Probing the Role of a Secondary Structure Element at the 5′- and 3′-Splice Sites in Group I Intron Self-Splicing: The *Tetrahymena* L-16 *Sca*I Ribozyme Reveals a New Role of the G·U Pair in Self-Splicing†

Katrin Karbstein,‡,§ Jihee Lee,|| and Daniel Herschlag*,‡,||

Department of Biochemistry and Department of Chemistry, Stanford University, Stanford, California 94305-5307

Received October 18, 2006; Revised Manuscript Received February 9, 2007

ABSTRACT: Several ribozyme constructs have been used to dissect aspects of the group I self-splicing reaction. The *Tetrahymena* L-21 *Sca*I ribozyme, the best studied of these intron analogues, catalyzes a reaction analogous to the first step of self-splicing, in which a 5′-splice site analogue (S) and guanosine (G) are converted into a 5′-exon analogue (P) and GA. This ribozyme preserves the active site but lacks a short 5′-terminal segment (called the IGS extension herein) that forms dynamic helices, called the P1 extension and P10 helix. The P1 extension forms at the 5′-splice site in the first step of self-splicing, and P10 forms at the 3′-splice site in the second step of self-splicing. To dissect the contributions from the IGS extension and the helices it forms, we have investigated the effects of each of these elements at each reaction step. These experiments were performed with the L-16 *Sca*I ribozyme, which retains the IGS extension, and with 5′- and 3′-splice site analogues that differ in their ability to form the helices. The presence of the IGS extension strengthens binding of P by 40-fold, even when no new base pairs are formed. This large effect was especially surprising, as binding of S is essentially unaffected for S analogues that do not form additional base pairs with the IGS extension. Analysis of a U·U pair immediately 3′ to the cleavage site suggests that a previously identified deleterious effect from a dangling U residue on the L-21 *Sca*I ribozyme arises from a fortuitous active site interaction and has implications for RNA tertiary structure specificity. Comparisons of the affinities of 5′-splice site analogues that form only a subset of base pairs reveal that inclusion of the conserved G·U base pair at the cleavage site of group I introns destabilizes the P1 extension >100-fold relative to the stability of a helix with all Watson−Crick base pairs. Previous structural data with model duplexes and the recent intron structures suggest that this effect can be attributed to partial unstacking of the P1 extension at the G·U step. These results suggest a previously unrecognized role of the G·U wobble pair in self-splicing: breaking cooperativity in base pair formation between P1 and the P1 extensions. This effect may facilitate replacement of the P1 extension with P10 after the first chemical step of self-splicing and release of the ligated exons after the second step of self-splicing.

The group I intron from *Tetrahymena thermophila*, the first catalytic RNA to be discovered, folds into a specific structure and performs a self-splicing reaction that includes chemical steps as well as RNA conformational rearrangements. This reaction is simple, relative to pre-mRNA splicing and other RNA-mediated processes, and may serve as a tractable system for learning more about RNA catalysis and its functional conformational changes. To date, much information about the structure and catalytic function of group I introns has been garnered, especially the originally discovered intron from *Tetrahymena thermophila* (*1*, *3*, *8*, *12*−*32*), but much less is understood about the required conformational changes (*33*, *34*).

The group I self-splicing reaction has been studied in the greatest depth with shortened ribozyme constructs derived from the self-splicing intron (*1*, *5*, *8*, *20*, *21*, *30*, *35*−*37*; see also refs *38* and *39*). The most commonly studied of these constructs is the L-21 *Sca*I ribozyme (Figure 1B). The 5′-end of this ribozyme positions the 5′-exon through formation

† This work was supported by NIH Grant GM49243. K.K. was supported in part by a predoctoral fellowship from the Boehringer Ingelheim Fonds.

* To whom correspondence should be addressed: Department of Biochemistry, Stanford University, Stanford, CA 94305-5307. E-mail: herschla@stanford.edu. Phone: (650) 723-9442. Fax: (650) 723-6783.

‡ Department of Biochemistry.

§ Present address: Department of Chemistry, University of Michigan, Ann Arbor, MI 48109-1055.

|| Department of Chemistry.

[1] Abbreviations: S, 5′-splice site analogue without specifying the identity of the 2′-substituents, which can be a hydrogen atom or a hydroxyl or methoxy group, and without specifying the length of the tail 3′ of the cleavage site (see Chart 1 for the specific oligonucleotides used); P, 5′-exon analogue (see Chart 1) without specifying the identity of 2′-substituents as for S; *S, radioactively labeled S; *P, radioactively labeled P; E, L-16 *Sca*I or L-21 *Sca*I ribozyme; PCR, polymerase chain reaction; HPLC, high-pressure liquid chromatography; IGS, internal guide sequence, the sequence of the ribozyme that is complementary to the 5′-exon (see Figure 1D); Me, methyl group (CH_3); HEPES, 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid; MES, 2-(*N*-morpholino)ethanesulfonic acid; MOPS, 3-(*N*-morpholino)propanesulfonic acid; EDTA, ethylenediaminetetraacetic acid.

10.1021/bi062169g CCC: $37.00 © 2007 American Chemical Society
Published on Web 03/27/2007

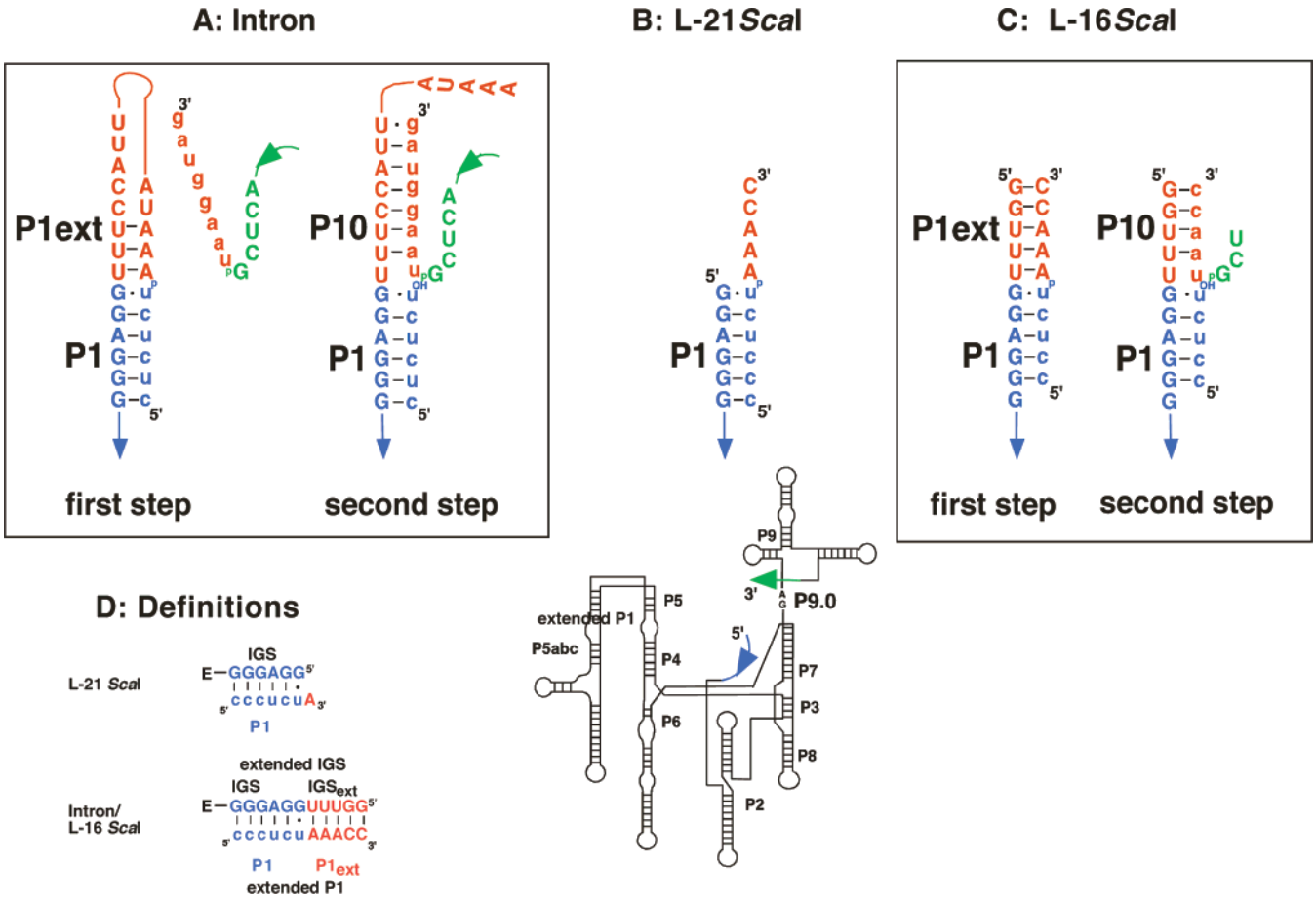


FIGURE 1: Comparison of the self-splicing RNA with the L-21 *Sca*I and L-16 *Sca*I ribozymes. Exon sequences are shown in lowercase letters, while intron or ribozyme sequences are shown in uppercase letters. The P1 duplex is colored blue (A−D), while the P1 extension and P10 (A, C, and D) are colored red. The five 3′-terminal nucleotides that are deleted in the L-16 and L-21 *Sca*I ribozymes are colored green for the intron. The secondary structure of the remaining RNA is identical in all cases and shown for orientation, with arrows denoting the covalent connections between the enlarged parts and the rest of the ribozyme. The reactive phosphoryl group is shown explicitly in each case as p. (A) Secondary structure of the region around the cleavage site in the first and second transesterification steps of self-splicing. (B) Secondary structure of the L-21 *Sca*I ribozyme around the cleavage site. (C) Secondary structure of the L-16 *Sca*I ribozyme around the cleavage site. Analogues of the 5′- and 3′-splice sites are shown as substrates, with U·A and U·U pairs, respectively, following the G·U wobble pair. Relative to the L-21 *Sca*I ribozyme, the L-16 *Sca*I ribozyme contains a five-nucleotide 5′-extension ($^{5'}$GGUUU$^{3'}$); shown in panel A are the naturally occurring nucleotides at this position ($^{5'}$CCUUU$^{3'}$). The 5′-GG residues that replace the 5′-CC residues in the L-16 *Sca*I ribozyme were included to facilitate transcription by T7 RNA polymerase. The natural 5′-exon analogue contains a U five residues upstream of the cleavage site, and the resulting G·U wobble pair appears to strengthen the tertiary binding interactions with the P1 duplex (*6*−*8*). Herein, as in previous work, we have replaced this U with a C residue. Finally, to prevent binding from being too strong, we have removed the 5′-most C of the 5′-exon mimic for most of the reactions studied herein. (D) Definition of the nomenclature used herein. The additional sequence present in the self-splicing intron and the L-16 *Sca*I ribozyme, but absent in the L-21 *Sca*I ribozyme, extends the IGS and is therefore termed the IGS extension. Together, the IGS and the IGS extension are termed the extended IGS. Formation of base pairs with the IGS results in formation of the P1 duplex (colored blue), which can be formed with both the L-21 and L-16 *Sca*I ribozymes. Analogously, in the L-16 *Sca*I ribozyme, formation of base pairs with the IGS extension results in formation of the P1 extension (colored red). Together, P1 and the extended P1 duplexes are called the extended P1 duplex.

of base pairs and is therefore called the internal guide sequence (IGS,[1] Figure 1D). However, the L-21 *Sca*I ribozyme lacks a 5′-extension of the IGS, present in the self-splicing intron, that has the potential to form a longer helix (P1 and the P1 extension form the extended P1 duplex) during the first step of splicing (compare panels A and B of Figure 1). Thus, the immediate environment of the cleavage site differs between the L-21 *Sca*I ribozyme and the self-splicing intron. Furthermore, the 5′-extension of the IGS forms another helix in the second step of self-splicing, termed P10. This helix is formed between the IGS extension and the 3′-exon [Figure 1A (*40*−*43*); see also ref *44*)]. Thus, base pairs in the P1 extension must be broken for self-splicing to proceed, and the P1 extension and P10 are dynamic secondary structure elements in self-splicing (Figure 1A).

To understand the kinetic and thermodynamic contributions from the IGS extension, the P1 extension, and P10 to the self-splicing reaction, we have employed an IGS-extended ribozyme. This construct, termed the L-16 *Sca*I ribozyme (Figure 1C), retains the ability to form the P1 extension and P10. Rate and equilibrium constants for individual steps in the reaction of the L-16 *Sca*I ribozyme with different bound substrates have been compared to those previously obtained with the L-21 *Sca*I ribozyme. This stepwise approach has allowed us to isolate the effects from the IGS extension and formation of the extended P1 or the P10 duplex on each step of the ribozyme reaction.

Table 1: Binding and Dissociation Rate Constants for 5′-Splice Site Analogue Substrates[a]

IGS sequence ($^{3'}$GAGGUUUGG$^{5'}$)	k_{off} (min^{-}1)	K_d (nM)	$\Delta\Delta G$ (kcal/mol)
CCUCdTMe	1.5[b]	15[c]	−0.2
CCUCdTA	1.9[c]	19[c]	0
CCUCdTAAACC	2.5×10^{-4}	0.0025	−5.4
CCUCdTAAAAA	0.17	1.7	−1.5

[a] Dissociation rate constants from the E•S complex at 30 °C and pH 6.0. Equilibrium dissociation constants were calculated from the measured dissociation rate constants and the association rate constant of 1×10^8 M^{-1} min^{-1} for CCUCUAAACC and CCUCUAAAAA (data not shown). This association rate constant reflects the formation of the ribozyme−substrate duplex and is the same, within error, as that determined previously for the L-21 *Sca*I ribozyme (*1*, *2*, *5*). Deoxyribose substitution at position −1 (Chart 1) has no effect on the association rate constant (*57*). Values are averages of two or more independent experiments giving rate constants that were the same within 3-fold. [b] The dissociation rate constant for CCUCdTMe is too fast to be measured accurately and was therefore calculated from the measured dissociation rate constant for CCCUCUMe, with the effect from the deletion of −6C (Chart 1), and the deoxyribose substitution at position −1 identified by comparing the dissociation rate constants for CCUCdTAAAAA and CCCUCdTAAAAA (for −6C) and CCCUCdTA and CCCUCUA (for the −1 deoxyribose). [c] The dissociation rate constant for CCUCdTA is too fast to be measured accurately and was therefore calculated from the dissociation constant for CCCUCdTA, accounting for the effect from the −6 base pair as described in footnote b.

MATERIALS AND METHODS

Materials. The L-16 *Sca*I ribozyme, compared to the L-21 *Sca*I ribozyme, has a five-nucleotide extension of the internal guide sequence (IGS) at the 5′-end ($^{5'}$GGUUU$^{3'}$; see Figure 1 for relation of L-16 *Sca*I and L-21 *Sca*I ribozymes and the full-length intron). The naturally occurring sequence at these positions is $^{5'}$CCUUU$^{3'}$. The 5′-GG in the L-16 *Sca*I ribozyme replacing the 5′-most CC residues are included to facilitate transcription by T7 RNA polymerase. The DNA template for the L-16 *Sca*I ribozyme was constructed from the L-21 *Sca*I-containing plasmid by PCR mutagenesis and ligation into pUC19. Individual clones were sequenced to verify the L-16 *Sca*I sequence.

Conditions for transcription of the L-16 *Sca*I ribozyme were modified to prevent endonucleolytic cleavage, which apparently occurs at the 5′-end to yield the L-21 *Sca*I ribozyme (refs *45* and *46* and data not shown). Transcriptions were carried out for 30 min at 30 °C in the presence of 4 mM $MgCl_2$, NTPs (0.5 mM each), 40 mM DTT, 2 mM spermidine, 40 mM Tris (pH 8.1), 0.01% Triton X-100, 5 μg/mL *Sca*I-digested, purified template, and T7 RNA polymerase. Reactions were stopped by the addition of excess EDTA, and mixtures were ethanol precipitated, resuspended, and purified over an RNeasy column (Qiagen) as described previously (*5*, *47*). Analytical 5′-^{32}P phosphorylation and comparison with authentic L-21 *Sca*I confirmed that the full-length L-16 *Sca*I ribozyme was obtained (∼95%). This conclusion is further supported by the functional data that show that the L-16 *Sca*I ribozyme binds stronger to fully base paired oligonucleotides than to oligonucleotides in which the 5′-most residues on the ribozyme cannot form base pairs (Table 1).

RNA oligonucleotides were purchased from Dharmacon Research, Inc. (Lafayette, CO), and DNA oligonucleotides were purchased from Operon Technologies Inc. (Alameda, CA). RNA and DNA oligonucleotides were 5′-end-labeled with [γ-^{32}P]ATP and purified with nondenaturing gels using standard procedures (*48*). Unlabeled oligonucleotides were purified by anion exchange HPLC and desalted on Sep-Pak C-18 columns (Waters, Franklin, MA) as described previously (*5*).

General Kinetic Methods. All reactions were single-turnover reactions with ribozyme in excess of 5′-^{32}P-labeled S or P and, unless otherwise noted, were carried out at 30 °C in 50 mM Na-MES (pH 6.0) and 10 mM $MgCl_2$. Prior to reaction, the ribozyme was preincubated for 30 min at 50 °C and pH 6.0 with 10 mM $MgCl_2$ to allow folding to the active state (*1*, *49*). Control experiments performed by incubating the radiolabeled L-16 *Sca*I ribozyme and monitoring its length on denaturing polyacrylamide gels confirmed that no significant processing of the L-16 *Sca*I ribozyme occurred during folding (data not shown). Reactions were initiated by addition of ribozyme, and aliquots were removed at specified times to be quenched with 2 volumes of a solution containing 20 mM EDTA and 85% formamide. Radiolabeled oligonucleotides were separated by denaturing gel electrophoresis (7 M urea and 20% acrylamide) and quantitated using Phosphorimager analysis (Molecular Dynamics) with ImageQuant. For slow reactions, rate constants were obtained from initial rates with end points of 95% assumed. All other reactions were followed to completion, and good first-order fits to the data with end points of ≥90% were obtained in all cases ($R^2 \geq 0.98$) (Kaleidagraph, Synergy Software, Reading, PA).

Dissociation Rate Constants for Substrates and Product. Dissociation rate constants for *S (see Chart 1 for oligonucleotides that were used) in the absence of guanosine and for P in the presence and absence of UCGAAACC[2] were determined in pulse−chase gel-shift experiments essentially as described previously (*5*, *36*). Oligonucleotides with a 2′-deoxyribose substitution at the cleavage site were used to prevent reaction during the dissociation reaction. Trace *S (or *P) was bound to saturating amounts of ribozyme (pulse), and a large excess of unlabeled S (chase) was then added. Control experiments in which the chase was preincubated with the labeled oligonucleotide verified that the chase was effective in preventing rebinding of dissociated *S (or *P). At specified times, aliquots were loaded onto a running native gel in THEM buffer (33 mM Tris, 67 mM HEPES, 1 mM EDTA, and 10 mM $MgCl_2$) to separate bound and free *S (or *P). Dissociation rate constants were obtained by fitting the fraction of free *S (or *P) to a double exponential. Double-exponential fitting was necessary because a small fraction of ribozyme (∼8%) released *S or *P considerably faster than the bulk of the population, consistent with release from the L-21 *Sca*I ribozyme. Presumably, these ribozymes had lost their IGS extension through self-processing (*44*−*46*). Control experiments monitoring the length of radioactively labeled L-16 *Sca*I confirmed that L-16 *Sca*I is not substantially processed during the dissociation reactions (≥80% remaining, data not shown; for reference, processing of 50% in parallel with dissociation, to give an apparent dissociation rate constant that is approximately the sum of the rate constants for dissociation and processing, would

[2] The guanosine residue that binds in the guanosine binding site is underlined throughout for clarity.

Chart 1[a]

Abbreviation	Oligonucleotide
IGS sequence[b]:	*$_{3'}$rG rG rG rA rG rG rU rU rU rG rG$_{5'}$*
Position:	**-6 -5 -3 -1 +1 +3 +5**
-1d,$^{\Delta}$rSA$_3$C$_2$	rC rC rU rC dT rA rA rA rC rC
-1d,$^{\Delta}$rSUA$_2$C$_2$	rC rC rU rC dT rU rA rA rC rC
-1d,$^{\Delta}$rSA$_5$	rC rC rU rC dT rA rA rA rA rA
-1d,$^{\Delta\Delta}$rSA$_5$	rC rU rC dT rA rA rA rA rA
-1d,rSA$_5$	rC rC rC rU rC dT rA rA rA rA rA
-1r,dSA$_5$	dC dC dC dU dC rU dA dAdA dA dA
dSA$_5$	dC dC dC dU dC dT dA dAdA dA dA
-1d,$^{\Delta}$rSA	rC rC rU rC dT rA
-1d, rSA	rC rC rC rU rC dT rA
rSA	rC rC rC rU rC rU rA
rSMe	rC rC rC rU rC rU CH$_3$
-3m,-1d,$^{\Delta}$rSA$_3$C$_2$	rC rC mU rC dT rA rA rA rC rC
-3m,-1d, rSA	rC rC rC mU rC dT rA
-1d,$^{\Delta}$P	rC rC rU rC dT
d$^{\Delta}$P	dC dC dT dC dT

[a] r = 2′-OH. d = 2′-H. m = 2′-OCH$_3$. [b]The IGS sequence is listed in the 3′ to 5′ direction to show complementarity to the listed oligonucleotides.

result in an only 2-fold deviation of the observed dissociation rate constant from the actual rate constant).

The rate constant for dissociation of *S from the E·*S·G complex (or *P from the E·*P·UCGAAACC complex) was obtained from experiments that measured the partitioning of molecules in the E·*S·G (or E·*P·UCGAAACC) complex between dissociation of *S (k_{off}^{S} or k_{off}^{P} for dissociation of *P) and reaction of the ternary complex to form E·*P+GA (k_c or k_{-c} for the reaction of E·*P·UCGAAACC to form E·*S+UCG). To accurately determine the rate constant for dissociation in these experiments, the values of k_{off}^{S} and k_c must be similar. To assess the partitioning, radiolabeled *S (or *P) was added to the ribozyme (final concentration of 50 nM) to allow formation of E·*S (or E·*P). Reaction was initiated by addition of G ([G] = 2 mM; K_d^G = 35 μM) or UCGAAACC ([UCGAAACC] = 10 μM; $K_d^{UCGAAACC}$ = 108 nM) to rapidly form the E·*S·G (or E·*P·UCGAACC) complex, along with a large excess (470 nM) of unlabeled S, to prevent rebinding of dissociated *S or *P. The observed rate constant for formation of *P from *S (or *S from *P) is the sum of the rate constants for the chemical step and dissociation. Thus, by independently measuring the rate constant for the chemical step in reactions with the chase omitted, we can calculate the dissociation constant (eq 1A).

$$k_{off}^{S} = k_{obs} - k_c \quad (1A)$$

$$\text{frac}_{reacted} = \frac{k_c}{k_c + k_S^{off}} \quad (1B)$$

Alternatively, the dissociation rate constant can be obtained from the fraction of *S that reacted relative to reactions in which no chase was added ($\text{frac}_{reacted}$) and the observed rate constant of the reaction in the presence of the chase (k_{obs}, eqs 1A and 1B). The results from the two methods agreed within 2-fold. The dissociation rate constant for −1d,$^{\Delta}$P (Chart 1) in the presence of UCGAAACC obtained in partitioning experiments and pulse−chase gel-shift experiments (see above) agreed within 2-fold.

Association Rate Constants for S. Association rate constants for S in the absence of G were determined in pulse−chase gel-shift experiments (*5*, *36*). Varying concentrations of ribozyme (2−10 nM) were added to trace *S to initiate the binding reaction ("pulse"). After specified times (t_1), a large excess of unlabeled S (340 nM) was added to prevent any additional binding of *S ("chase"). Control experiments confirmed that the chase effectively prevented subsequent binding of *S. Plotting the fraction of bound *S against t_1 yields a first-order rate constant, which varies with ribozyme concentration. Plotting the first-order rate constants against the ribozyme concentration at which they were obtained gives the second-order association rate constant.

Previous work with the L-21 *Sca*I ribozyme has shown that in the presence of saturating concentrations of G the rate of reaction of rS is limited by the association of S (*1*, *50*) so that $(k_{cat}/K_M)^S$ values represent the association rate constants for S in the presence of G. Moreover, $(k_{cat}/K_M)^S$ values represent the association rate constant if $k_{off} \ll k_c$ (*51*). Because these conditions hold for the L-16 *Sca*I ribozyme as well, $(k_{cat}/K_M)^S$ values were measured to obtain association rate constants in the presence of G. The similarity of the $(k_{cat}/K_M)^S$ values, which represent the formation of the P1 duplex, for the L-21 *Sca*I and L-16 *Sca*I ribozymes (1.5×10^8 and 1.0×10^8 M^{-1} min^{-1}, respectively) further supports the notion that $(k_{cat}/K_M)^S$ values describe the association rate constants for S. To obtain $(k_{cat}/K_M)^S$ values, the reaction rate constant was measured using varying concentrations of ribozyme (2−10 nM). The observed rate constants were plotted as a function of ribozyme concentration to yield the apparent second-order rate constant $(k_{cat}/K_M)^S$.

Measurement of the Equilibrium Constant for Docking. S and P bind to the ribozyme in two steps: base pairing to form the P1 duplex between S (or P) and the ribozyme [$(E\cdot S)_{open}$] followed by docking of this duplex into tertiary interactions at the active site [$(E\cdot S)_{open}$, eq 2A; e.g., refs *8*, *35*, and *52*].

$$E + S \overset{K_{d,open}}{\rightleftharpoons} (E\cdot S)_{open} \overset{K_{tert}}{\rightleftharpoons} (E\cdot S)_{closed} \quad (2A)$$

$$K_{d,obs} = \frac{K_{d,open}}{1 + K_{tert}} \quad (2B)$$

2′-Hydroxyl residues on S (and the IGS) as well as the exocyclic amino group of G22 of the ribozyme's IGS stabilize docking of the ribozyme−substrate duplex (*2*, *7*, *22*, *53*, *54*). Modification of the 2′-hydroxyl group at position −3 to a 2′-methoxy group destabilizes the docked ribozyme−substrate duplex conformation while not affecting the stability of the P1 duplex (*7*, *35*, *55*). Thus, substrate analogues with this modification approximate the affinity of the duplex between substrate and IGS (for $K_{tert} < 1$, $K_{d,obs}^{undocked\ S} \approx K_{d,open}$; eq 2B). Comparing the observed affinity of substrates S* that have been undocked in this way and are thus held only by base pairing interactions ($K_{d,obs}^{undocked\ S}$) to the affinity of the wild-type substrate ($K_{d,obs}^{wt\ S}$) allows the evaluation of tertiary interactions that stabilize the docked conformation for the wild-type substrate [eq 2C obtained from eq 2B for the wild-type substrate (*5*, *22*, *35*, *56*)].

$$K_{tert} = \frac{K_{d,open}}{K_{d,obs}^{wt\ S}} - 1 \qquad (2C)$$

Measurement of Affinities for G and GAAACC Analogues. To determine the affinities of ribozyme complexes for G or UCG, the rate of reaction of *S was determined with 0−2000 μM G (or 0−1000 μM UCG). A substrate with a 2′-deoxyribose residue at position −1 was used in these experiments to ensure that the chemical step was rate-limiting (*23*, *57*). The observed rate constant for cleavage of *S was plotted as a function of G (or UCG) concentration and fit to eq 3.

$$k_{obs} = \frac{k_{max}[G]}{[G] + K_{1/2}^{G}} \qquad (3)$$

Nonlinearity of ribozyme concentration dependences was observed at low ribozyme concentrations (≤1 nM), suggesting nonspecific losses of ribozyme to tube walls. To avoid these losses, ribozyme concentrations of ≥2 nM were used in all experiments described herein. This meant, however, that S was bound to the ribozyme under all equilibrium conditions. To obtain the affinity of G for the free ribozyme, we used a 5′-shortened oligonucleotide, CUCdTAAAAA. The presence of the 5′-terminal nucleotides affects binding of oligonucleotides via formation of a base pair, but docking of the resulting P1 duplex is not affected (*58*). At 2 nM L-16 *Sca*I ribozyme, most CUCdTAAAAA is unbound ($K_d \sim$ 100 nM; data not shown), as suggested from the rate of reaction compared to that for a substrate that is fully bound to the ribozyme. Binding of G and UCG to the open complex was assessed with 5′-splice site analogues with a 2′-methoxyribose substitution at position −3 (−3m,rSA$_5$, Chart 1) or with 2′-deoxyribose substitutions at all positions other than the cleavage site [−1r,dSA$_5$, Chart 1 (*4*, *35*)].

To determine the affinity of UCGA and its analogue UCGAAACC for E·P, the rate of formation of *S from *P was followed as a function of UCGA or analogue concentration, as described above ([UCGA] = 10−4000 μM; [UCGAAACC] = 0.02−2 μM). In these experiments, caution was taken to ensure that UCGAAACC remained in excess of E by lowering the ribozyme concentration to 10 nM when necessary. A 5′-exon analogue (P) with a 2′-deoxyribose substitution at position −1 was used to ensure that the chemical step was rate-limiting in these experiments. To measure the affinity of UCGA or UCGAAACC for free L-16 *Sca*I ribozyme, the shortened all-deoxyribose 5′-exon analogue d$^{\Delta}$P (Chart 1) was used (*5*). d$^{\Delta}$P binds weakly to the ribozyme so that at 10 nM ribozyme most of the d$^{\Delta}$P is unbound ($K_d \sim$ 20 nM; data not shown). The dissociation constant obtained in this way is the weighted average of binding to E and E·P, and the reported K_d values have been adjusted to account for the fraction of E that has P bound. In addition, due to coupled binding of P and UCGAAACC, binding of UCGAAACC increases the affinity for P so that at high UCGAAACC concentrations most E will have P bound as well as UCGAAACC (see Results and Discussion). The dissociation constant K_d^{UCGA} in Figure 2D is therefore a limit. The rate of reaction was determined as a function of UCGA or UCGAAACC concentration as described above ([UCGA] = 200−8000 μM; [UCGAAACC] = 0.1−10 μM).

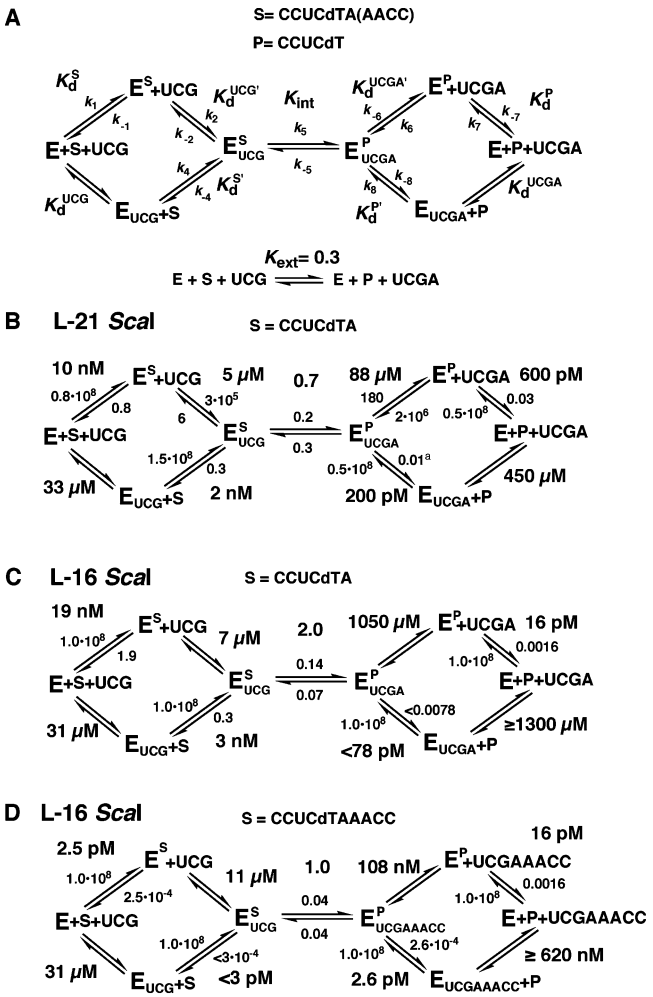
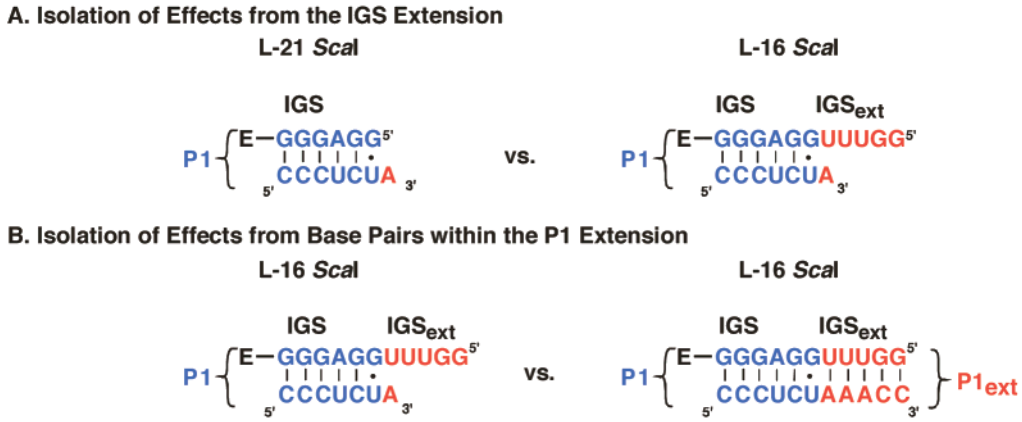


FIGURE 2: Kinetic and thermodynamic framework for the ribozyme reaction. (A) Definition of rate and equilibrium constants used herein. Association rate constants (k_1, k_4, k_7, and k_8) are reported in M^{-1} min^{-1}; all other rate constants are in units of min^{-1}. The internal and external equilibrium constants are unitless and describe conversion of substrates to products. Frameworks are defined at 30 °C in 10 mM $MgCl_2$ at pH 6.2. (B) Kinetic and thermodynamic framework for the L-21 *Sca*I ribozyme with CCUCdTA bound. This framework was adapted from ref *5*. Rate constants for dissociation of CCUCdTA and CCUCdT were determined herein for direct comparison with the L-16 *Sca*I reactions. (C) Kinetic and thermodynamic framework for the L-16 *Sca*I ribozyme with CCUCdTA bound. (D) Kinetic and thermodynamic framework for the L-16 *Sca*I ribozyme with CCUCdTAAACC bound.

Scheme 1: (A) Isolation of Effects from the IGS Extension and (B) Isolation of Effects from Base Pairs within the P1 Extension



In the case of binding of UCGA, inhibition was observed at higher oligonucleotide concentrations. This inhibition precluded determination of a reliable value for the K_d so that only a lower limit is reported in Figure 2C.

Measurement of the Rate Constant for the Chemical Step and the Internal Equilibrium. The rate constant for the chemical step in the forward (and reverse) direction was measured as described previously (*5*). In these reactions, the rate constant for conversion of the tertiary E·*S·UCG complex to E·*P+UCGAAACC (or E·*P·UCGAAACC to E·*S+UCG) was followed ([E] = 50 nM, and $K_d^{-1d,rP}$ = 2.6 pM; [UCG] = 100 μM, and K_d^{UCG}= 11 μM; [UCGAAACC] = 2 μM, and $K_d^{UCGAAACC}$ = 108 nM). Control experiments in which the concentrations of UCG (or UCGAAACC) were varied indicated that the ternary complexes were quantitatively formed. S and P with a 2′-deoxyribose substitution at position −1 were used to ensure that the chemical step was rate-limiting (*57*). This conclusion was confirmed by the observation of a log-linear pH dependence between pH 6.2 and 8.1 (data not shown and refs *5*, *59*, and *60*).

The internal equilibrium constant, K_{int}, which describes the equilibrium between the ternary E·S·UCG and E·P·UCGAAACC complexes, was obtained from the rate constant for the chemical step in the forward and reverse direction (k_c and k_{-c}, respectively; eq 4).

$$K_{int} = \frac{[E{\cdot}P{\cdot}UCGAAACC]}{[E{\cdot}S{\cdot}UCG]} = \frac{k_c}{k_{-c}} \quad (4)$$

Measurement of the External Equilibrium. The external equilibrium constant K_{ext} describes the equilibrium between unbound substrates and products (eqs 5A and 5B). The external equilibrium constant was determined as described previously (*5*), except that the cleavage site also contained a deoxyribose substitution at position −1 (dSA_5 and dP) to relate it to the internal equilibrium obtained with −1-deoxyribose-substituted oligonucleotides. Briefly, the L-21 *Sca*I ribozyme was used to catalyze the equilibration of all deoxyribose S and P forms at subsaturating concentrations of UCG and UCGd(AAAAA). At low ribozyme concentrations, most S and P is unbound [20−50 nM; $K_d \approx$ 85 and 150 nM for binding of dSA_5 and dP to the L-21 *Sca*I ribozyme, respectively (*2*, *5*)] so that the external equilibrium is observed (ref *5* and data not shown). Previous control experiments showed that the presence of the deoxyribose residues at positions other than the cleavage site and the length of the "tail" 3′ to the cleavage site have no effect on the equilibrium constant, so external equilibrium constants obtained with d(CCCTCTAAAAA) can be compared to internal equilibrium constants obtained with r(CCCUC)dTr(AAACC) (*5*).

$$E + S + UCG \overset{K_{ext}}{\rightleftharpoons} E + P + UCGA \quad (5A)$$

$$K_{ext} = \frac{[E][P][UCGAAACC]}{[E][S][UCG]} \quad (5B)$$

$$frac_P = \frac{K_{ext}}{K_{ext} + \frac{[UCGAAACC]}{[UCG]}} \quad (5C)$$

Reactions were carried out with 20−50 nM E, 1−5 μM UCG, and 5−40 μM $UCGdA_5$. *S or *P was used in a trace amount so that UCG and UCGd(AAAAA) concentrations were essentially unchanged throughout the course of the reaction. The value of K_{ext} was calculated from the fraction of *P formed from *S at different UCGd(AAAAA):UCG ratios according to eq 5C, which was derived from eq 5B. Control experiments in which the ribozyme concentration or the concentrations of UCG and UCGd(AAAAA) were varied while a constant [UCGd(AAAAA)]:[UCG] ratio was maintained gave faster reaction but no change within 8% in the observed extent of reaction, indicating that substrates and products were subsaturating. The same value of K_{ext}, within 15%, was obtained from reactions initiated with *P.

Measurement of RNA Duplex Stabilities. Melting temperatures were measured using a Varian Cary 1 UV−vis spectrophotometer equipped with a thermo programmer. RNA samples at various concentrations were prepared in 1 M NaCl, 50 mM sodium phosphate buffer, and 10 mM EDTA (pH 7.0) and transferred to a quartz cuvette, and the cuvette was covered with a Teflon cap. To anneal duplexes, samples were heated to 95 °C and then slowly cooled to 10 °C. Duplex stability was measured subsequently by heating the sample at a rate of 1 °C/min while recording the absorbance at 260 nm every 1 °C. MeltWin version 3.0 (*61*) was used to fit the resulting absorbance data to obtain the melting temperatures of each sample. The thermodynamic constants ΔH and ΔS were determined by linear fitting of

the reciprocal of the melting temperature (T_m) against the logarithm of concentrations (C) using eq 6A. The value of ΔG at 30 °C was calculated from ΔH and ΔS using eq 6B.

$$\frac{1}{T_m} = \frac{R}{\Delta H}\ln\left(\frac{C}{4}\right) + \frac{\Delta S}{\Delta H} \quad (6A)$$

$$\Delta G = \Delta H - T\Delta S \quad (6B)$$

$\Delta\Delta G$ values represent the stabilization of an A·U pair and a G·C pair, respectively, relative to a G·U pair and were calculated by subtracting the stability of the latter from the stability of the former (eq 6C).

$$\Delta\Delta G = \Delta G_{AU} - \Delta G_{GU} \text{ or } \Delta\Delta G = \Delta G_{GC} - \Delta G_{GU} \quad (6C)$$

RESULTS AND DISCUSSION

Several ribozymes have been constructed in an effort to learn more about group I intron self-splicing (*1*, *36*−*39*, *62*). The most-studied construct, the L-21 *Sca*I ribozyme, lacks 21 nucleotides at the 5′-end of the intron and five nucleotides at the 3′-end [corresponding to a *Sca*I restriction site in the encoding plasmid, Figure 1A,B (*62*)]. Previous investigations added back the 3′-sequence to assess the effect and importance of the P9.0 helix and the 3′-terminal guanosine residue (*36*, *37*). Here we take an analogous stepwise approach, adding five 5′-terminal nucleotides to give the L-16 *Sca*I ribozyme. This ribozyme allows formation of the P1 extension, base pairing that is analogous to base pairing between the 5′-most portion of the intron and the extended IGS (Figure 1A,C; see below for a description of the nomenclature). This construct also allows formation of P10, a helix that forms in the second step of self-splicing between the extended IGS and the 3′-exon (Figure 1A,C).

We have used pre-steady-state kinetics to first dissect the contributions arising from the presence of the IGS extension and then from formation of base pairs with this extended sequence. This dissection is accomplished by comparing kinetic and thermodynamic reaction frameworks for the L-16 *Sca*I ribozyme reaction with appropriate substrates to that of the L-21 *Sca*I ribozyme reaction (Figure 2). The second section addresses the effects from the presence and absence of a base pair immediately 3′ to the cleavage site, as this difference reflects the scenario in the first and second chemical steps of self-splicing, respectively (Figure 1A,C).

Nomenclature

Compared to the L-21 *Sca*I ribozyme, the L-16 *Sca*I ribozyme has a five-nucleotide 5′-extension. Because this sequence extends the IGS (internal guide sequence), it is called the IGS extension herein (Figure 1D). Together, the IGS and the IGS extension (IGS_{ext}) are called the extended IGS. While binding of S (or P) to the L-21 *Sca*I ribozyme results in formation of P1 only (blue in Figure 1D), the IGS extension in the L-16 *Sca*I ribozyme allows formation of another helix, termed the P1 extension ($P1_{ext}$) herein (red in Figure 1D). By analogy to the extended IGS above, the P1 and P1 extension helices together are called the extended P1 duplex.

Kinetic and Thermodynamic Frameworks for the L-16 ScaI Ribozyme Reaction

We first describe practical considerations for obtaining the kinetic and thermodynamic frameworks. The details of the pre-steady-state kinetic methods that were employed are given in Materials and Methods and follow directly from previous studies (*1*, *5*, *6*, *23*, *35*, *36*). We then present and compare the frameworks. The rate and equilibrium constants are defined in Figure 2A, and the data presented below are summarized in Figure 2B−D.

Equilibrium binding constants for UCG, UCGA, and UCGAAACC were determined from the concentration dependences of the single-turnover reaction rate constants as described in Materials and Methods. The affinity for S and P could not be determined directly by varying the ribozyme concentration. This is because these oligonucleotides bind very strongly to the ribozyme (Table 1 and Figure 2), and nonspecific losses to the tube walls occur at very low ribozyme concentrations. Instead, we measured association and dissociation rate constants and calculated affinity constants from these (*5*, *26*, *36*). Because the dissociation rate constant of the fully base paired 5′-splice site analogue CCCUCdTAAACC is too slow to be measured accurately at 30 °C, we instead used 5′-splice site analogues that lacked the 5′-terminal C residue (−6C$^{\Delta}$, Chart 1). Previous work has suggested that this residue contributes to binding solely via base pairing interactions (*50*) and is therefore not expected to affect the comparisons herein. This conclusion is supported by comparisons at 50 °C with full-length 5′-splice site analogues (i.e., containing residue −6C), which produced effects similar to those obtained with −6C$^{\Delta}$ 5′-splice site analogues at 30 °C (Table S1 of the Supporting Information).[3]

Effects from Extending the IGS on Individual Steps in the Ribozyme Reaction. Effects from lengthening the IGS without formation of any new base pairing interactions have been dissected through comparison of the framework for the reaction of CCUCdTA with the L-21 *Sca*I ribozyme and the L-16 *Sca*I ribozyme (Figure 2B,C and Scheme 1A; see also ref *5*). Residue A(+1) of CCUCdTA can in principle form a base pair with the L-16 *Sca*I ribozyme, but results described below suggest that this base pair does not form. Reactions with this oligonucleotide can therefore be used to determine the effects on the reaction kinetics and thermodynamics that arise from the IGS extension that is present in the L-16 *Sca*I ribozyme but absent in the L-21 *Sca*I ribozyme (Scheme 1A).

Binding of the 5′-Exon Analogue P Is Strengthened by the IGS Extension. Binding of S to free ribozyme is unaffected or only slightly affected by the IGS extension (K_d^S = 19 and 10 nM for binding of CCUCdTA to the L-16 and L-21 *Sca*I ribozymes, respectively; Figure 2B,C). CCUCdTMe and CCUCdTA bound to the L-16 *Sca*I ribozyme with the same affinity, within error (K_d^S = 15 and 19 nM, respectively; Table 1). This similarity suggested that the base pair between A(+1) and U(21) in the IGS extension (Scheme 1) is not formed; CCUCdTA was used subsequently to dissect effects that arise from the presence of the IGS extension without formation of the P1 extension.

The 5′-exon analogue (P) binds ~40-fold stronger to the free L-16 *Sca*I ribozyme than to the L-21 *Sca*I ribozyme

[3] The 5′-splice site analogues used herein and in previous work (e.g., refs *1*−*5*) also replace a G·U wobble pair 5 bp upstream of the cleavage site with a G·C pair. It appears that this wobble pair contributes to tertiary binding of the P1 helix (*6*−*8*).

Table 2: Binding of G and UCG to the L-21 and L-16 *Sca*I Ribozymes[a]

	K_d (μM)								
	$E{\cdot}S_{open}$			E·SA			$E{\cdot}SA_3C_2$		
	L-21	L-16	ratio	L-21	L-16	ratio	L-21	L-16	ratio
G	930 ± 54	802 ± 67	1.1	124 ± 4	59 ± 9	2.1	76 ± 16	35 ± 3	2.2
UCG	33 ± 2	31 ± 7	1.1	5 ± 2	7 ± 1	0.7	28 ± 7	11 ± 2	2.5
ratio	28	26		25	8		2.7	3.2	

[a] All experiments were performed at 30 °C and pH 7.2 in the presence of 10 mM $MgCl_2$. Comparisons between L-21 *Sca*I and L-16 *Sca*I ribozymes were performed side by side and repeated two or more times. No significant difference was detected for binding of UCG to $E{\cdot}S_{open}$ with −1d,-3m,rSA or −1r,dSA_5 bound (K_d^{UCG} = 75 and 85 μM, respectively).

(K_d^P = 16 and 600 pM, respectively; Figure 2B,C). This large additional energetic contribution to binding of P was surprising, especially as there was essentially no change in binding of S. Previous work with the L-21 *Sca*I ribozyme has shown that binding of P is stronger than binding of S, and a metal ion that coordinates the 3′-OH group of P has been suggested to be responsible for this effect (*5*, *25*, *63*). It is possible that this metal ion interaction is further optimized on the L-16 *Sca*I ribozyme. Alternatively, the IGS extension could be forming fortuitous interactions with the core when P is bound. To account for the effect on binding of P but not S, these fortuitous interactions would have to be sensitive to the presence of the reactive phosphoryl group such that they would be prevented from forming with the phosphoryl group present. Regardless, these results underscore the high sensitivity of RNA interactions to local variations and rearrangements.

Binding of G Is Similar with and without the IGS Extension. Binding of G to the L-16 *Sca*I and L-21 *Sca*I ribozymes with the substrate bound in the open complex is similar (K_d^G = 930 ± 54 and 802 ± 67 μM for binding to L-16 $ScaI_{open}$ and L-21 $ScaI_{open}$, respectively; Table 2). With bound S, the binding of G to the L-16 *Sca*I ribozyme is ∼2-fold stronger than binding to the L-21 *Sca*I ribozyme (2.1- and 2.2-fold stronger binding to the L-16 *Sca*I ribozyme with CCUCdTA and CCUCdTAAACC bound, respectively; Table 2).

Formation of P9.0. The P9.0 duplex is formed between the residues directly 5′ of the 3′-terminal G residue of the intron and single-stranded residues of the intron adjacent to the G-binding site. Comparison of the affinity of UCG, which can form P9.0, and G for the L-16 *Sca*I ribozyme with CCCUCdTA bound in the open complex shows that UCG binding is stronger than G binding, suggesting that P9.0 is formed on the L-16 *Sca*I ribozyme akin to its formation on the L-21 *Sca*I ribozyme (*47*, *64*) (K_d^G = 930 ± 54 μM and K_d^{UCG} = 33 ± 2 μM for binding to the L-21 *Sca*I ribozyme, and K_d^G = 802 ± 67 μM and K_d^{UCG} = 31 ± 7 μM for binding to the L-16 *Sca*I ribozyme each with −1d,rSA bound in the open complex; Table 2). Furthermore, the energetic contribution from formation of P9.0 is similar for the L-21 *Sca*I and L-16 *Sca*I ribozymes (28- and 26-fold stronger binding to L-21 *Sca*I and L-16 *Sca*I ribozymes, respectively; Table 2).

Previous work has shown that interference between P9.0 and the residue at position A(+2) of the 3′-tail of the 5′-splice site analogue weakens binding of UCG (ref *47* and Table 2). The observation that the contribution from P9.0 is also reduced when the longer substrates are used on the L-16 *Sca*I ribozyme suggests that formation of a duplex with the +2 residue does not eliminate the interference (only 2.7- and 3.2-fold stronger binding of UCG vs G to E·CCUCdTAAACC with the L-21 *Sca*I and the L-16 *Sca*I ribozymes, respectively; Table 2**)**. Furthermore, the reduced contribution from P9.0 on the L-16 *Sca*I ribozyme with CCUCdTA bound suggests that an analogous interference with P9.0 can arise from the IGS extension, which includes a residue complementary to residue +2 (25- and 8-fold stronger binding of UCG vs G to the L-21 *Sca*I and L-16 *Sca*I ribozymes, respectively, both with CCUCdTA bound in the closed complex; Table 2).

In the recently determined structure of the *Azoarcus* intron with P10, the distance between P9.0 and P10 (which contains both the IGS extension and the substrate tail) exceeds 8 Å, too long to explain steric interference via a direct effect (*65*). It is possible that these distances are different in the *Tetrahymena* and *Azoarcus* RNAs or in different conformations of the ribozyme. Alternatively, the interference may not be direct but mediated by other parts of the ribozyme. Residues of P9 come within less than 4 Å of P9.0, making it the closest neighbor of P9.0 and thereby a candidate for mediating indirect interactions.

Binding of UCGA. Above, we have shown that CCUCdTA can be used to isolate effects from the IGS extension only, as the A(+1) residue does not base pair with U21 of the extended IGS. Analogously, UCGA can be used to dissect the effects from the IGS extension on the reverse reaction if the corresponding base pair with U21 is not formed. To investigate this potential effect, we compared the contribution of the base pairs with the IGS extension when formed with CCUCdTAAACC or UCGAAACC relative to CCUCdTA or UCGA, respectively. If the base pair between the IGS extension and UCGA were formed, the increase in stability for UCGAAACC relative to UCGA would be expected to be larger than the increase in stability of CCUCdTAAACC relative to CCUCdTA. This is because for CCUCdTAAACC part of the energy from formation of the base pairs in the P1 extension must pay for an energetic penalty to allow formation of the first A·U pair (see below). If this A·U pair were preformed in UCGA, formation of the additional base pairs in the P1 extension would not incur this energetic penalty for helix initiation and thus would be predicted to provide additional stabilization. However, the increased affinity for CCUCdTAAACC relative to CCUCdTA is the same, within error, as the increased affinity for UCGAAACC relative to UCGA (Table 3). These results, although indirect, suggest that the 3′-A of UCGA does not form its potential base pair with the IGS extension. UCGA was therefore used

Table 3: Thermodynamic Contribution from the P1 Extension[a]

oligonucleotide	K_d (μM)	$\Delta\Delta G$ (kcal/mol)
CCUCdTA	3.0×10^{-3}	−5.4
CCUCdTAAACC	0.4×10^{-6} [b]	
UCGA	1050	−5.5
UCGAAACC	0.11	

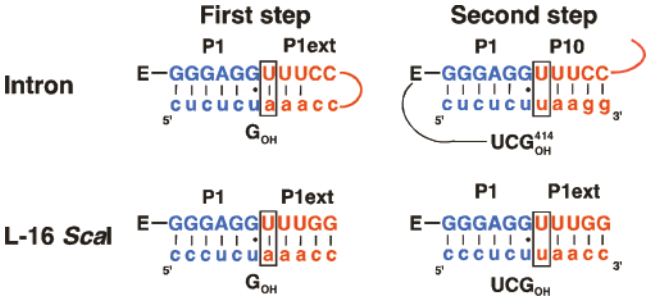
[a] Dissociation constants from E·S·G and E·P·UCGA(AACC) ternary complexes were obtained at 30 °C. S and P had 2′-deoxyribose substitutions at position −1. [b] Calculated from the rate constant for dissociation of S from E·S and the observation of 6-fold cooperativity of S and G binding with this oligonucleotide substrate (Table 2). Direct measurement was not possible because dissociation was slower than the reaction that forms P.

to probe the reverse of the L-16 *Sca*I ribozyme reaction without base pairs to the IGS extension formed.[4]

UCGA binds to the L-16 *Sca*I ribozyme with P bound in the closed complex with a dissociation constant of 1050 μM, 12-fold weaker than the binding to the L-21 *Sca*I ribozyme (Figure 2B,C**)**. This destabilization could arise from interference, direct or indirect, between P9.0 and the IGS extension, as suggested previously and above for binding of UCG (*47*). With the L-21 *Sca*I ribozyme, UCGAAAAA binding is 3-fold weaker than UCGA binding ($K_d^{UCGAAACC} = 220$ μM and $K_d^{UCGA} = 88$ μM for binding to E·P; data not shown), consistent with an interfering effect with the extended IGS. Alternatively, it is possible that the weaker binding of UCGA to the L-16 *Sca*I ribozyme is due to an increased contribution from an unfavorable interaction with the reactive phosphoryl group that was previously suggested for the L-21 *Sca*I ribozyme[5] (*5*).

Effects of Formation of the Extended P1 Duplex on Individual Steps in the Ribozyme Reaction. In the previous sections, we have shown that extending the IGS strengthens binding of G and P ~2- and ~40-fold, respectively, while binding of UCGA is weakened 12-fold. We next present results for the L-16 *Sca*I ribozyme with CCUCdTAAACC bound (Figure 2D). These data suggest that the P1 extension is formed but with a decreased stability relative to a continuously stacked duplex. Furthermore, formation of the P1 extension has no significant effect on the kinetics or equilibria of subsequent steps.

Scheme 2



Effects from the formation of the P1 extension were isolated through comparison of the frameworks of the L-16 *Sca*I ribozyme reactions with CCUCUdTA and CCUCdTAAACC bound (Scheme 1B). This was possible as the IGS extension is present in both cases, but only CCUCdTAAACC, not CCUCdTA, forms the base pairs of the P1 extension.

Formation of the Extended P1 Duplex. To test formation of the base pairs with the IGS extension, we compared the affinity of a series of 5′-splice site analogues (Table 1). The 10^4-fold stronger binding of CCUCdTAAACC (nucleotides with complementarity to the IGS extension of the L-16 *Sca*I ribozyme are underlined in this section) relative to CCUCdT-Me suggests that at least some of the potential base pairs are formed. We next compared the affinities of 5′-splice site analogues that could form only subsets of the additional P1 base pairs. The 600-fold stronger binding of CCUCdTAAACC compared to CCUCUdTAAAAA suggests that the 3′-terminal C·G base pairs are formed. The stronger affinity of CCUCUdTAAAAA versus CCUCdTMe suggests that at least some of the three A·U base pairs are formed. Nevertheless, addition of these three potential base pairs gave an only 10-fold increase in affinity, whereas an effect of ~1000-fold, ~100-fold more, would be expected for this addition to a Watson−Crick duplex (*9*). This difference is consistent with the absence of stronger binding with a single A(+1) relative to a methyl group in this position as discussed above. To test whether A(+1) formed a base pair in the longer duplex, we changed the base pairing at this position from an A·U pair to a U·U pair, using CCUCdTUAACC. This mismatched oligonucleotide base pair resembles the ligated exon product of the second step of self-splicing (Scheme 2).

Changing +1A to +1U in the context of the base-paired oligonucleotide (CCUCdTAAACC and CCUCdTUAACC) decreases the affinity for the oligonucleotide 100-fold ($\Delta\Delta G$ = 2.4 kcal/mol;[6] Table 4). These data suggest that in the context of a duplex the +1 A·U base pair is formed. In contrast, a single A·U pair at that position is not formed, as described above. These observations suggest that the G·U wobble pair that precedes the A·U pair does not behave like

[4] If an A·U base pair were formed, the observed 12-fold weaker binding of UCGA due to the IGS extension would be an underestimate.

[5] The IGS extension adds five nucleotides to give the L-16 *Sca*I ribozyme. Therefore, the 5′-terminal triphosphate group, which is present from transcription, is farther from the active site. The proximity of the triphosphate to the cleavage site, its large negative charge, and its capacity to bind a divalent metal ion raised the possibility that the differences between the L-16 and L-21 *Sca*I ribozyme originated from displacement of the triphosphate group rather than the added nucleotides. To test this possibility, we prepared the L-21 *Sca*I ribozyme with a 5′-terminal hydroxyl group (L-21 *Sca*I^{OH}) by transcription with a 10-fold excess of G over GTP. The presence of a predominantly (≥85%) 5′-OH-containing ribozyme was confirmed via incorporation of [^{32}P]ATP at the 5′-end without preincubation with phosphatase to remove the 5′-phosphate (data not shown). Binding of G and P to L-21 *Sca*I^{OH} and to L-21 *Sca*I was tested as these equilibria showed the largest difference between the L-21 *Sca*I and L-16 *Sca*I ribozymes. The dissociation constants for G and for P from L-21 *Sca*I^{OH} and L-21 *Sca*I were the same within error (K_d^G = 87 and 93 μM and K_d^P = 460 and 600 pM for binding to L-21 *Sca*I^{OH} and L-21 *Sca*I, respectively; Figure S1A and Figure S1B) and differed significantly from those for the L-16 *Sca*I ribozyme [K_d^G = 35 μM, and K_d^P = 16 pM (Figure 2C and Table 2)], suggesting that the presence of additional 5′-nucleotides rather than the absence of the nearby triphosphate group is responsible for the observed effects.

[6] Because in the absence of G the +1U-containing substrate makes nonproductive tertiary interactions that contribute ~10-fold to its affinity, the affinities of CCUCdTUAACC and CCUCdTAAACC are compared in the presence of G, which disrupts these alternative interactions (see Probing the Effect of a U·U Mismatch after the Cleavage Site). This analysis isolates the contribution of the A·U base pair to the stability of the reactive complex.

Table 4: Binding of the +1 U·U Mismatched Substrate[a]

	$\Delta\Delta G_{bp}$ (kcal/mol)			
CCUCdT-	without G	with G	k_c (min^{-1})	K_{tert}
-AAACC	−5.4[b]	−5.4	0.039	23
-UAACC	−3.9	−3.0	0.041	230

[a] All experiments were performed at 30 °C and pH 7.2. All substrates contained a 2′-deoxyribose substitution at the cleavage site (position −1) to prevent cleavage. K_{tert} values were obtained by comparison of docked and undocked substrates as described in Materials and Methods, with oligonucleotide substrates undocked by the introduction of a 2′-methoxyribose substitution at position −3 (*26*, *35*, *55*). $\Delta\Delta G_{bp}$ values were obtained by comparison of the affinity of the CCUCdTAAACC or CCUCdTUAACC oligonucleotide with CCUCdTA. [b] Value from Table 1 included for comparison.

Table 5: Effect of Introduction of a G·U Pair in Model Duplexes[a]

	internal			terminal		
	GU	AU	GC	GU	AU	GC
ΔG (kcal/mol)	−14.6	−15.3	−15.8	−9.6	−9.2	−10.2
$\Delta\Delta G$ (kcal/mol)	0	−0.7	−1.2	0	0.4	−0.6

[a] Duplex stabilities (ΔG) were measured in thermal denaturation experiments in 1 M NaCl, 10 mM EDTA, and 50 mM sodium phosphate (pH 7.0). The sequence of the duplexes was as follows (in the 5′ to 3′ direction, with the base pair being varied underlined): GCCUCUAAAC/GUUUGGAGGC, GCCUCUAAAC/GUUUAGAGGC, and GCCUCCAAAC/GUUUGGAGGC for duplexes containing internal GU, AU, and GC pairs and GCCUCU/GGAGGC, GCCUCU/AGAGGC, and GCCUCC/GGAGGC for duplexes with terminal GU, AU, and GC pairs, respectively. $\Delta\Delta G$ values represent stabilities for AU and GC pairs relative to GU pairs. $\Delta\Delta\Delta G$ values represent the additional destabilization of internal GU pairs relative to terminal GU pairs and are calculated by subtracting the $\Delta\Delta G$ values for terminal GU pairs from the $\Delta\Delta G$ values for internal GU pairs.

a Watson−Crick base pair for which formation of a subsequent A·U pair would be favored.

The observation that the G·U wobble pair at the cleavage site differs from Watson−Crick base pairs is consistent with observations from structural work on model duplexes containing a G·U wobble pair. This work shows that stacking of the G·U wobble pair with the 3′-helix (relative to the G, in this case the P1 duplex) is continuous, while stacking with the 5′-helix (relative to G, in this case the P1 extension) is altered and involves only the G and the opposite strand of the helix (the substrate "3′-tail"), with the U and its opposite strand (the IGS extension) remaining unstacked on its 3′-side (*66*−*69*; see also ref *70*).

To evaluate the ribozyme results, we compared the destabilization arising from an internal G·U pair (relative to a G·C or A·U pair) to that of a terminal G·U pair in model duplexes (Table 5). The destabilization was calculated by comparing the stabilities of model duplexes with internal and terminal G·U, A·U, and G·C pairs. Duplex stabilities were measured in thermal denaturation experiments as described in Materials and Methods. The destabilization from an internal G·U pair relative to an A·U pair or a G·C pair is 0.7 or 1.2 kcal/mol, respectively, while a terminal G·U pair is stabilized relative to an A·U pair by 0.4 kcal/mol and destabilized relative to a G·C pair by only 0.6 kcal/mol (Table 5**)**. This suggests that there is an added destabilizing effect from internally located G·U pairs that amounts to 0.6−1.1 kcal/mol ($\Delta\Delta\Delta G_{AU}$ = 1.1 kcal/mol, and $\Delta\Delta\Delta G_{GC}$ = 0.6 kcal/mol). Thus, this energetic analysis suggests that internal G·U wobble pairs behave differently from terminal G·U pairs, in that they have an additional detrimental effect that is not explained simply by the change in base pairing. These results are consistent with the structural studies of duplexes with internal G·U pairs described above (*66*−*69*; see also ref *70*).

The data on model duplexes show that internal G·U duplexes are intrinsically destabilized. This intrinsic destabilization appears to be exacerbated when the extended P1 duplex is bound to the L-16 ribozyme, where the destabilization amounts to >2.8 kcal/mol[7] instead of 0.6−1.1 kcal/mol. This difference suggests that the ribozyme may make use of features in the duplex architecture to further destabilize the duplex with an internal G·U wobble pair demarcating the cleavage site G·U pair.

Consistent with these biochemical and structural results on model duplexes and the extended P1 duplex in the *Tetrahymena* ribozyme, the recently determined X-ray crystal structures of the group I intron from *Azoarcus* also exhibit disrupted helical stacking at the G·U wobble pair (*65*, *71*). The potential importance of partial unstacking of the extended P1 duplex for self-splicing is discussed in Summary and Implications.

Accommodation of the Chemical Transition State within a Helix. In the L-21 *Sca*I ribozyme, the cleavage site is located at the end of a helical segment, whereas for self-splicing and for the L-16 *Sca*I ribozyme with CCUCdTAAACC bound, the cleavage site is embedded in a duplex (Figure 1). This constraint within a duplex could increase or decrease the rate of the chemical step if the helical geometry were preferred or not accommodated in the transition state, respectively. We therefore compared the rate constants for the forward and reverse chemical steps (k_5 and k_{-5}, respectively) with a fully base paired substrate (CCUCUdTAAACC) and a substrate that does not form a base pair in the P1 extension (CCUCdTA). Log-linear pH dependences with a slope of 1 between pH 6.2 and 8.1 suggest that the chemical step is rate-limiting (*59*, *60*) in these experiments (data not shown). The rate constants for the reaction of the base-paired and non-base-paired substrates are similar (k_5 = 0.04 and 0.14 min^{-1} and k_{-5} = 0.04 and 0.07 min^{-1} for CCUCdTAAACC and CCUCdTA, respectively; Figure 2C,D).

The similar rate constants suggest that the constraint on the helical conformation by the extended P1 duplex has at most a small effect on adoption of the transition-state structure. This effect could be small because the transition-state conformation can be reached within the helix or because the extended P1 duplex is not highly constrained due to the limited stacking following the G·U wobble pair at the cleavage site.

Cooperativity in Binding of Reactants. Previous work with the L-21 *Sca*I ribozyme showed that binding of S strength-

[7] This value is calculated from the observed 10-fold stronger binding to the L-16 *Sca*I ribozyme of CCUCdTAAAAA relative to CCUCdTMe compared to the >1000-fold stronger binding calculated from nearest neighbor rules for extended P1 duplexes with an A·U base pair or the >4000-fold stronger binding calculated for extended P1 duplexes with a G·C pair replacing the cleavage site G·U pair (*9*). Comparison of model duplex data to nearest neighbor rules suggests that nearest neighbor estimates are accurate for these specific sequences (Table 5 and refs *9*−*11*).

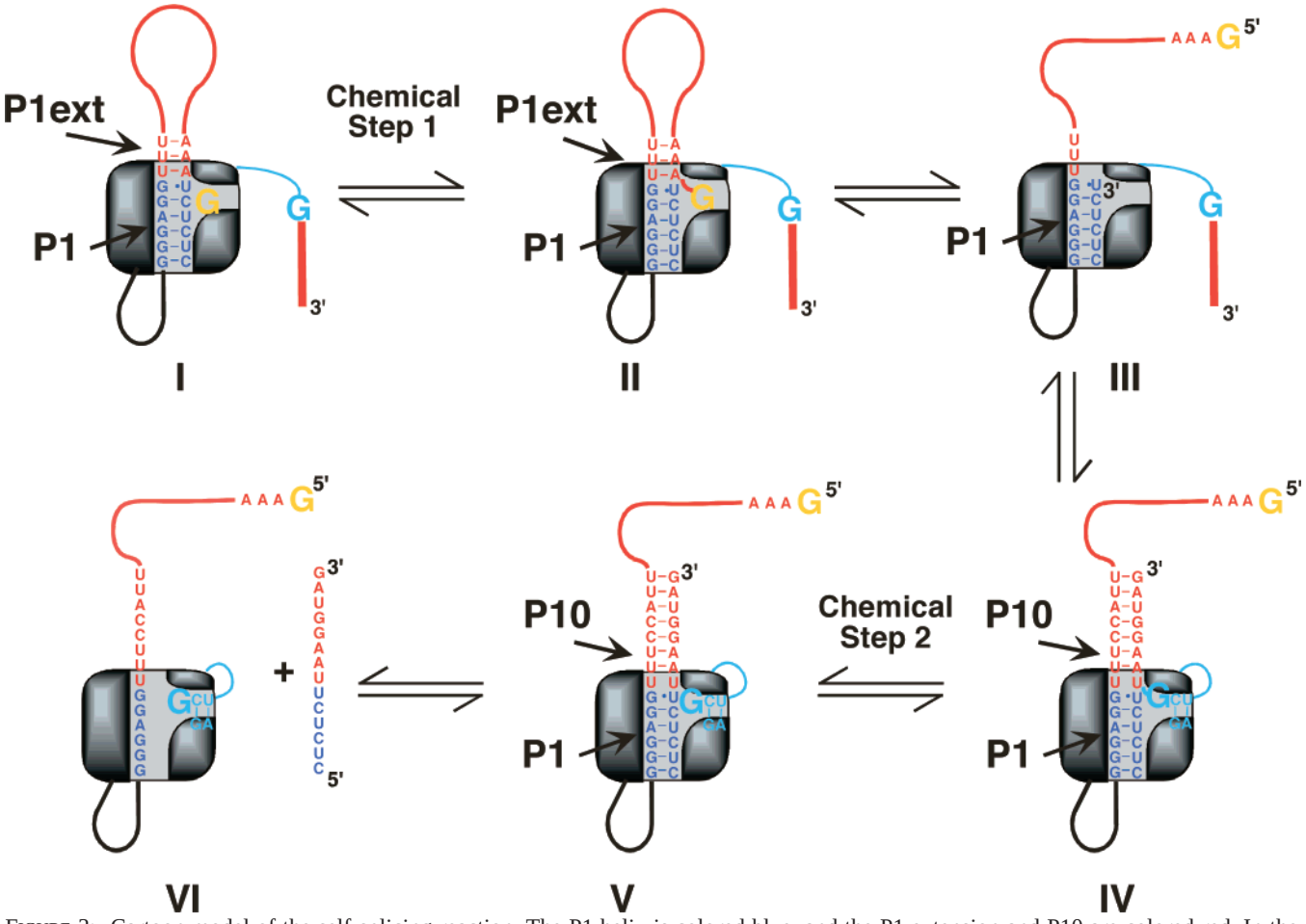


FIGURE 3: Cartoon model of the self-splicing reaction. The P1 helix is colored blue, and the P1 extension and P10 are colored red. In the first chemical step of self-splicing (transition from I to II), an exogenous G (colored orange) attacks at the 5′-splice site. Note that in II the extended P1 duplex is nicked. In a subsequent conformational step (transition from II to III), the G, now covalently linked to the P1 extension, leaves the G binding site. Next, G414, the last nucleotide of the intron (colored cyan), enters the G binding site, and the 3′-exon is aligned in P10 for ligation (III to IV). In the second chemical step (IV to V), exons are ligated to be released in a last step (V to VI).

ened binding of G (or UCG) and vice versa (*5*, *23*, *28*). This thermodynamic coupling of substrates is also observed for the L-16 *Sca*I ribozyme with CCUCdTA or CCUCdTAAACC bound (Figure 2C,D). The ~10-fold coupling effect with the L-16 *Sca*I ribozyme is similar to the 6-fold effect with the L-21 *Sca*I ribozyme (ref *5* and Figure 2B−D). Thus, neither the IGS extension nor the base pairs formed with the IGS extension to give the extended P1 duplex substantially affect the thermodynamic coupling of substrates.

Binding of the 5′-exon analogue (P) also strengthens binding of UCGA to the L-21 *Sca*I ribozyme ~5-fold (ref *5* and Figure 2B). In the presence of the extended IGS alone, weak binding of UCGA prevented determination of the extent of product coupling. When the P1 extension is formed, cooperative binding of products is maintained with a coupling effect of at least 6-fold (Figure 2D).

No Significant Contribution to 5′-Splice Site Analogue (S) Binding from Tertiary Interactions with the P1 Extension. A large body of previous work has shown that binding of S or P to the ribozyme occurs in two steps: formation of the P1 duplex and subsequent docking of this duplex into a binding site at the active site (*1*, *8*, *52*, *53*). In the "docked" state, specific tertiary interactions between the ribozyme and the P1 duplex contribute to binding of S and P (*1*, *7*, *54*, *72*−*74*). The strength of these tertiary interactions is measured as K_{tert}, the equilibrium constant for binding in tertiary interactions. To determine whether there were additional tertiary interactions from the P1 extension, we compared the K_{tert} values for the L-21 *Sca*I and L-16 *Sca*I ribozymes.

Single-molecule FRET measurements at 22 °C have shown that K_{tert} values for the L-21 *Sca*I and L-16 *Sca*I ribozymes (with all base pairs in the P1 extension formed) are similar with values of 22 ± 5 and 32 ± 4, respectively (*55*). These data are consistent with bulk measurements using modified substrate analogues at 30 °C (K_{tert} = 15 and 23 for binding to the L-21 *Sca*I and L-16 *Sca*I ribozyme, respectively; data not shown). The simplest interpretation of these data is that there are no significant interactions between the P1 extension and the ribozyme. The recently determined crystal structure of the *Azoarcus* intron suggests the formation of a single hydrogen bond between the 2′-OH group of the last residue in J4/5 (G116 in *Tetrahymena*) and the 2′-OH group of the first residue of the IGS extension (G21 in *Tetrahymena*) (*65*). A previous measurement of the effect of removing tertiary interactions involving 2′-OH groups (by modification to 2′-H) gave effects of ~0.5 kcal/substitution, corresponding to an effect of ~2-fold (*75*). Such a small effect is consistent with the similar tertiary stabilization observed for ribozymes with and without extended P1 duplexes.

Probing the Effect of a U·U Mismatch after the Cleavage Site

As described above, there is a change in base pairing between the first and second chemical steps in self-splicing (Figure 3 and Scheme 2) that originates from the replacement of the 5′-end of the intron with the 3′-exon. These segments differ in the identity of the +1 residue, A or U, resulting in formation of an A·U or U·U base pair in the first or second step, respectively. Previous work with the L-21 *Sca*I ribozyme has shown that a U(+1)-containing substrate weakened G binding (*76*). These results were surprising and could have arisen from fortuitous, nonproductive interactions of the ligated exon analogue that interfered with catalysis. Alternatively, it was possible that a previously unrecognized mechanism used to render self-splicing irreversible had been uncovered. To distinguish between these possibilities, we further investigated the effect from the U(+1) substitution using the L-16 *Sca*I ribozyme.

Comparison of the contributions of tertiary interactions toward binding of CCUCdTUAACC and CCUCdTAAACC shows that in the absence of G the U(+1)-containing substrate docks more strongly into the active site than the A(+1)-containing substrate (K_{tert} = 230 and 23 for CCUCdTUAACC and CCUCdTAAACC, respectively; Table 4), providing evidence for tertiary interactions with U(+1). Further, the tertiary binding energy of CCUCdTUAACC is 4.0 kcal/mol in the absence of bound G, 1 kcal/mol more than that observed in the presence of bound G (Table 4). The stronger binding energy without G present suggests that G binding disrupts the tertiary interactions with CCUCdTUAACC.

Comparison of the rate of reaction from the L-16 *Sca*I E·S·G complex with U(+1)- and A(+1)-containing substrates shows that both reactions proceed with the same rate constant (k_c = 0.039 and 0.041 min^{-1} for the reaction of CCUCdTAAACC and CCUCdTUAACC, respectively; Table 4). This observation suggests that the presence of guanosine restores interactions that are important for catalysis, supporting the notion that the effect from U(+1) is fortuitous and nonproductive.

The previous results and the analysis of the +1U effect suggest that the interactions of the U(+1)-containing substrate with the L-21 *Sca*I are fortuitous, providing another example of RNA's conformational promiscuity (e.g., refs *30*, *55*, *77*, and *78*). The promiscuous interaction can be overcome by occupying other portions of the active site in their cognate interactions with G and by immersing the U·U pair within a duplex, effects that have implications for understanding the origins of tertiary structure specificity in RNA (*58*).

SUMMARY AND IMPLICATIONS

Summary of Effects from the IGS Extension and the P1 Extension. A stepwise approach to dissecting the contribution from individual elements within the group I intron to the self-splicing reaction was taken. Starting from the well-characterized L-21 *Sca*I ribozyme, which lacks 21 nucleotides at the 5′-terminus and five nucleotides at the 3′-terminus, we restored sequence elements at the 5′-end and determined their effect on individual steps of the ribozyme reaction. Figure 4 summarizes effects from the IGS extension

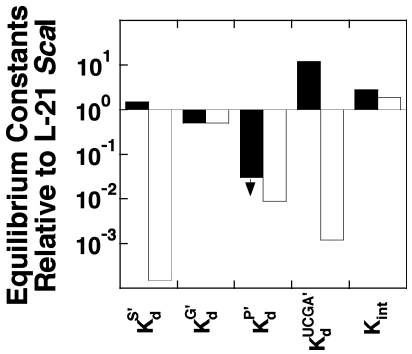


FIGURE 4: Summary of the effects on equilibrium constants from IGS extension (black columns) and the formation of the P1 extension (white columns) on the ribozyme reaction. Equilibrium constants are defined in Figure 2A. $K_d^{UCGA'}$ refers to binding of UCGA (black columns) or UCGAAACC (white columns) in the presence of bound P. Similarly, $K_d^{S'}$ refers to the binding of −1d,-rSA (black columns) or −1d,rSA$_3$C$_2$ (white columns) in the presence of bound UCG. All values are relative to the L-21 *Sca*I reaction and were calculated from Figure 2. Effects from the IGS extension were determined from comparison of reactions with the L-16 *Sca*I ribozyme that has CCUCdTA bound. The combined effects from the IGS extension and formation of base pairs with that extension were determined from reactions of the L-16 *Sca*I ribozyme with CCUCdTAAACC bound.

(dark columns) and from formation of the P1 extension (light columns). These effects are summarized briefly below.

The presence of the IGS extension (dark columns in Figure 4) strengthens binding of G and P and weakens binding of UCGA (second and third set of columns in Figure 4, respectively). Further addition of the base pairs with the IGS extension (white columns in Figure 4) affects only reaction steps in which the base pairs directly tether the ligand to the ribozyme (first and fourth set of columns in Figure 4). These data suggest that the IGS extension modulates kinetics and thermodynamics in self-splicing beyond its role in positioning the 3′-exon for ligation in the second step of self-splicing. P bound to the L-16 *Sca*I ribozyme 40-fold stronger than to the L-21 *Sca*I ribozyme (Figure 2B,C). This large effect from the IGS extension was especially surprising as binding of the 5′-splice site analogue (i.e., S = rSMe or rSA) lacking the ability to form the P1 extension was essentially unaffected.

Last, the effects from the IGS extension on the product side of the framework are larger than the effects on the substrate side, with 2−3-fold stronger binding of (UC)G and 2-fold weaker binding of S, compared to 40-fold stronger binding of P and 12-fold weaker binding of UCGA (Figure 2B,C). This asymmetry suggests that the communication between the ribozyme extension and the active site is different when S and UCG or P and UCGA are bound. The origin of these differences is not understood.

Role of the G·U Pair in Self-Splicing. The P1 extension is a dynamic secondary structure element. During the first step of self-splicing, three A·U pairs are formed in the P1 extension. These base pairs have to be broken to allow positioning of the 3′-splice site, which occurs through base pairing of the 3′-exon with the IGS extension to form the P10 helix [Figure 3 (*40*−*43*)]. A recent mutational analysis showed that a modest stabilization of the P1 extension, which is expected to slow dissociation of this duplex, results in splicing defects (*79*). It thus appears possible that disruption

of the three A·U pairs could be rate-limiting or nearly rate-limiting for the first step of self-splicing. These considerations lead to the following questions. How are the base pairs in the P1 extension preferentially broken while those in P1 remain intact to prevent dissociation of the 5′-exon before the second step of self-splicing? Also, how is rapid dissociation of the ligated exons, the final product, ensured?

As noted above, previously published structural data with model duplexes suggested that stacking is disrupted by an internal G·U wobble pair so that stacking between the P1 and the P1 extension helices might not be expected to be continuous (*66*−*68*, *70*). This structural picture is supported by the recently determined X-ray structures of the group I intron (*65*, *71*). Furthermore, this structural picture of internal G·U pairs is reflected in their energetics; substitution of a Watson−Crick pair with a G·U pair in a model duplex has a larger effect within a helix than at the end (see Formation of the Extended P1 Duplex) and an even larger effect of >2.8 kcal/mol measured on the ribozyme.

The discontinuity in stacking between P1 and the P1 extension presumably causes the observed loss in cooperativity of base pair formation at the junction of these helices such that a lone base pair beyond the G·U pair is not formed. In self-splicing, this interruption of cooperativity might facilitate release of the ligated exons after the second step of self-splicing. The discontinuity in stacking could also facilitate exchange of the P1 extension with the 3′-exon after the first chemical step in splicing (Figure 3, II−IV) while holding onto the 5′-exon in P1 for exon ligation. For exchange of the P1 extension with P10 to be facilitated by the G·U pair, the G·U pair is required to exert its destabilizing effect even in the case of a nicked duplex [as found after the first chemical step of self-splicing (Figure 3, II)]. While it is possible that the disruptive effect from the G·U wobble pair is dependent on a covalent linkage and would then disappear with a nick, the data in Table 3 show that the contribution from the P1 extension is independent of covalent linkage to P1, suggesting that the G·U pair also disrupts the noncovalently linked, nicked helix. This notion is also consistent with the crystal structures on the group I intron in the intermediate form (*65*, *71*). Thus, the G·U pair may serve a previously unappreciated role in self-splicing. Besides helping to define the position of the cleavage site through tertiary interactions (*2*, *73*), the G·U wobble pair uncouples the stability of the P1 and P1 extension helix, allowing coordination of conformational transitions needed to accomplish the two transesterification steps in self-splicing.

Proline is an amino acid that disrupts α-helices or β-sheets in proteins. This disruption arises because proline does not have a free backbone amide for hydrogen bonding with carbonyl groups within helices or sheets. We suggest that G·U wobble pairs in RNA molecules can act analogously. Internal G·U wobble pairs weaken the stacking interactions between two helical segments, thereby reducing or removing the cooperativity between these segments so that they become quasi-independent, rigid modules.

SUPPORTING INFORMATION AVAILABLE

Plots showing that the lack of a 5′-terminal triphosphate group is not responsible for the stronger binding of G to the L-16 *Sca*I ribozyme (Figure S1) and dissociation rate constants of different substrates at 30 and 50 °C (Table S1). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Herschlag, D., and Cech, T. R. (1990) Catalysis of RNA Cleavage By the *Tetrahymena thermophila* Ribozyme. 1. Kinetic Description of the Reaction of an RNA Substrate Complementary to the Active Site, *Biochemistry 29*, 10159−10171.
2. Knitt, D. S., Narlikar, G. J., and Herschlag, D. (1994) Dissection of the Role of the Conserved G·U Pair in Group I RNA Self-Splicing, *Biochemistry 33*, 13864−13879.
3. Narlikar, G. J., and Herschlag, D. (1998) Direct Demonstration of the Catalytic Role of Binding Interactions in an Enzymatic Reaction, *Biochemistry 37*, 9902−9911.
4. Shan, S., Yoshida, A., Sun, S. G., Piccirilli, J. A., and Herschlag, D. (1999) Three metal ions at the active site of the *Tetrahymena* group I ribozyme, *Proc. Natl. Acad. Sci. U.S.A. 96*, 12299−12304.
5. Karbstein, K., Carroll, K. S., and Herschlag, D. (2002) Probing the *Tetrahymena* Group I Reaction in Both Directions, *Biochemistry 41*, 11171−11183.
6. Herschlag, D., and Cech, T. R. (1990) Catalysis of RNA Cleavage by the *Tetrahymena thermophila* Ribozyme. 2. Kinetic Description of the Reaction of an RNA Substrate That Forms a Mismatch at the Active Site, *Biochemistry 29*, 10172−10180.
7. Narlikar, G. J., Khosla, M., Usman, N., and Herschlag, D. (1997) Quantitating tertiary binding energies of 2′ OH groups on the P1 duplex of the *Tetrahymena* ribozyme: Intrinsic binding energy in an RNA enzyme, *Biochemistry 36*, 2465−2477.
8. Bevilacqua, P. C., Kierzek, R., Johnson, K. A., and Turner, D. H. (1992) Dynamics of Ribozyme Binding of Substrate Revealed by Fluorescence-Detected Stopped-Flow Methods, *Science 258*, 1355−1357.
9. Serra, M. J., and Turner, D. H. (1995) Predicting Thermodynamic Properties of RNA, *Methods Enzymol. 259*, 242−261.
10. Kierzek, R., Burkard, M. E., and Turner, D. H. (1999) Thermodynamics of single mismatches in RNA duplexes, *Biochemistry 38*, 14214−14223.
11. Bevilacqua, J. M., and Bevilacqua, P. C. (1998) Thermodynamic analysis of an RNA combinatorial library contained in a short hairpin, *Biochemistry 37*, 15877−15884.
12. Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982) Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*, *Cell 31*, 147−157.
13. Michel, F., and Westhof, E. (1990) Modeling of the 3-Dimensional Architecture of Group I Catalytic Introns Based on Comparative Sequence Analysis, *J. Mol. Biol. 216*, 585−610.
14. Celander, D. W., and Cech, T. R. (1991) Visualizing the Higher-Order Folding of a Catalytic RNA Molecule, *Science 251*, 401−407.
15. Wang, J. F., and Cech, T. R. (1992) Tertiary Structure around the Guanosine-Binding Site of the *Tetrahymena* Ribozyme, *Science 256*, 526−529.
16. Lehnert, V., Jaeger, L., Michel, F., and Westhof, E. (1996) New Loop-Loop Tertiary Interactions in Self-Splicing Introns of Subgroup Ic and Id: A Complete 3D Model of the *Tetrahymena thermophila* Ribozyme, *Chem. Biol. 3*, 993−1009.
17. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K. H., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) Crystal Structure of a Group I Ribozyme Domain: Principles of RNA Packing, *Science 273*, 1678−1685.
18. Golden, B. L., Gooding, A. R., Podell, E. R., and Cech, T. R. (1998) A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme, *Science 282*, 259−264.
19. Zaug, A. J., Grabowski, P. J., and Cech, T. R. (1983) Autocatalytic Cyclization of an Excised Intervening Sequence RNA Is a Cleavage Ligation Reaction, *Nature 301*, 578−583.
20. Zaug, A. J., and Cech, T. R. (1986) The Intervening Sequence RNA of *Tetrahymena* Is an Enzyme, *Science 231*, 470−475.
21. Pyle, A. M., McSwiggen, J. A., and Cech, T. R. (1990) Direct Measurement of Oligonucleotide Substrate Binding to Wild-Type and Mutant Ribozymes from *Tetrahymena*, *Proc. Natl. Acad. Sci. U.S.A. 87*, 8187−8191.

22. Herschlag, D., Eckstein, F., and Cech, T. R. (1993) Contributions of 2′-Hydroxyl Groups of the RNA Substrate to Binding and Catalysis by the *Tetrahymena* Ribozyme: An Energetic Picture of an Active Site Composed of RNA, *Biochemistry 32*, 8299−8311.
23. McConnell, T. S., Cech, T. R., and Herschlag, D. (1993) Guanosine Binding to the *Tetrahymena* Ribozyme: Thermodynamic Coupling with Oligonucleotide Binding, *Proc. Natl. Acad. Sci. U.S.A. 90*, 8362−8366.
24. Bevilacqua, P. C., Johnson, K. A., and Turner, D. H. (1993) Cooperative and Anticooperative Binding to a Ribozyme, *Proc. Natl. Acad. Sci. U.S.A. 90*, 8357−8361.
25. Narlikar, G. J., Gopalakrishnan, V., McConnell, T. S., Usman, N., and Herschlag, D. (1995) Use of Binding-Energy by an RNA Enzyme for Catalysis by Positioning and Substrate Destabilization, *Proc. Natl. Acad. Sci. U.S.A. 92*, 3668−3672.
26. Shan, S., and Herschlag, D. (1999) Probing the role of metal ions in RNA catalysis: Kinetic and thermodynamic characterization of a metal ion interaction with the 2′-moiety of the guanosine nucleophile in the *Tetrahymena* group I ribozyme, *Biochemistry 38*, 10958−10975.
27. Shan, S., Kravchuk, A. V., Piccirilli, J. A., and Herschlag, D. (2001) Defining the catalytic metal ion interactions in the *Tetrahymena* ribozyme reaction, *Biochemistry 40*, 5161−5171.
28. Shan, S., and Herschlag, D. (2002) Dissection of a metal-ion-mediated conformational change in *Tetrahymena* ribozyme catalysis, *RNA 8*, 861−872.
29. Szewczak, A. A., Ortolevadonnelly, L., Ryder, S. P., Moncoeur, E., and Strobel, S. A. (1998) A Minor-Groove RNA Triple Helix within the Catalytic Core of a Group I Intron, *Nat. Struct. Biol. 5*, 1037−1042.
30. Karbstein, K., and Herschlag, D. (2003) Extraordinarily Slow Binding of Guanosine to the *Tetrahymena* Group I Ribozyme: Implications for RNA Structure and Preorganization, *Proc. Natl. Acad. Sci. U.S.A. 100*, 2300−2305.
31. Zarrinkar, P. P., and Williamson, J. R. (1994) Kinetic intermediates in RNA folding, *Science 265*, 918−924.
32. Zarrinkar, P. P., and Williamson, J. R. (1996) The P9.1−P9.2 peripheral extension helps guide folding of the *Tetrahymena* ribozyme, *Nucleic Acids Res. 24*, 854−858.
33. Zaug, A. J., McEvoy, M. M., and Cech, T. R. (1993) Self-Splicing of the Group I Intron From *Anabaena* Pre-Transfer-RNA: Requirement for Base Pairing of the Exons in the Anticodon Stem, *Biochemistry 32*, 7946−7953.
34. Golden, B. L., and Cech, T. R. (1996) Conformational Switches Involved in Orchestrating the Successive Steps of Group I RNA Splicing, *Biochemistry 35*, 3754−3763.
35. Narlikar, G. J., and Herschlag, D. (1996) Isolation of a Local Tertiary Folding Transition in the Context of a Globally Folded RNA, *Nat. Struct. Biol. 3*, 701−710.
36. Mei, R., and Herschlag, D. (1996) Mechanistic Investigations of a Ribozyme Derived From the *Tetrahymena* Group I Intron: Insights into Catalysis and the Second Step of Self-Splicing, *Biochemistry 35*, 5796−5809.
37. Zarrinkar, P. P., and Sullenger, B. A. (1998) Probing the interplay between the two steps of group I intron splicing: Competition of exogenous guanosine with omega G, *Biochemistry 37*, 18056−18063.
38. Zaug, A. J., Davilaaponte, J. A., and Cech, T. R. (1994) Catalysis of RNA Cleavage by a Ribozyme Derived from the Group I Intron of *Anabaena* Pre-tRNA(Leu), *Biochemistry 33*, 14935−14947.
39. Kuo, L. Y., Davidson, L. A., and Pico, S. (1999) Characterization of the *Azoarcus* ribozyme: Tight binding to guanosine and substrate by an unusually small group I ribozyme, *Biochim. Biophys. Acta 1489*, 281−292.
40. Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A., and Scazzocchio, C. (1982) Making ends meet: A model for RNA splicing in fungal mitochondria, *Nature 300*, 719−724.
41. Price, J. V., and Cech, T. R. (1988) Determinants of the 3′ splice site for self-splicing of the *Tetrahymena* pre-rRNA, *Genes Dev. 2*, 1439−1447.
42. Michel, F., Hanna, M., Green, R., Bartel, D. P., and Szostak, J. W. (1989) The Guanosine Binding Site of the *Tetrahymena* Ribozyme, *Nature 342*, 391−395.
43. Suh, E. R., and Waring, R. B. (1990) Base pairing between the 3′ exon and an internal guide sequence increases 3′ splice site specificity in the *Tetrahymena* self-splicing rRNA intron, *Mol. Cell. Biol. 10*, 2960−2965.
44. Been, M. D., and Cech, T. R. (1985) Sites of circularization of the *Tetrahymena* rRNA IVS are determined by sequence and influenced by position and secondary structure, *Nucleic Acids Res. 13*, 8389−8408.
45. Zaug, A. J., Kent, J. R., and Cech, T. R. (1984) A Labile Phosphodiester Bond at the Ligation Junction in a Circular Intervening Sequence RNA, *Science 224*, 574−578.
46. Zaug, A. J., Kent, J. R., and Cech, T. R. (1985) Reactions of the Intervening Sequence of the *Tetrahymena* Ribosomal Ribonucleic Acid Precursor: pH Dependence of Cyclization and Site-Specific Hydrolysis, *Biochemistry 24*, 6211−6218.
47. Russell, R., and Herschlag, D. (1999) Specificity from steric restrictions in the guanosine binding pocket of a group I ribozyme, *RNA 5*, 158−166.
48. Zaug, A. J., Grosshans, C. A., and Cech, T. R. (1988) Sequence-Specific Endoribonuclease Activity of the *Tetrahymena* Ribozyme: Enhanced Cleavage of Certain Oligonucleotide Substrates That Form Mismatched Ribozyme Substrate Complexes, *Biochemistry 27*, 8924−8931.
49. Russell, R., and Herschlag, D. (1999) New pathways in folding of the *Tetrahymena* group I RNA enzyme, *J. Mol. Biol. 291*, 1155−1167.
50. Narlikar, G. J., Bartley, L. E., Khosla, M., and Herschlag, D. (1999) Characterization of a local folding event of the *Tetrahymena* group I ribozyme: Effects of oligonucleotide substrate length pH and temperature on the two substrate binding steps, *Biochemistry 38*, 14192−14204.
51. Fersht, A. (1984) *Enzyme Structure and Mechanism*, W. H. Freeman and Co., New York.
52. Herschlag, D. (1992) Evidence For Processivity and Two-Step Binding of the RNA Substrate from Studies of J1/2 Mutants of the *Tetrahymena* Ribozyme, *Biochemistry 31*, 1386−1399.
53. Pyle, A. M., and Cech, T. R. (1991) Ribozyme Recognition of RNA by Tertiary Interactions with Specific Ribose 2′-OH Groups, *Nature 350*, 628−631.
54. Strobel, S. A., and Cech, T. R. (1993) Tertiary Interactions with the Internal Guide Sequence Mediate Docking of the P1 Helix into the Catalytic Core of the *Tetrahymena* Ribozyme, *Biochemistry 32*, 13593−13604.
55. Bartley, L. E., Zhuang, X., Das, R., Chu, S., and Herschlag, D. (2003) Exploration of the transition state for tertiary structure formation between an RNA helix and a large structured RNA, *J. Mol. Biol. 328*, 1011−1026.
56. Shan, S., and Herschlag, D. (2000) An unconventional origin of metal-ion rescue and inhibition in the *Tetrahymena* group I ribozyme reaction, *RNA 6*, 795−813.
57. Herschlag, D., Eckstein, F., and Cech, T. R. (1993) The Importance of Ribose Being at the Cleavage Site in the *Tetrahymena* Ribozyme Reaction, *Biochemistry 32*, 8312−8321.
58. Narlikar, G. J., Bartley, L. E., and Herschlag, D. (2000) Use of duplex rigidity for stability and specificity in RNA tertiary structure, *Biochemistry 39*, 6183−6189.
59. Herschlag, D., and Khosla, M. (1994) Comparison of pH Dependencies of the *Tetrahymena* Ribozyme Reactions with RNA 2′-Substituted and Phosphorothioate Substrates Reveals a Rate-Limiting Conformational Step, *Biochemistry 33*, 5291−5297.
60. Knitt, D. S., and Herschlag, D. (1996) pH Dependencies of the *Tetrahymena* Ribozyme Reveal an Unconventional Origin of an Apparent pK_a, *Biochemistry 35*, 1560−1570.
61. McDowell, J. A., and Turner, D. H. (1996) Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: Solution structure of $(rGAGGUCUC)_2$ by two-dimensional NMR and simulated annealing, *Biochemistry 35*, 14077−14089.
62. Zaug, A. J., Been, M. D., and Cech, T. R. (1986) The *Tetrahymena* Ribozyme Acts Like an RNA Restriction Endonuclease, *Nature 324*, 429−433.
63. Piccirilli, J. A., Vyle, J. S., Caruthers, M. H., and Cech, T. R. (1993) Metal Ion Catalysis in the *Tetrahymena* Ribozyme Reaction, *Nature 361*, 85−88.
64. Moran, S., Kierzek, R., and Turner, D. H. (1993) Binding of Guanosine and 3′ Splice Site Analogs to a Group I Ribozyme: Interactions with Functional Groups of Guanosine and with Additional Nucleotides, *Biochemistry 32*, 5247−5256.
65. Adams, P. L., Stahley, M. R., Kosek, A. B., Wang, J., and Strobel, S. A. (2004) Crystal structure of a self-splicing group I intron with both exons, *Nature 430*, 45−50.

66. Allain, F. H., and Varani, G. (1995) Structure of the P1 helix from group I self-splicing introns, *J. Mol. Biol. 250*, 333−353.
67. Baeyens, K. J., De Bondt, H. L., and Holbrook, S. R. (1995) Structure of an RNA double helix including uracil-uracil base pairs in an internal loop, *Nat. Struct. Biol. 2*, 56−62.
68. Masquida, B., Sauter, C., and Westhof, E. (1999) A sulfate pocket formed by three GoU pairs in the 0.97 Å resolution X-ray structure of a nonameric RNA, *RNA 5*, 1384−1395.
69. Shi, K., Wahl, M., and Sundaralingam, M. (1999) Crystal structure of an RNA duplex r(G GCGC CC)$_2$ with non-adjacent G*U base pairs, *Nucleic Acids Res. 27*, 2196−2201.
70. Masquida, B., and Westhof, E. (2000) On the wobble G·U and related pairs, *RNA 6*, 9−15.
71. Stahley, M. R., and Strobel, S. A. (2005) Structural evidence for a two-metal-ion mechanism of group I intron splicing, *Science 309*, 1587−1590.
72. Strobel, S. A., and Cech, T. R. (1994) Translocation of an RNA Duplex on a Ribozyme, *Nat. Struct. Biol. 1*, 13−17.
73. Strobel, S. A., and Cech, T. R. (1995) Minor-Groove Recognition of the Conserved G·U Pair at the *Tetrahymena* Ribozyme Reaction Site, *Science 267*, 675−679.
74. Strobel, S. A., and Cech, T. R. (1996) Exocyclic Amine of the Conserved G·U Pair at the Cleavage Site of the *Tetrahymena* Ribozyme Contributes to 5′-Splice Site Selection and Transition-State Stabilization, *Biochemistry 35*, 1201−1211.
75. Silverman, S. K., and Cech, T. R. (1999) Energetics and cooperativity of tertiary hydrogen bonds in RNA structure, *Biochemistry 38*, 8691−8702.
76. Liao, X. M., Anjaneyulu, P. S. R., Curley, J. F., Hsu, M., Boehringer, M., Caruthers, M. H., and Piccirilli, J. A. (2001) The *Tetrahymena* ribozyme cleaves a 5′-methylene phosphonate monoester similar to 10^2-fold faster than a normal phosphate diester: Implications for enzyme catalysis of phosphoryl transfer reactions, *Biochemistry 40*, 10911−10926.
77. Walstrum, S. A., and Uhlenbeck, O. C. (1990) The self-splicing RNA of *Tetrahymena* is trapped in a less active conformation by gel purification, *Biochemistry 29*, 10573−10576.
78. Shan, S., Narlikar, G. J., and Herschlag, D. (1999) Protonated 2′-aminoguanosine as a probe of the electrostatic environment of the active site of the *Tetrahymena* group I ribozyme, *Biochemistry 38*, 10976−10988.
79. Guo, F., and Cech, T. R. (2002) In vivo selection of better self-splicing introns in *Escherichia coli*: The role of the P1 extension helix of the *Tetrahymen*a intron, *RNA 8*, 647−658.

BI062169G