and Kethoxal Modification. DMS modification was performed as described (*41*) in 50 mM potassium·HEPES, pH 6.8, with 1 mM EDTA, pH 8.0, and 0.5 μg/mL $tRNA^{Phe}$. Substituting sodium·MES, pH 5.8, for potassium·HEPES, pH 6.8, had no effect on the modification patterns. $E^{\Delta P5abc}$ or $E^{\Delta P5abc}$·P5abc was preincubated for 5 min at 50 °C and then 20 min at 37 °C. Reactions were initiated by the addition of 0.02 volume of a 1:2 DMS−ethanol mixture and quenched after 5 min with β-mercaptoethanol. DMS reactions performed on 3′-labeled $E^{\Delta P5abc}$ (∼25 nM) were treated with aniline to induce strand scission 3′ of guanosine residues methylated at the N7 position (*47*). Reactions performed on unlabeled RNA (100 nM) were followed by reverse transcription using 5′-labeled DNA primers to visualize modification of the N1 position of adenosines and the N3 position of cytosines (*41, 48*). Kethoxal modification to detect solvent- accessible N1 and N2 of guanosine residues was performed with 100 nM unlabeled RNA under the same conditions used for DMS modification (*49*). Reactions were initiated by addition of 0.02 volume of 250 μM kethoxal and quenched after 10 min by ethanol precipitation in the presence of 20 mM Tris−borate, pH 8.2, to stabilize the kethoxal adducts. Modifications were visualized by reverse transcription, performed as for DMS (*41, 48*). The primers selected for reverse transcription allowed positions 25−370 of the ribozyme to be probed (see Figure 2 for numbering).

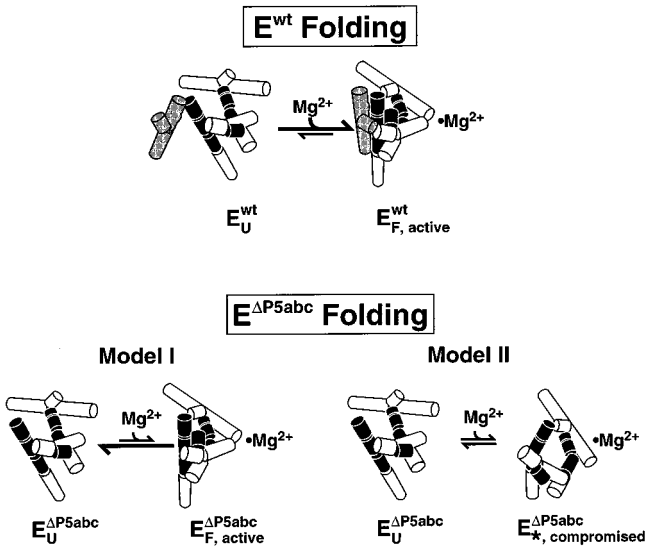

FIGURE 3: Models for ribozyme folding. The full-length ribozyme folds cooperatively to the fully active ribozyme conformation (E^{wt} folding). Two simple models describe $E^{\Delta P5abc}$ folding. Model I: $E^{\Delta P5abc}$ has an increased Mg^{2+} requirement for folding, but folds to the active conformation, analogous to E^{wt}. Model II: Instead of folding to the active structure, $E^{\Delta P5abc}$ folds to an alternative structure or population of structures ($E^{\Delta P5abc}_{*,compromised}$). E^{wt}_U and $E^{\Delta P5abc}_U$ represent inactive unfolded ribozyme lacking tertiary structure, and $E^{wt}_{F,active}$ and $E^{\Delta P5abc}_{F,active}$ represent the fully folded and active state of the ribozyme. The conserved core of the ribozyme is shown in black, P5abc is shown in gray, and other peripheral elements are shown in white.

RESULTS

Models for the Contribution of P5abc to the Active Ribozyme Fold. Previously, Inoue and co-workers showed that P5abc is essential for efficient self-splicing of the Tetrahymena group I intron, but that considerable rescue of an intron lacking P5abc could be achieved at high Mg^{2+} concentration (50 mM) (*25*). A simple model to explain this is that P5abc stabilizes the active conformation of the ribozyme relative to unfolded forms, and thus lowers the Mg^{2+} requirement for folding (Figure 3, E^{wt} folding versus $E^{\Delta P5abc}$ folding, model I). According to model I, $E^{\Delta P5abc}$ adopts the correct fold at high Mg^{2+} concentrations, but remains predominantly unfolded at lower Mg^{2+} concentrations, possibly because the ribozyme lacking P5abc does not efficiently stabilize the juxtaposition of negatively charged helix backbones present in the active structure. In contrast, according to model II, P5abc is required to selectively stabilize the active structure of the ribozyme with respect to

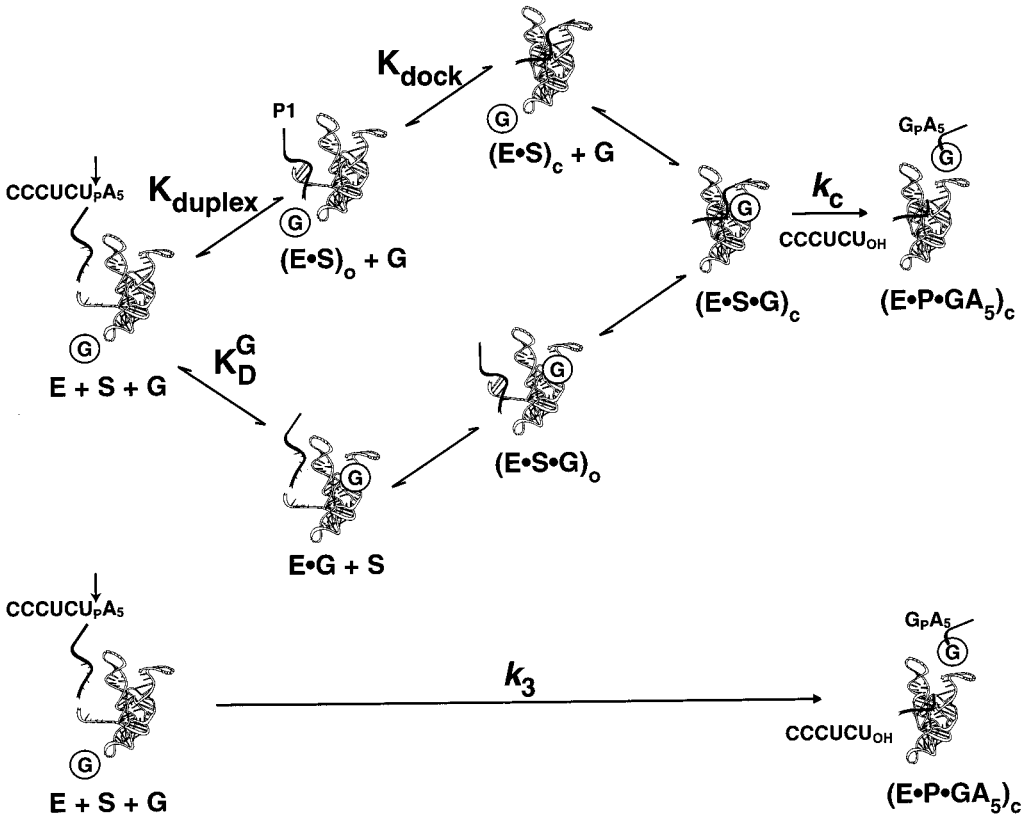

FIGURE 4: Framework for the *T. thermophila* group I ribozyme reaction (*44, 55-64*). $CCCUCU_PA_5$ is the oligonucleotide substrate (S), which is cleaved in a phosphoryl transfer reaction, with the $_PA_5$ tail transferred to G_{OH} (G), the guanosine nucleophile, to yield the products G_PA_5 and $CCCUCU_{OH}$ (P). The substrates of the ribozyme, S and G, can bind in either order to form the E·S·G ternary complex. K_{duplex} is the equilibrium dissociation constant for S binding to the IGS to form the P1 duplex in the open complex, $(E·S)_o$. Duplex formation is followed by the formation of tertiary contacts between P1 and the ribozyme core, a step referred to as docking. K_D^G is the equilibrium dissociation constant for G binding, and k_c is the rate constant for the chemical step of the phosphoryl transfer reaction which occurs from the E·S·G ternary complex. The arrow denotes the scissile bond. The tertiary structure representation is a depiction of the conserved core helices and was adapted from a drawing by L. Jaeger.

alternative conformations that are favorable in its absence. By this alternative model, $E^{\Delta P5abc}$ does not fold into the fully active conformation even at high concentrations of Mg^{2+}, but rather folds into a structure or population of structures that is functionally compromised (Figure 3, $E^{\Delta P5abc}_{*,compromised}$). To distinguish between the models for $E^{\Delta P5abc}$ folding presented in Figure 3 and investigate the role of P5abc in ribozyme catalysis, we have carried out a detailed structure−function analysis using a P5abc deletion mutant of the L-21 ScaI ribozyme derived from the *Tetrahymena* self-splicing intron.

The folded Form of $E^{\Delta P5abc}$ Is Not Fully Active. The models for $E^{\Delta P5abc}$ folding depicted in Figure 3 give rise to distinct predictions. If the folding of $E^{\Delta P5abc}$ follows model I, then the Mg^{2+} dependencies of global folding and activity should superimpose; i.e., the fully folded form (fractional protection 1) should also have full activity ($k_3 = 1$), as is true for E^{wt} (*33*). In contrast, model II predicts that Mg^{2+} will promote an equilibrium folded structure of $E^{\Delta P5abc}$ that is not the fully active structure, due to the accumulation of $E^{\Delta P5abc}_{*,compromised}$. To test these models, the Mg^{2+} dependence of the rate constant for the overall reaction (k_3), which monitors binding of both substrates and the chemical step (Figure 4, E + G + $CCUCUA_5$ (S) → products), was determined and compared to the Mg^{2+} dependence for global folding assayed previously by Fe(II)·EDTA cleavage (*32*).

The Mg^{2+} dependencies of global folding of the $E^{\Delta P5abc}$ core and of k_3 are shown in Figure 5. Protection of the $E^{\Delta P5abc}$ core from Fe(II)·EDTA cleavage increases with Mg^{2+} concentration up to about 20 mM Mg^{2+}, where it reaches a plateau, suggesting that $E^{\Delta P5abc}$ is globally folded at this Mg^{2+} concentration (*32*). The Fe(II)·EDTA cleavage pattern for $E^{\Delta P5abc}$ is very similar to that observed for E^{wt}, suggesting that deletion of P5abc does not grossly perturb the structure of the core and remaining peripheral elements (*32*; unpublished results). Despite this global structural similarity, $E^{\Delta P5abc}$ reactivity is greatly compromised, with overall activity decreased by 10^4-fold even at 100 mM Mg^{2+}. Thus, $E^{\Delta P5abc}$ folds, but does so predominantly to a less active or inactive state (Figure 3, model II).[2]

What Is the Nature of the Folded but Functionally Compromised State of $E^{\Delta P5abc}$? To further test the model that the folded structure of $E^{\Delta P5abc}$ is not the fully active structure, the disposition of the nucleoside bases in the catalytic core

[2] A trivial alternative explanation for the impaired function of $E^{\Delta P5abc}$ is that the $E^{\Delta P5abc}$ construct is unable to achieve the fully active structure because J5/5a is closed by a covalent bond between G195 and A125 instead of the P5a helix. However, the hinge region (J5/5a) of the ribozyme is highly tolerant to mutation (*50*). Further, addition of P5abc *in trans* fully rescues $E^{\Delta P5abc}$ function: the equilibrium binding constant for the guanosine and the oligonucleotide substrates and the rate constant of the chemical step are the same for $E^{\Delta P5abc}$·P5abc and E^{wt} (data not shown). Analogous full rescue was previously observed for the self-splicing reaction (*28*).

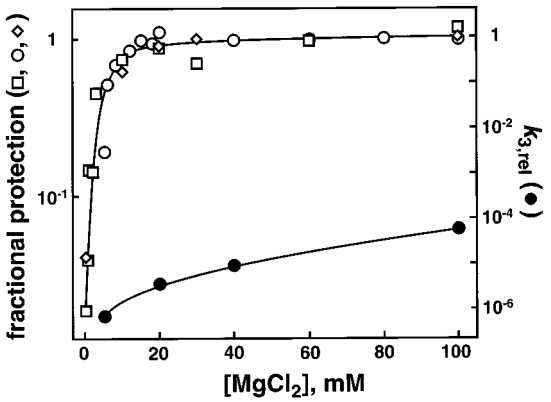

FIGURE 5: The globally folded form of $E^{\Delta P5abc}$ is not the fully active structure. Open symbols represent the solvent accessibility of residues in J7/3 (○), P3 (□), and J8/7 (◇) in the $E^{\Delta P5abc}$ ribozyme core, measured by Fe(II)·EDTA-generated hydroxyl radical cleavage (from ref *32*; J7/3 denotes the unpaired sequence joining P7 and P3). The fractional protection is the ratio of the degree of protection of $E^{\Delta P5abc}$ core residues at a given Mg^{2+} concentration relative to that at 100 mM Mg^{2+}. Fe(II)·EDTA cleavage protection analyses were performed at 37 °C in 20 mM potassium·HEPES, pH 6.8. The closed circles represent the Mg^{2+} dependence of $k_{3,rel}$, which is the ratio of the third-order rate constant for the cleavage reaction: E + S + G → products catalyzed by $E^{\Delta P5abc}$ relative to that catalyzed by E^{wt} at the same concentration of Mg^{2+}. Activity was determined at 37 °C in 50 mM sodium·MES, pH 6.0.

at 100 mM Mg^{2+} was probed by DMS and kethoxal modification and compared to results with $E^{\Delta P5abc}$·P5abc. Chemical modification was performed with $E^{\Delta P5abc}$·P5abc instead of E^{wt} to simplify side-by-side comparison with $E^{\Delta P5abc}$. E^{wt} and $E^{\Delta P5abc}$·P5abc are indistinguishable when compared by Fe(II)·EDTA protection or DMS modification (data not shown).

Consistent with the Fe(II)·EDTA cleavage data, the DMS modification patterns for the conserved core residues of $E^{\Delta P5abc}$ are similar to those of the $E^{\Delta P5abc}$·P5abc complex and do not change between 20 and 100 mM Mg^{2+} (data not shown). Also, as expected, there were differences in DMS and kethoxal modification along the P5abc contact interface in the P5, P6, and P6a helical regions and the L2 loop in $E^{\Delta P5abc}$ relative to $E^{\Delta P5abc}$·P5abc (Figure 6 and open symbols in Figure 2).

There were also differences in the DMS modification patterns for $E^{\Delta P5abc}$ and $E^{\Delta P5abc}$·P5abc at positions that lie on the opposite side of the catalytic core from the interface with P5abc, in P7, J7/3, and J8/7 (Figure 6 and closed symbols in Figure 2). Upon removal of P5abc, methylation of N1 of A306 in J8/7 is enhanced by 2-fold, while A269 in J7/3 is protected by 2-fold (Figure 6A). Also, N7 of G303 in J8/7 and G264 in P7 are protected from methylation by 2- and 3-fold, respectively (Figure 6B). In addition, modification of G303 by kethoxal is enhanced by 2-fold (Figure 6C). These reproducible changes in modification suggest that there are local differences in the structure of $E^{\Delta P5abc}$ relative to E^{wt}.

The differences in DMS and kethoxal modification cluster near the guanosine substrate binding pocket and the P1 docking site near the scissile bond. N7 of G264 interacts directly with the guanosine nucleophile (*51, 52*). In addition, chemical cleavage from Fe(II)·EDTA tethered to GMP and the 5 Å X-ray diffraction data suggest that A306, a second affected residue, lies adjacent to the guanosine binding site (*17, 53*).

The differences in chemical modification also suggest structural differences near the oligonucleotide substrate binding site. In addition to being positioned near the guanosine site, A306 is situated at the 3′-end of the J8/7 strand; this strand has been proposed to form extensive interactions with 2′-hydroxyls of the P1 duplex, which contains the oligonucleotide substrate (*13, 18*). A306 has been proposed to hydrogen bond to A269 (*17*), which also exhibits changes in DMS modification upon deletion of P5abc. In addition, metal ion rescue experiments suggest that the 5′-phosphate of A306 interacts with a functionally important divalent cation (*19*). Finally, G303 is situated in J8/7 and has been proposed to interact directly with the 2′-hydroxyl of C(−2) of the substrate (*18*; see Chart 1 for numbering).

Functional Analysis of Substrate Binding Steps. The large decrease in overall activity and the structural changes in the core of $E^{\Delta P5abc}$ support the model that $E^{\Delta P5abc}$ folds to an alternate, less active conformation and suggest that interactions with the substrates may be compromised (Figure 3, model II). To probe the functional properties of this alternatively folded state, we first compared the binding of substrates to $E^{\Delta P5abc}$ and E^{wt} (Figure 4).

The equilibrium dissociation constant for binding to the guanosine site was determined by measuring the dependence of the rate of substrate cleavage on GUCG concentration under conditions in which the chemical step is rate-limiting so that $K_{1/2}$ equals K_D (Figure 7). The 3′-terminal guanosine of GUCG binds to the G-binding site within helix P7, and the additional 5′-nucleotides make interactions with G313 and A314 to form P9.0 (Figure 2 and refs *45* and *54*). Saturation of the reaction catalyzed by E^{wt} was observed with increasing concentrations of GUCG, but the results with $E^{\Delta P5abc}$ were complicated by inhibition at high GUCG concentrations that was not well resolved from the stimulatory phase of the curve (see the Experimental Procedures). Therefore, only a limit for the dissociation constant of GUCG from $E^{\Delta P5abc}$ was obtained, from the half-maximal value of the linear stimulatory region that extends to ~600 μM GUCG (Figure 7). The conservative limit of ≥300 μM obtained for K_D^{GUCG} for $E^{\Delta P5abc}$ corresponds to a ≥25-fold decrease in binding affinity upon deletion of P5abc (Figure 7). GMP binding is similarly compromised by ≥6-fold (data not shown; inhibition was observed above 4 mM GMP).

Oligonucleotide binding was measured to determine whether tertiary interactions of the P1 duplex with the ribozyme core were also compromised upon deletion of P5abc. Oligonucleotide binding to the ribozyme occurs in two steps (Figure 4, refs *55-64*). In the first step the open complex is formed, in which the oligonucleotide is held in place solely by base-pairing interactions with the IGS of the ribozyme. In the second step, the P1 duplex docks into tertiary interactions with the ribozyme core to form the closed complex. The open complex has the stability expected for a simple duplex, whereas binding is stronger in the closed complex due to the tertiary interactions.

The docking step was isolated by comparing binding of product oligonucleotides (i.e., oligonucleotides lacking the transferred 3′-residues) with and without a 2′-methoxy

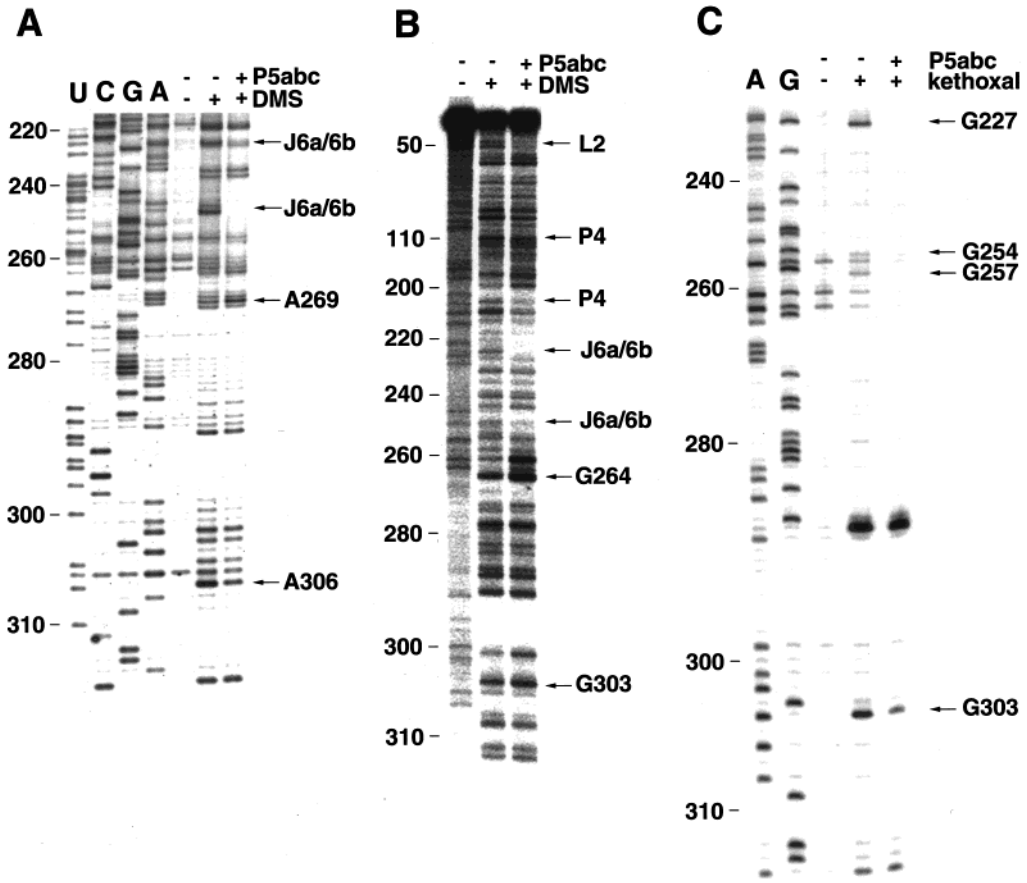

FIGURE 6: Local differences in the core structure of $E^{\Delta P5abc}$ and $E^{\Delta P5abc}$·P5abc. DMS and kethoxal modifications were performed at 37 °C, 100 mM Mg^{2+} in 50 mM potassium·HEPES, pH 6.8. Band positions were mapped by comparison to dideoxy sequencing ladders (which run one nucleotide higher than primer extension lanes). (A) DMS modification in the conserved core of $E^{\Delta P5abc}$ detected by primer extension. Primer extension of DMS-modified RNA detects methylation at N1 of adenine and N3 of cytosine. (B) DMS modification of $E^{\Delta P5abc}$ detected by aniline cleavage. Aniline cleavage of DMS-modified RNA detects methylation at N7 of guanines. The bands on the gel at non-guanosine positions likely represent degradation or nonspecific aniline cleavage. (C) Kethoxal modification of $E^{\Delta P5abc}$ detected by primer extension. Primer extension of kethoxal-modified RNA detects modification at N1 and N2 of guanosine residues.

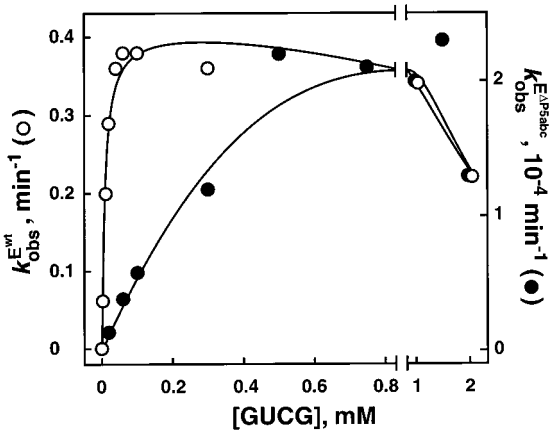

FIGURE 7: Compromised binding of GUCG to $E^{\Delta P5abc}$. The GUCG dependence of the observed rate constant is shown for the reaction E + GUCG + −1r,dSA$_5$ (^{32}P-labeled) → products, for E^{wt} (○) and $E^{\Delta P5abc}$ (●), in 50 mM sodium·MOPS, pH 6.8, at 37 °C and 100 mM Mg^{2+}. Fits to the data give $K_D^{GUCG} = 12 \pm 1$ μM for E^{wt} and $K_D^{GUCG} \geq 300$ μM for $E^{\Delta P5abc}$ (see the Experimental Procedures). Because of the inhibition at the higher GUCG concentrations, the K_D for $E^{\Delta P5abc}$ is a lower limit.

substitution at the −3 position (Figure 4; Chart 1, rP and −3m,rP). The 2′-methoxy substitution greatly destabilizes docking, by about 1000-fold, without affecting the stability of the open complex (*63*; and G. J. Narlikar and D. Herschlag, unpublished results). The product oligonucleotide was used because it forms stronger tertiary interactions than the substrate oligonucleotide with the ribozyme core, giving a larger window for comparison of E^{wt} and $E^{\Delta P5abc}$ (*64*, *65*).

Binding was assayed by measuring inhibition of the cleavage reaction over a range of product oligonucleotide concentrations under conditions such that K_I is equal to the equilibrium dissociation constant, K_D (see the Experimental Procedures). This approach allowed side-by-side comparisons of E^{wt} and $E^{\Delta P5abc}$ at each product oligonucleotide concentration. Consistent with previous results, rP bound to E^{wt} much more strongly than did −3m,rP, with K_D values of 0.02 nM and 7 nM, respectively (Figure 8). In contrast to this 350-fold stronger binding of rP to E^{wt}, rP and −3m,rP bound to $E^{\Delta P5abc}$ with the same affinity (K_D values of 18 and 19 nM, respectively). The equivalent binding affinity with and without the −3-methoxy substitution suggests that rP binds to $E^{\Delta P5abc}$ in the open complex, without stable tertiary

[3] The value of 350 is derived from comparison of the binding constants for rP and −3m,rP binding to E^{wt} and is indicative of the degree to which the docked complex is more stable than the open complex for E^{wt}. It represents a lower limit for the destabilization of tertiary interactions with $E^{\Delta P5abc}$ relative to E^{wt} because any destabilization greater than 350-fold would result in the oligonucleotide binding in the open complex, which lacks all tertiary interactions, and would not change the observed binding affinity (for a more detailed explanation, see ref *64*).

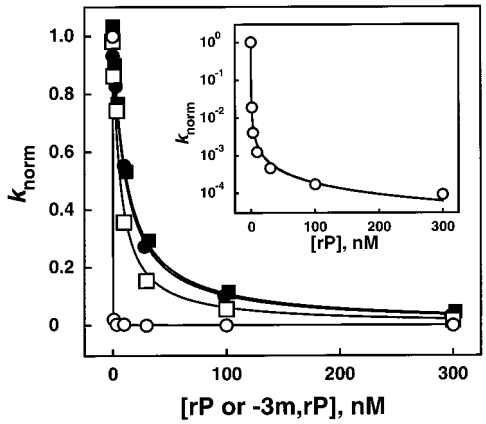

FIGURE 8: Compromised tertiary interactions between the P1 duplex and the $E^{\Delta P5abc}$ core. K_I^{rP} and $K_I^{-3m,rP}$ were measured with trace 5′-labeled $CCUCUA_5$ and subsaturating ribozyme in 50 mM sodium·MOPS, pH 6.8, at 37 °C and 100 mM Mg^{2+}. The plot shows inhibition of the reaction of E^{wt} by rP (○), E^{wt} by −3m,rP (□), $E^{\Delta P5abc}$ by rP (●), and $E^{\Delta P5abc}$ by −3m,rP (■). The symbols for $E^{\Delta P5abc}$ were moved slightly from the line so that the data for both rP and −3m,rP could be seen. Due to the strong binding of rP to E^{wt}, data for their interaction could only be obtained well above the K_I (measurement at lower inhibitor concentrations was not performed due to technical limitations in the use of ribozyme at concentrations less than 1 nM). Nevertheless, the binding constant obtained by fitting to a simple competitive inhibition model is expected to provide an accurate value for the equilibrium constant for rP dissociation from E^{wt} (*59, 65, 88*). The inset shows the concentration dependence for inhibition of E^{wt} reaction by rP, and is presented with k_{norm} on a log scale to show the fit to the data more clearly (k_{norm} is the rate constant obtained at a given inhibitor concentration normalized to the rate constant obtained with no inhibitor present). Fits to three independent data sets give $K_I^{rP}(E^{wt}) = 0.021 \pm 0.003$ nM, $K_I^{-3m,rP}(E^{wt}) = 7.4 \pm 0.7$ nM, $K_I^{rP}(E^{\Delta P5abc}) = 18 \pm 1$ nM, and $K_I^{-3m,rP}(E^{\Delta P5abc}) = 19 \pm 2$ nM.

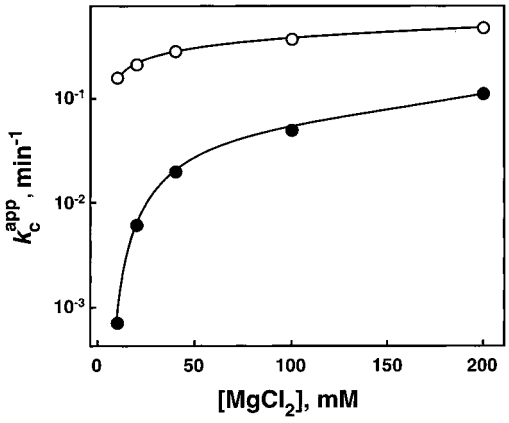

FIGURE 9: The apparent rate constant for the chemical step catalyzed by $E^{\Delta P5abc}$ approaches that measured for E^{wt} at high Mg^{2+}. k_c^{app} was measured at a GUCG concentration (2 mM) that gives a maximal rate and ribozyme saturating with respect to −1d,rSA, in 50 mM sodium·MOPS, pH 6.8, at 37 °C. A representative experiment showing the Mg^{2+} concentration dependence for the first-order rate constant for cleavage from the ternary complexes with (○) E^{wt} and (●) $E^{\Delta P5abc}$ is shown. Similar results were obtained using −1d,rSA_5 or −1d,$CCUCUA_5$ in place of −1d,rSA, and GMP in place of GUCG.

interactions. Thus, docking of the P1 duplex into the $E^{\Delta P5abc}$ core is destabilized by at least 350-fold relative to that into E^{wt}.[3]

The dependence of the rate constant for $CCUCUA_5$ cleavage on ribozyme concentration gave K_D values of 1 and 180 nM for E^{wt} and $E^{\Delta P5abc}$, respectively, providing additional evidence for a large disruption of tertiary interactions between the P1 duplex and the core in the $E^{\Delta P5abc}$ mutant (100 mM Mg^{2+}, subsaturating GMP, data not shown). The affinity of $CCUCUA_5$ for $E^{\Delta P5abc}$ is that expected for simple duplex formation between this oligonucleotide and an oligonucleotide having the IGS sequence [$K_D^{IGS} = 280$ nM at 10 mM Mg^{2+} (*66*), and a small decrease in K_D^{IGS} is expected at 100 mM Mg^{2+} (*67*)], suggesting that this oligonucleotide binds $E^{\Delta P5abc}$ in the open complex.

Effect of P5abc on the Reaction of Bound Substrates. To assess the contribution of P5abc in the chemical step, the rate constant for oligonucleotide cleavage from the E·−1d,-rSA·GUCG ternary complex was measured (Figure 9). At 10 mM Mg^{2+}, the rate constant for cleavage from the ternary complex with $E^{\Delta P5abc}$ is 200-fold lower than that with E^{wt}. However, as the Mg^{2+} concentration is increased, the reactivity of the $E^{\Delta P5abc}$ ternary complex approaches that of E^{wt}, with relative rate constants of 7 and 4 at 100 and 200 mM Mg^{2+}, respectively. Analogous results were obtained for reaction from the E·−1d,rSA_5·GMP complex (data not shown). These small differences in the observed rate constant for reaction from the ternary complex represent an upper limit for the effect on the actual chemical step, as the oligonucleotide substrate may not be fully docked (see Functional Analysis of Substrate Binding Steps above), and inhibition by GUCG and GMP prevents observation of full saturation of $E^{\Delta P5abc}$.

The ≥25-fold decrease in GUCG binding and ≥350-fold decrease in docking of the P1 duplex account for essentially all of the 10^4-fold decrease in the overall reaction of $E^{\Delta P5abc}$ observed in Figure 5 at 100 mM Mg^{2+} ($25 \times 350 = 9 \times 10^3$), consistent with at most a small effect of P5abc deletion on the chemical step.

If $E^{\Delta P5abc}$ and E^{wt} have similar rate constants for the chemical step, they are expected to have similar transition-state conformations. Substitution of hydrogen for 2′-hydroxyl groups at position G22 or G25 of the ribozyme and at position U(−1), U(−2), or U(−3) of the oligonucleotide substrate, and mutation of G22 to I22 or U(−1) to C(−1) confirm that each of these groups is involved in interactions in the chemical transition state for the $E^{\Delta P5abc}$ reaction (unpublished results), as they are for the wild-type reaction (*60, 64, 68*).

These results suggest that the chemical transition-state barriers and transition-state structures for E^{wt} and $E^{\Delta P5abc}$ complexes are similar. Therefore, the compromised overall activity of $E^{\Delta P5abc}$ is likely to arise from an energetically unfavorable $E^{\Delta P5abc}$ rearrangement from a misfolded ground-state structure or population of structures to the active structure adopted stably by E^{wt}.

DMS Modification of $E^{\Delta P5abc}$ with Bound Substrates. The similar reactivity of E^{wt} and $E^{\Delta P5abc}$ ternary complexes but not of E^{wt} and $E^{\Delta P5abc}$ starting with unbound substrates suggests that binding of the substrates helps the ribozyme adopt its active conformation. To test whether there are structural changes in $E^{\Delta P5abc}$ upon substrate binding as suggested by the preceding functional data, DMS modifications of $E^{\Delta P5abc}$ and $E^{\Delta P5abc}$·P5abc were compared in the presence and absence of product oligonucleotide and GMP (Figure 10).

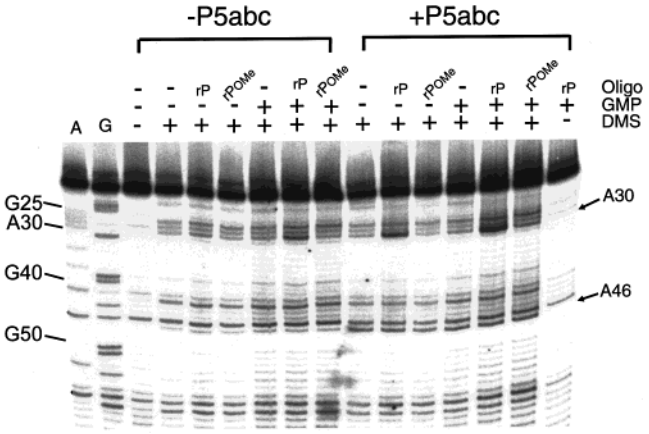

FIGURE 10: Structural effects of substrate binding on $E^{\Delta P5abc}$ and $E^{\Delta P5abc}$·P5abc. DMS modification was performed at 37 °C, 100 mM Mg^{2+} in 50 mM potassium·HEPES, pH 6.8, and detected by primer extension. Reactions contained 100 nM ribozyme and the indicated combination of 400 nM P5abc, 4 mM GMP, and 1 μM rP or −3m,rP. The lanes without DMS are controls for background cleavage and were treated the same way as the other reactions except that DMS was omitted. Numbers on the far left of the gel refer to backbone sequence positions mapped by comparison to standards generated by dideoxy sequencing.

As expected, residues G25, G26, and G27, which form Watson−Crick base pairs with C(−4), C(−5), and C(−6), respectively, of rP and −3m,rP were protected from modification by kethoxal whenever rP or −3m,rP was present with either $E^{\Delta P5abc}$ or E^{wt} (data not shown). However, changes were also observed at A30, which is located in the unpaired region between the oligonucleotide-containing P1 duplex and the P2 peripheral element (Figure 2).

Upon rP binding to $E^{\Delta P5abc}$·P5abc, a 5-fold increase in DMS modification was consistently observed at A30. This change was not observed upon binding of −3m,rP, which is expected to bind in the open complex, or upon binding of GMP alone. Also, there was no significant increase in modification at A30 upon binding of GMP to $E^{\Delta P5abc}$·P5abc·rP, consistent with the observation that E^{wt} binds rP in the closed complex even in the absence of GMP. These differences suggest that the increased modification at A30 provides a readout for docking of rP in the ribozyme core, allowing this change to serve as a structural assay for the interaction between docking and GMP binding in $E^{\Delta P5abc}$.

Binding of rP to $E^{\Delta P5abc}$ produced an increase of only 50% in DMS modification at A30, consistent with the functional results suggesting that rP binds to $E^{\Delta P5abc}$ predominantly in the open rather than closed complex. No change in DMS protection was observed upon binding of either GMP alone or −3m,rP (with or without GMP), as was observed for $E^{\Delta P5abc}$·P5abc. However, with rP and GMP both bound, modification of A30 was enhanced by 3-fold, suggesting that the two ligands act cooperatively to fold the core of the ribozyme and promote partial docking of bound rP. Analogous results were obtained using the oligonucleotide −1d,rSA_5 (data not shown).

The cooperative structural effects observed in the binding of rP and GMP to $E^{\Delta P5abc}$ suggest that there are functionally important rearrangements upon substrate binding to $E^{\Delta P5abc}$. This cooperativity is absent for E^{wt} (*69*), consistent with greater preorganization of the wild-type active site.

The smaller increase in DMS modification observed at A30 upon binding of substrates when P5abc is absent than when it is present presumably reflects a fraction of the molecules in which the P1 duplex is not fully docked and/or the GMP binding site is not fully saturated (see Effect of P5abc on the Reaction of Bound Substrates, above). This could also account for the inability of saturating rP and 4 mM GMP to restore the wild-type extent of DMS modification at residues A269, G303, and A306. Because of the smaller effects at these positions, partial restoration would not be expected to be detectable.

The remaining position that consistently displayed a difference in DMS modification in $E^{\Delta P5abc}$ relative to E^{wt}, N7 of G264, interacts directly with the guanosine nucleophile (*51*) and was protected upon GMP binding to E^{wt} (data not shown). However, removal of P5abc resulted in protection at this position even in the absence of guanosine (Figure 6B), obscuring the potential signal for guanosine binding.

DISCUSSION

We have used structural and functional probes to distinguish whether the compromised activity of $E^{\Delta P5abc}$ arises simply because the ribozyme folds less stably to the active structure than E^{wt}, or because a folded state that is inactive is favored when P5abc is absent (Figure 3). The results indicate that when $E^{\Delta P5abc}$ folds, it folds into a structure or population of structures that is not preorganized for substrate binding and catalysis. Thus, organization of the active site of the *Tetrahymena* ribozyme is linked to P5abc binding, and this interaction leads to specific stabilization of the active ribozyme structure relative to one or more alternative less active or inactive forms (Figure 3, model II).

Nature of the Folded State of $E^{\Delta P5abc}$. The Fe(II)·EDTA protection and DMS modification data suggest that $E^{\Delta P5abc}$ has the same overall structure as E^{wt} at high Mg^{2+} concentrations, but that some differences are present in the ribozyme core. The previously reported moderate reduction in protection from cleavage by Fe(II)·EDTA-generated hydroxyl radicals throughout the $E^{\Delta P5abc}$ core suggests that the core is looser or more dynamic when the interactions between P5abc and $E^{\Delta P5abc}$ are absent (*32*). Consistent with a more dynamic structure, DMS modification of $E^{\Delta P5abc}$ at A306 is increased relative to that of E^{wt}. Further, the increased protection from DMS modification upon removal of P5abc at positions G264, A269, and G303 suggests that the $E^{\Delta P5abc}$ structure is not simply looser, but that there are new interactions. Thus, the structure adopted by $E^{\Delta P5abc}$ appears to have some qualitative differences from that adopted by E^{wt}.

The results strongly suggest that the altered $E^{\Delta P5abc}$ structure has important changes within the core. The changes in accessibility of G264, A269, G303, and A306 suggest a functionally important disruption of the substrate binding sites upon deletion of P5abc, and this prediction was borne out by the functional studies that showed a substantial disruption in the binding of both substrates.

The P5abc Peripheral Element Arranges Ribozyme Structure through an Array of Organized Interactions. $E^{\Delta P5abc}$ activity increases 100-fold when the Mg^{2+} concentration is increased from 10 to 100 mM (Figure 5). Even at 100 mM Mg^{2+}, however, $E^{\Delta P5abc}$ is 10^4-fold compromised in activity compared to E^{wt}, and structural differences in the $E^{\Delta P5abc}$ core are detected by DMS modification. Why is Mg^{2+} alone

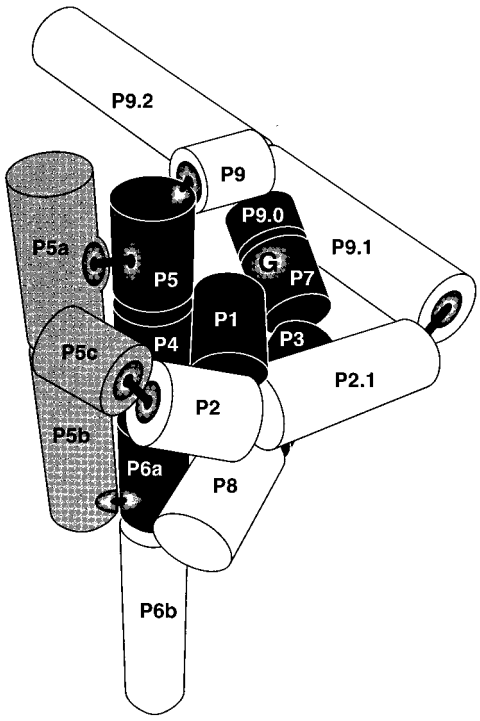

FIGURE 11: A network of peripheral element interactions within the *Tetrahymena* group I ribozyme structure. The noncovalent connections made by the guanosine substrate and the peripheral elements are shown as linked disks. The P5abc peripheral element is shown in gray, other peripheral elements are shown in white, and the conserved core helices are shown in black. Adapted from refs *15* and *17*.

unable to efficiently rescue activity? The answer may lie in the small size of the Mg^{2+} cation, which can form only local interactions with groups that fall within its coordination sphere. In contrast, the independently-folding P5abc element interacts with several distinct elements that are distant from each other in the tertiary structure, forming an interface of ~3000 Å^2 with the remainder of the ribozyme (Figure 11, refs *15*−*17* and *32*−*35*, and see below). The ability of P5abc to define the relative positions of interaction partners that are distant from each other in the tertiary fold may be the key to its ability to promote formation of the active catalytic core conformation.

What Parts of the Ribozyme Are Affected by P5abc? Structural, phylogenetic, and mutational data show that P5abc contacts the P2 peripheral element, as well as P4 and J6a/6b of the P5−P4−P6 helical stack that forms part of the catalytic core (Figure 11, refs *15-17* and *34*). As these elements either are directly connected to or are predicted to interact indirectly with the P1 duplex, changes in their relative positions upon removal of P5abc may produce part of the observed defect in P1 docking. Nevertheless, changes in DMS modification are observed at positions in P3−P7, far from sites of contact with P5abc. Further, the binding of the guanosine nucleophile to its site in P7 is severely compromised, suggesting P5abc deletion affects the structure within the catalytic core. These long-range effects of P5abc may be mediated through the network of interactions between peripheral elements shown in Figure 11 (*15, 17*). Thus, the alternative conformational state of $E^{\Delta P5abc}$ may be, in part, the result of a loosening of the constraints on the positioning of the remaining peripheral elements when P5abc is removed.

Cooperative Folding of the Ribozyme to the Active Structure. Deletion of the 3′-peripheral elements (P9.1 and P9.2) results in decreased protection of the conserved core from Fe(II)·EDTA cleavage and in changes in RNase T1 protection in the IGS, as well as large decreases in intron self-splicing activity (*30, 31, 70*). Similarly, removal of the P2 and P2.1 elements results in reduced 5′-exon cleavage activity (*27, 71*). This suggests that the 3′-peripheral elements and P2−P2.1 may, like P5abc, be important for cooperative folding to the active ribozyme structure, although measurement of individual reaction steps would be required to address this fully.

To benefit from a preorganized active site, a macromolecular catalyst must have an active fold that is thermodynamically more stable than the combination of all other possible conformations. Specific stabilization of the active fold can be achieved through cooperative interactions within the folded structure; i.e., if one component is correctly folded, it increases the stability of the correctly folded form of other components, preventing partially folded or misfolded forms from being significantly populated at equilibrium. P5abc appears to be one link in a chain of peripheral elements that wraps around the core of the *Tetrahymena* ribozyme (Figure 11); its removal results in the destabilization of the active structure relative to an inactive or less active structure or population of structures. The interactions among peripheral elements may cooperate to specifically promote formation of the active structure of the ribozyme through the covalent connections of these elements with the core and by steric effects that restrict core motions (*45*). This role has been previously suggested for peripheral elements in general by Lehnert et al. on the basis of inspection of tertiary structure models (*15*).

There Are Multiple Ways To Favor Formation of the Active Core Structure. Despite the near universal presence of peripheral elements, different group I introns vary significantly with respect to their complement of these subdomains, and there are subgroups that lack the P5abc element altogether (Figure 1, refs *8-11, 15,* and *72*). An extreme example is the 205 nucleotide *Azoarcus* intron, half the size of the 414 nucleotide *Tetrahymena* intron. Like the *Tetrahymena* intron, the *Azoarcus* intron is classified in subgroup IC, but it is missing elements P2.1, P9.1, P9.2, P5b, and P5c. This abbreviated intron may derive its exceptional stability from two loop−helix interactions that form between extensions of core helices, predicted to connect P9 to J5/5a and P2 to P8, rather than more extensive elements that wrap around the outside of the catalytic core (*72*). This suggests that there are multiple ways to provide the connectivity necessary for folding to the active structure.

There even appear to be ways to ensure formation of an active structure within the context of the *Tetrahymena* ribozyme lacking most or all of P5abc. Circularized *Tetrahymena* constructs with disruptions to two of three of the sites of specific contact between P5abc and the rest of the ribozyme can still perform efficient 3′-splice-site cleavage, perhaps due to the connectivity introduced by having covalently bound substrate (*73*). Also, the 5′-splice-site cleavage reaction catalyzed by a *Tetrahymena* intron that has an extended P1 helix is only compromised ~3-fold in activity when P5abc is deleted (*73*). Although the rate-limiting steps of the reactions were not identified in these

studies, it seems unlikely that deletion of P5abc or its elements results in an activity decrease of 10^4-fold as observed herein.

Our results suggest another means to stabilize the active structure of the ribozyme core in the absence of P5abc. The increased DMS accessibility at A30 upon binding of both the oligonucleotide and nucleophile substrates, but not upon binding of either one alone, suggests that together oligonucleotide and GMP binding can induce folding of the $E^{\Delta P5abc}$ core to make it more like that of E^{wt}. This is supported by the observation that the rate constant for chemical cleavage catalyzed by $E^{\Delta P5abc}$ is at most 4-fold compromised once both substrates are bound, suggesting the E^{wt}·S·GMP/GUCG and $E^{\Delta P5abc}$·S·GMP/GUCG ternary complexes are similar in structure. The similar core folding event induced by P5abc and substrates indicates a structural connection between P5abc and the ribozyme core, suggesting a network of mutually reinforcing interactions within the ribozyme structure.

Peripheral Elements As a General Strategy for Stabilization of Functional RNA Structures. The dearth of structural information in the primary sequence of RNA (*4*; also see the Introduction) may render it difficult to specifically stabilize a functional structure without the preorganization afforded by interconnected elements outside the catalytic core. The folding of the hammerhead ribozyme, which lacks peripheral sequences, appears to be so tenuous that mutation of almost any of the core residues is sufficient to reduce activity by more than 10^4-fold (*74*). Further, in contrast to the preorganization of the core of the *Tetrahymena* ribozyme suggested by structural and functional data (*15*, *17*, *18*, *65*), the hammerhead ribozyme lacks an active site crevice and seemingly requires a substantial conformational change prior to catalysis (*75*, *76*).

Group I introns appear to have adopted the strategy of adding interconnected structures around the outside of the functional domain to selectively stabilize their catalytic conformations. Unlike proteins, which fold around a highly stable hydrophobic core (*77*), the use of peripheral elements may be a strategy used by many large RNA molecules to stabilize functional structures. Phylogenetic comparisons of group II introns and also of bacterial RNase P RNA sequences indicate that these catalytic RNA molecules can also be divided into subclasses according to conservation of peripheral elements (*78-80*). Not all ribozymes require peripheral elements for stable folding, however. The extremely stable hepatitis delta ribozyme is held together in large part by strand-crossovers and backbone sugar interactions within a compact double pseudoknot core (*81*). This alternative strategy for structural stabilization may work less well for RNA molecules such as the group I introns that must rearrange to bind and release substrates through the course of their activity.

Proteins as Peripheral Elements. In some cases, proteins have taken the place of peripheral elements in group I introns (for a review see ref *82*). One example, CYT-18, appears to recognize conserved structural features of the multiple noncognate group I introns that it can activate for splicing (*83-86*). Its binding site overlaps with that of P5abc, and CYT-18 will give partial rescue of activity and structure for P5abc deletion mutants of the *Tetrahymena* group I intron (*86*, *87*). This conservation of peripheral elements through evolution, with proteins taking the place of RNA, highlights the importance of peripheral elements to intron function.

ACKNOWLEDGMENT

We thank members of the Herschlag and Doudna labs for helpful comments, and Marc Caprara and Alan Lambowitz for the construct used to synthesize $E^{\Delta P5abc}$.

REFERENCES

1. Herschlag, D., and Cech, T. R. (1990) *Biochemistry 29*, 10172-10180.
2. Thill, G., Vasseur, M., and Tanner, N. K. (1993) *Biochemistry 32*, 4254-4262.
3. Kurz, J. C., Niranjanakumari, S., and Fierke, C. A. (1998) *Biochemistry 37*, 2393-2400.
4. Sigler, P. B. (1975) Annu. Rev. Biophys. Bioeng. 4, 477-527.
5. Narlikar, G. J., and Herschlag, D. (1997) *Annu. Rev. Biochem. 66*, 19-59.
6. Jaeger, L., Michel, F., and Westhof, E. (1996) *Nucleic Acids Mol. Biol. 10*, 33-51.
7. Cech, T. R., and Herschlag, D. (1997) in *Catalytic RNA* (Eckstein, F., and Lilley, D. M. J., Eds.) pp 1-17, Springer-Verlag, Berlin.
8. Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A., and Scazzocchio, C. (1982) *Nature 300*, 719-724.
9. Michel, F., Jacquier, A., and Dujon, B. (1982) *Biochimie 64*, 867-881.
10. Burke, J. M., Belfort, M., Cech, T. R., Davies, R. W., Schweyen, R. J., Shub, D. A., Szostak, J. W., and Tabak, H. F. (1987) *Nucleic Acids Res. 15*, 7217-7221.
11. Michel, F., and Westhof, E. (1990) *J. Mol. Biol. 216*, 585-610.
12. Damberger, S. H., and Gutell, R. R. (1994) *Nucleic Acids Res. 22*, 3508-3510.
13. Pyle, A. M., Murphy, F. L., and Cech, T. R. (1992) *Nature 358*, 123-128.
14. Caprara, M. G., and Waring, R. B. (1993) *Biochemistry 32*, 3604-3610.
15. Lehnert, V., Jaeger, L., Michel, F., and Westhof, E. (1996) *Chem. Biol. 3*, 993-1009.
16. 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*, 1678-1685.
17. Golden, B. L., Gooding, A. R., Podell, E. R., and Cech, T. R. (1998) *Science 282*, 259-264.
18. Szewczak, A. A., Ortolevadonnelly, L., Ryder, S. P., Moncoeur, E., and Strobel, S. A. (1998) *Nat. Struct. Biol. 5*, 1037-1042.
19. Christian, E. L., and Yarus, M. (1993) *Biochemistry 32*, 4475-4480.
20. Collins, R. A. (1988) *Nucleic Acids Res. 16*, 2705-2715.
21. Suh, S. O., Jones, K. G., and Blackwell, M. (1999) *J. Mol. Evol. 48*, 493-500.
22. Latham, J. A., and Cech, T. R. (1989) *Science 245*, 276-282.
23. Heuer, T. S., Chandry, P. S., Belfort, M., Celander, D. W., and Cech, T. R. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 11105-11109.
24. Joyce, G. F., and Inoue, T. (1987) *Nucleic Acids Res. 15*, 9825-9840.
25. Joyce, G. F., van der Horst, G., and Inoue, T. (1989) *Nucleic Acids Res. 17*, 7879-7889.
26. Doudna, J. A., and Szostak, J. W. (1989) *Mol. Cell. Biol. 9*, 5480-5483.
27. Beaudry, A. A., and Joyce, G. F. (1990) *Biochemistry 29*, 6534-6539.
28. van der Horst, G., Christian, A., and Inoue, T. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 184-188.
29. Salvo, J. L., and Belfort, M. (1992) *J. Biol. Chem. 267*, 2845-2848.
30. Caprara, M. G., and Waring, R. B. (1994) *Gene 143*, 29-37.
31. Laggerbauer, B., Murphy, F. L., and Cech, T. R. (1994) *EMBO J. 13*, 2669-2676.

32. Doherty, E. A., Hershlag, D., and Doudna, J. A. (1999) *Biochemistry 38*, 2982-2990.
33. Celander, D. W., and Cech, T. R. (1991) *Science 251*, 401-407.
34. Murphy, F. L., and Cech, T. R. (1994) *J. Mol. Biol. 236*, 49-63.
35. Wu, M., and Tinoco, I., Jr. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 11555-11560.
36. Cate, J. H., Hanna, R. L., and Doudna, J. A. (1997) *Nat. Struct. Biol. 4*, 553-558.
37. Banerjee, A. R., Jaeger, J. A., and Turner, D. H. (1993) *Biochemistry 32*, 153-163.
38. Zaug, A. J., Grosshans, C. A., and Cech, T. R. (1988) *Biochemistry 27*, 8924-8931.
39. Herschlag, D., Eckstein, F., and Cech, T. R. (1993) *Biochemistry 32*, 8299-8311.
40. England, T. E., Bruce, A. G., and Uhlenbeck, O. C. (1980) *Methods Enzymol. 65*, 65-74.
41. Doherty, E. A., and Doudna, J. A. (1997) *Biochemistry 36*, 3159-3169.
42. Herschlag, D., and Khosla, M. (1994) *Biochemistry 33*, 5291-5297.
43. Knitt, D. S., and Herschlag, D. (1996) *Biochemistry 35*, 1560-1570.
44. Herschlag, D., and Cech, T. R. (1990) *Biochemistry 29*, 10159-10171.
45. Russell, R., and Herschlag, D. (1999) *RNA 5*, 158-166.
46. Herschlag, D., Eckstein, F., and Cech, T. R. (1993) *Biochemistry 32*, 8312-8321.
47. Peattie, D. A. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1760-1764.
48. Stern, S., Moazed, D., and Noller, H. F. (1988) *Methods Enzymol. 164*, 481-489.
49. Litt, M., and Hancock, V. (1967) *Biochemistry 6*, 1848-1854.
50. Szewczak, A. A., and Cech, T. R. (1997) *RNA 3*, 838-849.
51. Michel, F., Hanna, M., Green, R., Bartel, D. P., and Szostak, J. W. (1989) *Nature 342*, 391-395.
52. Lin, C. W., Hanna, M., and Szostak, J. W. (1994) *Biochemistry 33*, 2703-2707.
53. Wang, J. F., and Cech, T. R. (1992) *Science 256*, 526-529.
54. Bevilacqua, P. C., Sugimoto, N., and Turner, D. H. (1996) *Biochemistry 35*, 648-658.
55. Herschlag, D., and Cech, T. R. (1990) *Nature 344*, 405-409.
56. Pyle, A. M., McSwiggen, J. A., and Cech, T. R. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8187-8191.
57. Pyle, A. M., and Cech, T. R. (1991) *Nature 350*, 628-631.
58. Bevilacqua, P. C., Kierzek, R., Johnson, K. A., and Turner, D. H. (1992) *Science 258*, 1355-1358.
59. Herschlag, D. (1992) Biochemistry 31, 1386-1399.
60. Strobel, S. A., and Cech, T. R. (1993) *Biochemistry 32*, 13593-13604.
61. Wang, J. F., Downs, W. D., and Cech, T. R. (1993) *Science 260*, 504-508.
62. Bevilacqua, P. C., Li, Y., and Turner, D. H. (1994) *Biochemistry 33*, 11340-11348.
63. Narlikar, G. J., and Herschlag, D. (1996) *Nat. Struct. Biol. 3*, 701-710.
64. Narlikar, G. J., Khosla, M., Usman, N., and Herschlag, D. (1997) *Biochemistry 36*, 2465-2477.
65. Narlikar, G. J., Gopalakrishnan, V., McConnell, T. S., Usman, N., and Herschlag, D. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 3668-3672.
66. Narlikar, G. J., Bartley, L. E., Khosla, M., and Herschlag, D. (1999) *Biochemistry 38*, 14192-14204.
67. McConnell, T. S., Herschlag, D., and Cech, T. R. (1997) *Biochemistry 36*, 8293-8303.
68. Strobel, S. A., and Cech, T. R. (1994) *Nat. Struct. Biol. 1*, 13-17.
69. Bevilacqua, P. C., Johnson, K. A., and Turner, D. H. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8357-8361.
70. Barfod, E. T., and Cech, T. R. (1988) *Genes Dev. 2*, 652-663.
71. Ikawa, Y., Shiraishi, H., and Inoue, T. (1998) *J. Biochem. (Tokyo) 123*, 528-533.
72. Tanner, M. A., and Cech, T. R. (1996) *RNA 2*, 74-83.
73. Naito, Y., Shiraishi, H., and Inoue, T. (1998) *RNA 4*, 837-846.
74. Peracchi, A., Karpeisky, A., Maloney, L., Beigelman, L., and Herschlag, D. (1998) *Biochemistry 37*, 14765-14775.
75. Pley, H. W., Flaherty, K. M., and McKay, D. B. (1994) *Nature 372*, 68-74.
76. Wang, S., Karbstein, K., Peracchi, A., Beigelman, L., and Herschlag, D. (1999) *Biochemistry 38*, 14363-14378.
77. Dill, K. A. (1990) *Biochemistry 29*, 7133-7155.
78. Michel, F., Umesono, K., and Ozeki, H. (1989) *Gene 82*, 5-30.
79. Haas, E. S., Banta, A. B., Harris, J. K., Pace, N. R., and Brown, J. W. (1996) *Nucleic Acids Res. 24*, 4775-4782.
80. Massire, C., Jaeger, L., and Westhof, E. (1998) *J. Mol. Biol. 279*, 773-793.
81. Ferre-D'Amare, A. R., Zhou, K., and Doudna, J. A. (1998) *Nature 395*, 567-574.
82. Lambowitz, A. M., and Perlman, P. S. (1990) *Trends Biochem. Sci. 15*, 440-444.
83. Collins, R. A., and Lambowitz, A. M. (1985) *J. Mol. Biol. 184*, 413-428.
84. Akins, R. A., and Lambowitz, A. M. (1987) *Cell 50*, 331-345.
85. Mohr, G., Zhang, A., Gianelos, J. A., Belfort, M., and Lambowitz, A. M. (1992) *Cell 69*, 483-494.
86. Mohr, G., Caprara, M. G., Guo, Q., and Lambowitz, A. M. (1994) *Nature 370*, 147-150.
87. Caprara, M. G., Lehnert, V., Lambowitz, A. M., and Westhof, E. (1996) *Cell 87*, 1135-1145.
88. Knitt, D. S., Narlikar, G. J., and Herschlag, D. (1994) *Biochemistry 33*, 13864-13879.

BI992313G