Dissection of the Role of the Conserved G·U Pair in Group I RNA Self-Splicing†

Deborah S. Knitt,‡ Geeta J. Narlikar,§ and Daniel Herschlag*,‡,§

Department of Biochemistry, B400 Beckman Center, and Department of Chemistry, Stanford University, Stanford, California 94305

Received July 12, 1994; Revised Manuscript Received August 31, 1994⊗

ABSTRACT: Phylogenetic conservation among >100 group I introns and previous *in vitro* studies have implicated a G·U pair as defining the 5′-splice site for exon ligation. The U residue defines the 3′ end of the 5′ exon, and the complementary G residue is part of the internal guide sequence (IGS) that base pairs to the 5′ exon. We now quantitate the effect of this pair on individual reaction steps using the L-21ScaI ribozyme, which is derived from the group I intron of *Tetrahymena thermophila* pre-rRNA. The following results indicate that interactions with this G·U pair contribute to the binding of the 5′-exon, the positioning of the 5′-splice site with respect to the catalytic site, and the chemical step. The oligonucleotide, CCCUCU, binds to the ribozyme ~20-fold stronger than CCCUCC despite the fact that the U-containing oligonucleotide forms an ~5-fold less stable duplex with an oligonucleotide analog of the IGS, GGAGGG. This and two independent experimental observations indicate that the G·U pair contributes ~100-fold (3 kcal/mol, 50 °C) to tertiary interactions that allow the P1 duplex, which is formed between the 5′-exon and the IGS, to dock into the ribozyme's core. The ~50−80-fold increase in miscleavage of 5′-exon analogs upon replacement of the 3′-terminal U of CCCUCU with C or upon removal of the 3′-terminal U suggests that the tertiary interactions with the G·U pair not only contribute to docking but also ensure correct positioning of the 5′-splice site with respect to the catalytic site, thereby minimizing the selection of incorrect splice sites. Comparison of the rates of the chemical cleavage step with G·U vs G·C suggests that the G·U pair contributes ~10-fold to the chemical step. It was previously suggested that the 2′-hydroxyl of this U residue helps stabilize the 3′-oxyanion leaving group in the chemical transition state via an intramolecular hydrogen bond. Relative reactivities of oligonucleotide substrates with ribose and deoxyribose U and C are consistent with a model based on a recent X-ray crystallographic structure in which the exocyclic amino group of G helps orient the 2′-hydroxyl of U via a bridging water molecule, thereby strengthening the hydrogen bond donated from the 2′-hydroxyl group to the neighboring incipient 3′-oxyanion. Finally, kinetic and thermodynamic evidence for the formation of a G·C$^+$ wobble pair is presented. The tertiary energy of recognition of the G·U wobble pair appears to be sufficient to perturb the pK_a of C to favor a G·C$^+$ wobble pair instead of the Watson−Crick pair, despite the loss of a hydrogen bond in the base pair. This provides an example of RNA tertiary structure determining secondary structure.

Group I introns contain a conserved G·U pair, with the U residue defining the 3′-end of the 5′-splice site and the G residue being part of the internal guide sequence (IGS)[1] that base pairs to the 5′-exon (Michel & Westhof, 1990; Figure 1A). Previous mutational data supported a role for this G·U pair in splice site selection. Barfod and Cech (1989) observed a decrease in the fidelity of 5′-splice site selection in self-splicing of the *Tetrahymena* pre-rRNA intron upon replacement of this U residue with a C. Doudna et al. (1989) used a multipartite ribozyme and supplied a series of helices analogous to P1 (Figure 1) *in trans* in which the position of the G·U pair was varied throughout the helix; cleavage was strongly favored at the phosphodiester bond following the U residue.

Because G and U residues can form a stable wobble pair, it had commonly been assumed that such a wobble pair is formed in self-splicing. Evidence for this wobble pair was subsequently obtained by the observations that C·A, which can also form a wobble pair upon protonation of A, could partially substitute for G·U (Doudna et al., 1989) and that inosine, which can form the same hydrogen bonds as G in a wobble pair with U, could efficiently replace G in a ribozyme reaction (Green et al., 1991).

Protein enzymes can also distinguish between G·U wobble pairs and Watson−Crick pairs. For example, a G·U pair in the acceptor stem of tRNA^{Ala} serves as a major recognition element for *Escherichia coli* tRNA^{Ala} aminoacyl synthetase (Hou & Schimmel, 1988). In addition, a DNA glycosidase involved in mismatch repair is capable of cleaving uracil or thymine from G·U(T) pairs, but not from A·U(T) Watson−Crick pairs (Neddermann & Jiricny, 1994).

In this investigation, we have taken advantage of the previous kinetic and thermodynamic characterization of a ribozyme derived from the self-splicing *Tetrahymena thermophila* intron, the L-21ScaI ribozyme (Figure 1B), to ask the following question: How does the G·U pair contribute to accurate 5′-splice site selection? We have demonstrated and quantitated the contributions of the G·U pair in the formation of the P1 duplex, the docking of the P1 duplex into tertiary interactions, the proper positioning of the P1 duplex, and the chemical step. These results also lead to a

† This work was supported by NIH Grant GM49243 to D.H. D.H. is a Lucille P. Markey Scholar in Biomedical Sciences and a Searle Scholar (Chicago Community Trust). D.S.K. was supported by an NSF predoctoral fellowship.

* Author to whom correspondence should be addressed at the Department of Biochemistry.

‡ Department of Biochemistry.

§ Department of Chemistry.

⊗ Abstract published in *Advance ACS Abstracts,* October 15, 1994.

0006-2960/94/0433-13864$04.50/0

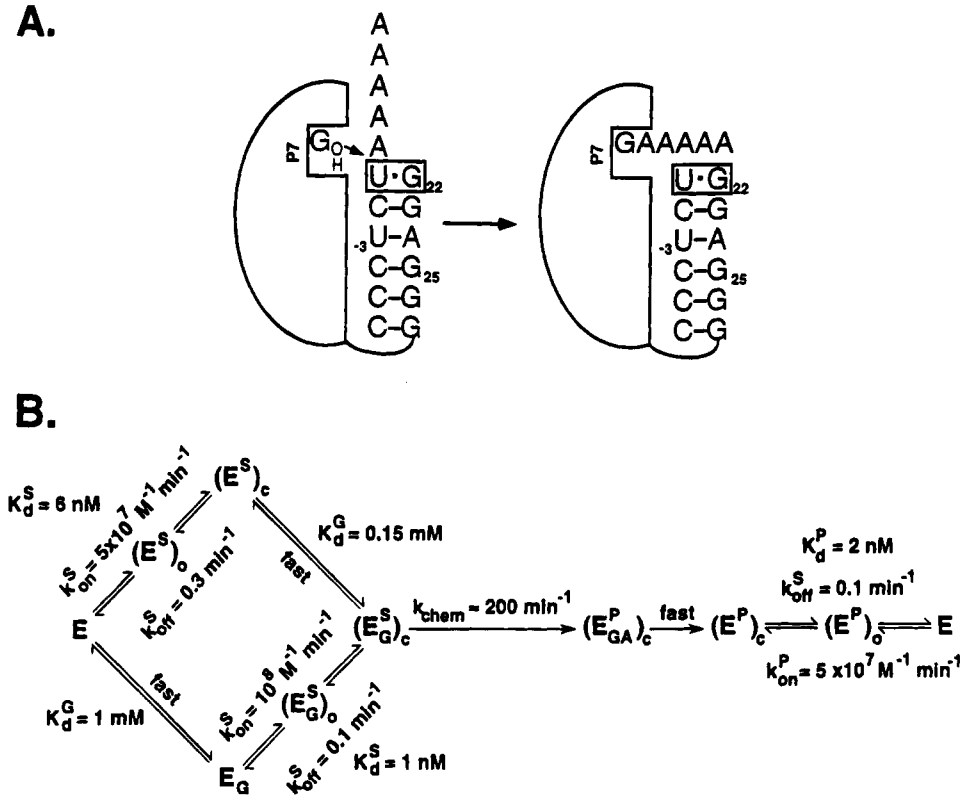


FIGURE 1: *Tetrahymena* ribozyme and its reaction. (A) The ribozyme endonuclease reaction. The internal guide sequence (IGS) base pairs with the oligonucleotide substrate to form the P1 duplex. G binds at region P7 and attacks following the conserved U residue (Michel et al., 1989; Cech, 1990). (B) Kinetic and thermodynamic framework for the L-21ScaI ribozyme reaction [pH 7, 10 mM $MgCl_2$, 50 °C; adapted from Herschlag and Khosla (1994) and references therein]. E is the ribozyme, S is the oligonucleotide substrate CCCUCUAAAAA, G is guanosine, and P is the oligonucleotide product CCCUCU. The subscript o refers to the open complex, in which the oligonucleotide is held in place solely by base-pairing interactions with the IGS, and the subscript c refers to the closed complex, in which the P1 duplex between the oligonucleotide and the 5′-exon docks into tertiary interactions with the ribozyme's core (Bevilacqua et al., 1992; Herschlag 1992; G. J. Narlikar and D. Herschlag, manuscript in preparation). These two binding steps are described further in Figure 2 and in the text. A conformational step that is associated with either G binding or the ternary complex and a deprotonation step of the ternary complex that occurs prior to the chemical step are omitted for clarity (Herschlag & Khosla, 1994; D. S. Knitt and D. Herschlag, manuscript in preparation).

model in which a directed hydrogen bond contributes specifically in the catalytic step. Finally, evidence for formation of a G·C$^+$ wobble pair demonstrates that the ribozyme's tertiary structure can alter secondary structure.

MATERIALS AND METHODS

Materials. Ribozyme was prepared by *in vitro* transcription and purified as described previously (Zaug et al., 1988). Oligonucleotides were made by solid phase synthesis and were supplied by Clontech (Palo Alto, CA) and by the Protein and Nucleic Acid Facility at Stanford University or were used and characterized in previous studies. Oligonucleotides were purified and characterized subsequent to 5′-end-labeling as described previously (Herschlag et al., 1993a). Oligonucleotides were 5′-end-labeled with approximately equimolar amounts of [γ-^{32}P]ATP using T4 polynucleotide kinase and purified by nondenaturing polyacrylamide gel electrophoresis, as described previously (Zaug et al., 1988; Herschlag et al., 1993a). All reaction products were determined by comparison of electrophoretic mobility vs appropriate markers from previously characterized ribozyme reactions or partial P1 nuclease digestions. The IGS analogs IGS′ (GGAGGG) and IGS′A (GGAGGGA) were kindly supplied by N. Usman of Ribozyme Pharmaceuticals, Inc., and were HPLC purified prior to use.

General Kinetic Methods. All reactions were single-turnover, with ribozyme (E) in excess of 5′-end-labeled S (~0.01−0.5 nM, depending on the [E] used) and with 0−4.3 mM G, and were carried out at either 50 or 30 °C in 50 mM buffer and 10 mM $MgCl_2$, as described previously (Herschlag

[1] Abbreviations: S is used generically to refer to an oligonucleotide substrate, CCCUCUAAAAA or CCCUCCAAAAA, without specification of the sugar identity; P is used similarly to refer to the oligonucleotide product, CCCUCU or CCCUCC, also without specification of the sugar identity. SU and SC refer to the substrates CCCUCUAAAAA and CCCUCCAAAAA, respectively. To specify the identity of sugar residues, the following nomenclature is employed: rS and dS refer to all-ribose and all-deoxyribose residues, respectively. Chimeric oligonucleotides are named in accord with the residues that are different from the ribose background of rS or the deoxyribose background of dS. For example, −1r,dSU signifies a ribose residue at position −1 (eq 1) in the deoxyribose background, i.e., with deoxyribose residues at all other positions, dCdCdCdUdCrUdAdAdAdAdA; likewise, −1d,rSU signifies a deoxyribose residue at position −1 in the ribose background, rCrCrCrUrCdUrArArArArA. A 2′-methoxy-2′-deoxyribose residue is signified with m. The shorthand notation rSU′ and rSC′ is used to denote substrates with 2′-methoxy-2′-deoxyribose residues at positions −6, −5, and −4 (e.g., mCmCmCrUrCrU-rArArArArA). This shorthand is adopted because these substrates have reactivities identical to those of rSU and rSC at the normal cleavage site, although they shut down the miscleavage reaction (see text). The following additional abbreviations are used: E, L-21ScaI ribozyme; P1, helix formed between the internal guide sequence (IGS) of the ribozyme and the oligonucleotide substrate or product, as depicted in Figure 1A; IGS′, all-ribose oligonucleotide GGAGGG; IGS′A, all-ribose oligonucleotide GGAGGGA; P′, CCCUC; P″, CCCU; MES, 2-morpholinoethanesulfonic acid; EDTA, ethylenedinitrilotetraacetic acid; EPPS, *N*-2-(hydroxyethyl)piperazine-*N*′-3-propanesulfonic acid; HEPES, *N*-2-(hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; CHES, (cyclohexylamino)ethanesulfonic acid.

& Cech, 1990a; Herschlag et al., 1993a). Briefly, reactions were initiated by the addition of labeled substrate following a 15 min/50 °C preincubation of E in $MgCl_2$ and buffer for reactions carried out at 50 °C or a 30 min/50 °C preincubation followed by a >5 min incubation at 30 °C for reactions carried out at 30 °C. At least six aliquots of 1−2 μL were removed from 20 μL reactions at specified times, and further reaction was quenched by the addition of ~2 vol of 20 mM EDTA in 90% formamide with 0.005% xylene cyanol, 0.01% bromophenol blue, and 1 mM Tris (pH 7.5). Substrates and product(s) were separated by electrophoresis on 20% polyacrylamide/7 M urea gels, and their ratio at each time point was quantitated with a Molecular Dynamics PhosphorImager. Reactions were typically followed for ~$3t_{1/2}$, except for the slowest reactions, which were followed for times such that there was no indication of the inactivation of E from curvature in reaction progress curves or from loss of activity in control reactions in which E was incubated under reaction conditions for varying times prior to the initiation of a $(k_{cat}/K_m)^S$ reaction that is first-order in [E]. End points of ~95% typically were obtained and were used in nonlinear first-order fits to the data (KaleidaGraph, Synergy Software, Reading, PA). If only initial rates could be measured, an end point of 95% was assumed. There were two exceptions with different end points. The first is the end point of ~50% for the phosphorothioate substrates. These substrates have sulfur replacing a nonbridging phosphoryl oxygen atom at the cleavage site. They were used as a mixture of the R_P and S_P isomers, but only the R_P substrate reacts on the time scale of these experiments (Rajagopal et al., 1989; Herschlag et al., 1991; J. A. Piccirilli and T. R. Cech, personal communication). The second exception is that two of the C-containing substrates, rSC and rSC′, had end points ranging from 5% to 20%, apparently due to minor side reactions that were especially prevalent at low [G] (data not shown); the small effects on the observed rate constants from this end point variation do not affect any of the conclusions herein. Details of specific kinetic experiments are described in the following sections.

The buffers used in these experiments, and their pH values at 50 °C, were as follows: sodium acetate, 4.6−5.1; NaMES, 5.2−6.8; NaHEPES, 7.2, 7.8; NaEPPS, 7.6, 7.9, 8.4; NaCHES, 8.8. Typically, the pH was measured at 25 °C and corrected to 50 °C using the following factors: sodium acetate, no correction; NaMES, −0.2; NaHEPES, −0.3; NaEPPS, −0.3; NaCHES, −0.6. These correction values were obtained by measuring the pH's at 25 and 50 °C for select pH values, and they agree reasonably well with literature values (Good et al., 1966; Gueffroy, 1978). The buffers and their pH values used at 30 °C were as follows: NaMES, pH 6.7; NaEPPS, pH 7.8; NaCHES, pH 9.3. pH corrections from 25 °C were −0.04, −0.06, and −0.12, respectively.

The following nomenclature is used for kinetic constants. The first-order rate constant with saturating ribozyme and saturating G is termed k_c, representing reaction of the E·S·G ternary complex. Pseudo-first-order reactions with saturating ribozyme and G concentrations below saturating are called k_c^{app} and hold for a specific G concentration. $(k_{cat}/K_m)^G$ represents the second-order rate constant for reaction of the E·S complex with G, and $(k_{cat}/K_m)^S$ represents the second-order rate constant for reaction of the E·G complex with S. The term K_m is used to define the concentration of ribozyme, G, or S that provides one-half of the maximal rate, even though the reactions are single-turnover and therefore not under steady state conditions. This nomenclature is used to emphasize the fact that this value does not *a priori* equal an equilibrium dissociation constant. For each of the saturation curves presented herein, there is strong independent evidence suggesting that K_m equals the true equilibrium dissociation constant (K_d), as will be summarized here. [See McConnell et al. (1993) for a more detailed description.]

Miscleavage of 5′-Exon Analogs: CCCUCU, CCCUCC, and CCCUC_. Miscleavage assays at pH 6.6 and 50 °C were carried out with subsaturating or near-saturating G (25 μM or 2 mM), ribozyme concentrations of 0−125 and 0−1000 nM for PU and PC, respectively, and ≤1 nM 5′-labeled PU or PC. P″ was the main product and P′ was a minor product for both oligonucleotides, as observed previously with CCCUCU (Herschlag, 1992). The following observations strongly suggest that the observed K_m is equal to the equilibrium dissociation constant, K_d, for PU and PC: (1) the values of k_{on}^P and k_{off}^P were measured independently for PU and PC by following the reverse reaction with GA using pulse chase methods as described previously (Herschlag & Cech, 1990a), and they give a value of K_d^P that is the same, within error, as the value of K_m obtained in this experiment (data not shown); (2) the maximal rate, k_c^{app}, is slower than k_{off}^P, suggesting that there is time to establish equilibrium binding prior to the cleavage reaction; (3) there is no lag observed as one might expect if binding were rate-limiting; and (4) the value of k_c^{app} for PU was log-linear with pH and that for PC was nearly log-linear over this same pH range (measured at pH 4.6−6.6), again suggesting that the chemical step rather than a binding step is rate-limiting [as discussed in Herschlag and Khosla (1994)]. (A small deviation from linearity for PC is discussed in Results.)

The values of k_c for CCCUCC and CCCUC_ (i.e., P′)[1] were compared at 30 °C because the binding of CCCUC_ was too weak at 50 °C to allow substantial formation of E·G·P′. The reactions were carried out in the presence of 2 mM G, and the ribozyme concentration was varied over 100−300 and 100−1000 nM for CCCUCC and CCCUC_, respectively, to ensure saturation. Comparison of CCCUCU and CCCUCC at 30 °C gave the same ratio of reactivities (k_c^U/k_c^C) that was obtained at 50 °C, indicating that the results were not significantly affected by the temperature of the comparison.

Reactions of −1r,dS and dS. Reactions of the U- and C-containing −1r,dS and dS substrates were carried out in 2 mM G at 30 °C because binding at 50 °C was too weak to allow saturation even at high ribozyme concentrations (2000−5000 nM). Reactions of −1r,dSU and −1r,dSC at pH 6.7 were followed with ribozyme concentrations of 0−5000 and 0−1000 nM, respectively. The following observations suggest that the observed values of K_m represent equilibrium dissociation constants: (1) there is no lag observed as one might expect if binding were rate-limiting; (2) the values of $(k_{cat}/K_m)^S$ are $<10^8$ M^{-1} min^{-1}, the expected rate for binding (Herschlag & Cech, 1990a; Herschlag et al., 1993a); and (3) there is an 11(±2)-fold increase in rate for both −1r,dS substrates with an increase in pH from 6.7 to 7.8 (with saturating ribozyme: 1000 and 200 nM for −1r,dSU and −1r,dSC, respectively); this suggests that the chemical step, and not a binding step, is rate-limiting [Herschlag & Khosla, 1994; D. S. Knitt and D. Herschlag, unpublished results; see also McConnell et al. (1993)]. The

reactions of −1r,dSA and −1r,dSG with 2 mM G and 0−2000 nM ribozyme at pH 6.7 were also investigated.

The reactions of dSU and dSC and 2 mM G at pH 6.7, 7.8, and 9.3 were followed with saturating ribozyme (500−1000 and 50−150 nM for dSU and dSC, respectively) to obtain values of k_c. The pH dependence revealed a log-linear increase with pH with a slope of ~1, again suggesting that k_c represents the chemical step [Herschlag & Khosla, 1994; see also McConnell et al. (1993)]. [Because the ribozyme is unstable during the 30 min preincubation at 50 °C and pH 9.3 (D. S. Knitt and D. Herschlag, manuscript in preparation), a pH jump experiment was performed in which the ribozyme was preincubated at pH 7 and 50 °C for 30 min and equilibrated to 30 °C for >5 min; the pH was subsequently raised prior to or concomitant with the addition of 5′-labeled S to initiate the reaction.]

Reactions of rS. The value of $(k_{cat}/K_m)^G$ was determined for rSU and rSC at 50 °C and pH 4.6−8.4 with the buffers indicated earlier. Ribozyme concentrations of 150−1000 nM were used, and controls indicated that the concentrations used at each pH were saturating with respect to the oligonucleotide substrate. Guanosine concentrations ranged from 0 to 25 μM, with 3−4 concentrations used for each $(k_{cat}/K_m)^G$ determination. The G concentrations were chosen to remain at least 10-fold below the dissociation constant of 250 μM for dissociation of G from the E·S·G complex (McConnell et al., 1993) and to give observed first-order rate constants of ≤2.5 min^{-1} in order to allow the accurate determination of rate constants by manual pipetting.

The main product from cleavage of rSC was CCCUCC at all pH values, but there were additional products due to miscleavage, accounting for no more than 15% of the total product. When the rates were determined, all products were included in the analysis. To discern whether these additional products were significantly affecting the rates, the cleavage of rSC′, which contains 2′-methoxy-2′-deoxyribose substituents at positions −4, −5, and −6, was followed. These substituents shut down miscleavage by preventing misdocking of the P1 helix [see Figure 2 and Herschlag (1992)]. The rates for the individual guanosine reactions were the same, within experimental error, for rSC and rSC′. The binding and reactivity of the U-containing substrate were also unaffected by substituting methoxy groups for the 2′-hydroxyl groups at positions −4, −5, and −6 (G. J. Narlikar and D. Herschlag, unpublished results).

The K_m(S) values for rSC and CCC(mU)CCAAAAA were measured by varying the ribozyme concentration in the presence of ~1 nM substrate and in the absence of guanosine at pH 6.6 and 50 °C. The reaction occurring in the absence of guanosine, referred to as hydrolysis (Herschlag & Cech, 1990a), was followed because the chemical step for hydrolysis of rSC is ≥10-fold slower than k_{off}, which is slow enough to allow the measured K_m to represent a K_d. [The value of k_{off} for rSC was determined in pulse chase experiments carried out as described previously (Herschlag & Cech, 1990a).] As the 2′-methoxy-2′-deoxyribose substitution at position −3 weakens binding in the closed complex [Herschlag et al. (1993a) and Results], and thereby slows k_c for CCC(mU)CCAAAAA (data not shown), the measured K_m for CCC(mU)CCAAAAA is also expected to reflect the actual equilibrium dissociation constant.

Reactions of −1d,rS. Reactions of −1d,rSU′ and −1d,-rSC′ were followed at 50 °C and pH 6.6 with saturating ribozyme (200−300 and 200−1000 nM E for −1d,rSU′ and −1d,rSC′, respectively). The substrates with 2′-methoxy-2′-deoxyribose substituents at positions −6, −5, and −4 again were used to prevent miscleavage reactions (see above). The rate constant for reaction of the ternary E·S·G complex, k_c, was also determined for these substrates at pH 4.6−7.6 in the presence of 2.8−4.3 mM G.

The values of k_c for cleavage of −1d,rSC′ at each pH were obtained by correcting values of k_{obsd} by 1.3- and 1.2-fold for values obtained with 2.8 and 4.3 mM G, respectively, to account for incomplete saturation of G. [The value of $K_d^{E·S·G}$ = 1.0 ± 0.2 mM for the E·(−1d,rSC′) complex used in this correction was obtained from reactions with 200 nM E, ~1 nM S, and G concentrations ranging from 0 to 4.3 mM at pH 6.6; this value and K_d^G for binding to free E were determined as described previously (McConnell et al., 1993). Note that binding of G is pH independent over this region (McConnell et al., 1993; D. S. Knitt and D. Herschlag, manuscript in preparation).] This correction is required for −1d,rSU′ only at pH <5.6, as the K_d for G dissociation from the E·(−1d,rSU′)·G complex is ~100 μM above this pH, but decreases to 1 mM below it (D. S. Knitt and D. Herschlag, manuscript in preparation).

Determination of Thio Effect for the rS and −1d,rS Reactions. A thio effect is defined as the ratio of the rate of reaction of the oxygen substrate relative to the sulfur substrate. $(k_{cat}/K_m)^G$ and k_c were measured for rS′ and −1d, rS′, respectively, for substrates with oxygen and sulfur at the *pro-*R_P oxygen position of the phosphate at the cleavage site. The only detectable reaction of the sulfur substrate comes from the *pro-*R_P substrate, which gives the R_P-thio isomer. The rSC′ substrate with $-OCH_3$ at positions −4 to −6 was used for a thio effect determination because miscleavage occurs at sites with nonbridging oxygens and will not allow a clean determination of the thio effect for cleavage solely at the bond of interest. Experiments to determine $(k_{cat}/K_m)^G$ for rSC′ and its R_P-thio isomer were carried out at pH 5.0, 6.6, and 7.6 using 300−1000 nM ribozyme to ensure complete formation of the E·S complex. For each $(k_{cat}/K_m)^G$ determination, 3−6 G concentrations were used ranging from 0 to 50 μM.

The thio effect on k_c was determined for −1drSC′ at pH values 5.0, 6.6, and 7.6. This experiment was carried out in quadruplicate for each k_c determination, along with a fifth experiment at a higher concentration of ribozyme to test for saturation. The enzyme concentrations were 1000−1500, 300−750, and 300−750 nM for pH values 5.0, 6.6, and 7.6, respectively. The rate of reaction did not change over this range of enzyme concentrations, suggesting that reaction of the E·S complex was followed.

Although variations of up to ~2-fold in individual rate constants have been observed in independent experiments (Herschlag & Cech, 1990a), the variation in side-by-side experiments carried out in multiplicate is much less, allowing the observed thio effects of 1.5−2.3 to be distinguished from a thio effect of 1.0.

Measurement of Duplex Stabilities by Substrate Inhibition. Reactions were carried out at 50 °C/pH 6.6 or 30 °C/pH 6.7with 2 mM G, 2−10 nM E, and 0.2−1 nM 5′-end-labeled oligonucleotide. After preincubation of ribozyme and guanosine in buffer and $MgCl_2$ at 50 °C, 2 μL of the appropriate 10× concentration of IGS′ (GGAGGG) or IGS′A (GGAGGGA) was added. Reactions were initiated by the addition of 2 μL of 5′-labeled S to give a final reaction volume of 20 μL. Duplex stabilities for CCCUCU and CCCUCC were

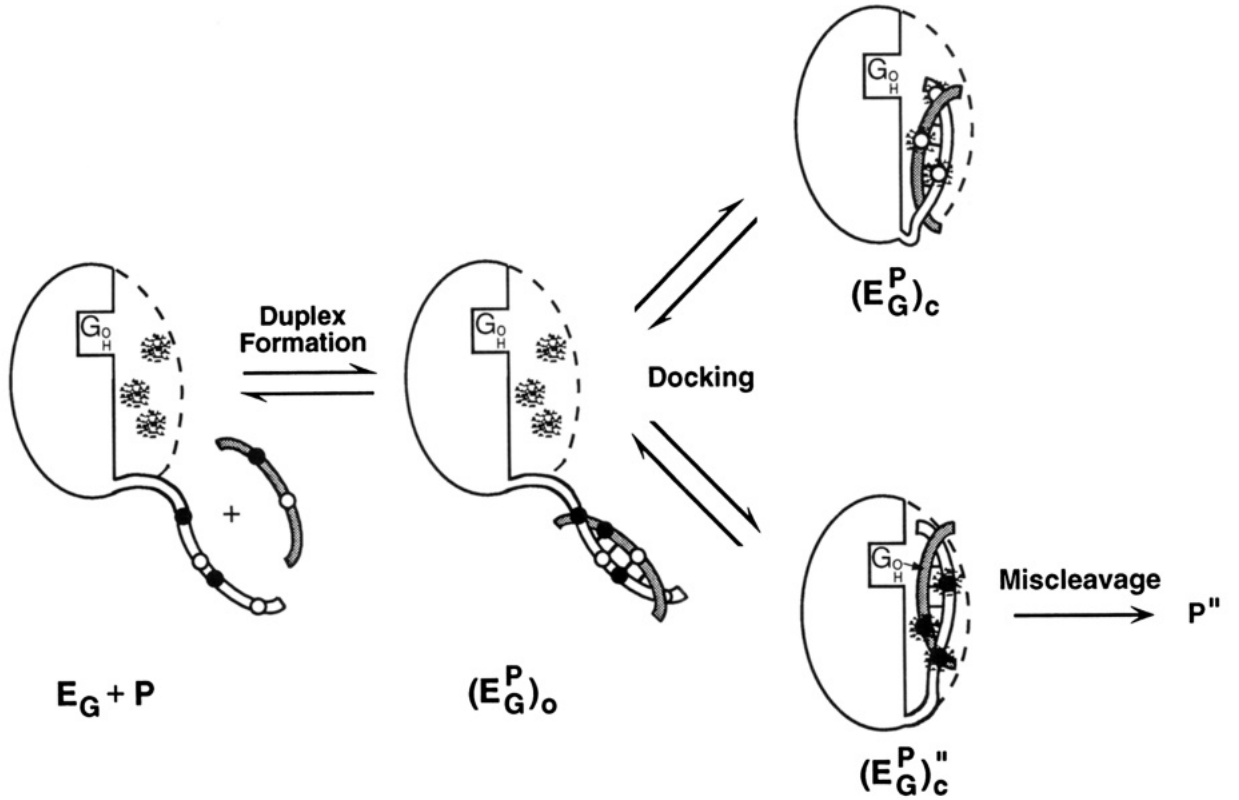


FIGURE 2: Two-step model for oligonucleotide binding. 5′-Exon analogs such as P (CCCUCU) initially bind in an open complex, $(E{\cdot}G{\cdot}P)_o$, in which the oligonucleotide appears to be held solely by base-pairing interactions in the P1 duplex. Subsequently, the P1 duplex can dock into tertiary interactions to give the closed complex, $(E{\cdot}G{\cdot}P)_c$ (Bevilacqua et al., 1992; Herschlag, 1992; G. J. Narlikar and D. Herschlag, manuscript in preparation). $(E{\cdot}G{\cdot}P)_c$ is the normal closed conformation, with the conserved U held at the catalytic site and tertiary interactions involving the 2′-hydroxyls at position −3 of P and at G22 and G25 of the IGS, which are represented by the open circles. However, misdocking has also been shown to occur, in which the P1 duplex docks into the tertiary interactions in different registers to form alternative closed complexes with different phosphodiester bonds positioned in the catalytic site; this leads to miscleavage (Herschlag, 1992; Strobel & Cech, 1994). Only one example of misdocking is shown in the figure for clarity: $(E{\cdot}G{\cdot}P)_c''$, which has tertiary interactions displaced two residues 5′ on P and the IGS (represented by the closed circles) and gives the P″ product, CCCU, upon cleavage. P″ is the predominant miscleavage product observed [see Figure 3A and Herschlag (1992)].

obtained by following a reaction analogous to the reverse reaction: E + GA + CCCUCX → E + G + CCCUCXA (X = U or C). In these reactions, 1 mM GA was used and reactions were initiated by the addition of 5′-labeled CCCUCU or CCCUCC to a final concentration of ~1 nM. Inhibition curves were obtained with six or more IGS′ or IGS′A concentrations in at least two independent experiments. Concentrations of the IGS analogs were used that spanned the observed K_d value by 5−10-fold on either side.

Control experiments were performed with varying concentrations of ribozyme (1−50 nM) to establish that the reactions were carried out under $(k_{cat}/K_m)^S$ conditions, such that the observed rate is directly proportional to the concentration of free S and the K_I value for IGS′ (or IGS′A) equals the K_d value for the model P1 duplex between S and IGS′ (or IGS′A). Analogous control experiments were performed to establish $(k_{cat}/K_m)^P$ conditions for the reverse reaction. To disrupt possible aggregates, aliquots of the IGS′ (or IGS′A) strand were heated to 95 °C for 5 min and cooled to room temperature before use. The following additional results established that free oligonucleotide, free ribozyme, and free IGS′ were in equilibrium: (1) good first-order kinetic behavior was observed for the disappearance of the oligonucleotide; (2) preincubation of IGS′ and ribozyme before addition of the oligonucleotide gave the same inhibition as preincubation of IGS′ and S before the addition of ribozyme; and (3) variation in the time course for the reaction by varying ribozyme concentration did not affect the observed inhibition. The reaction rate approached 0 at high IGS′ concentrations, indicating that there was no significant reaction of the P1 analog IGS·S *in trans*. Additional controls will be presented elsewhere (G. J. Narlikar and D. Herschlag, manuscript in preparation).

IGS′ (and IGS′A) concentrations were calculated from absorbance measurements at 260 nM and pH 7.0 using extinction coefficients calculated as a sum of the extinction coefficients of the individual bases.

RESULTS

Previous results have indicated that the conserved G·U pair of group I introns is important in 5′-splice site selection (see introduction). This pair could, in principle, affect one or a number of reaction steps such as folding of the intron, docking of the P1 duplex into the catalytic core, other conformational steps, binding of the G cofactor, and the chemical cleavage step. These individual reaction steps are best understood for the L-21ScaI ribozyme, which catalyzes a reaction analogous to the first step of self-splicing for its parent intron, that from *Tetrahymena thermophila* pre-rRNA (eq 1). We have therefore used the kinetic and thermody-

$$\overset{-5\ -3\ -1\ +2}{\text{CCCUCUAAAAA}} + \text{G} \rightleftharpoons \text{CCCUCU} + \text{GAAAAA} \qquad (1)$$

(S) (P)

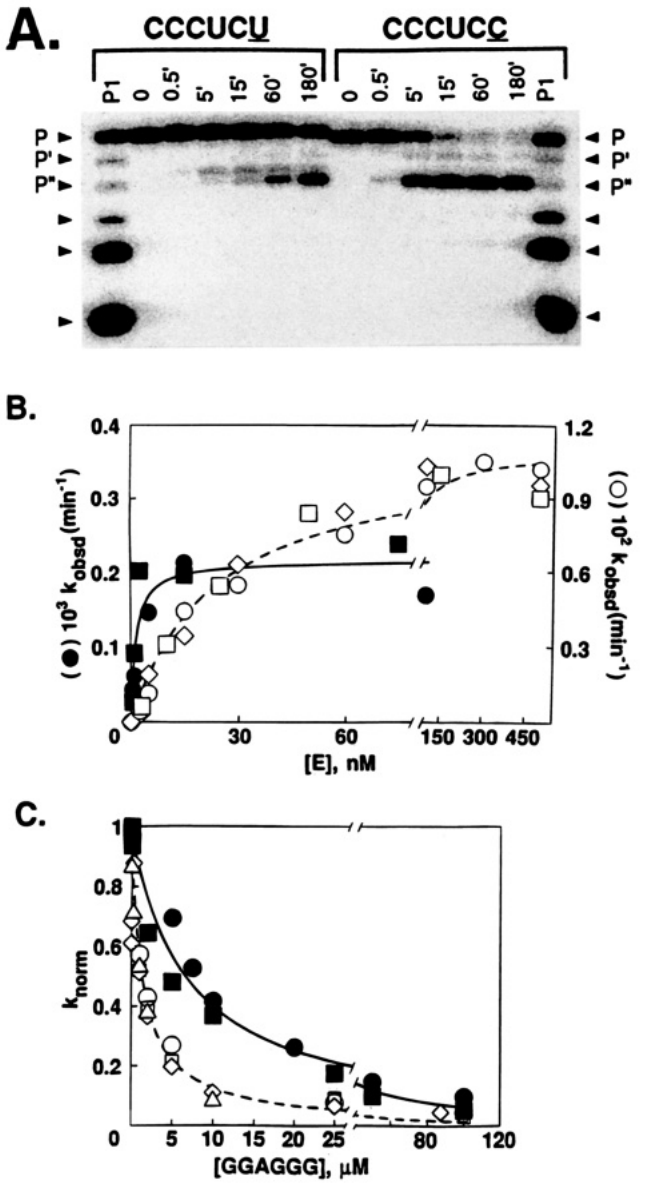


FIGURE 3: Comparison of miscleavage and binding affinities of CCCUCU and CCCUCC. (A) PhosphorImager printout showing time courses and products from the miscleavage reactions with 300 nM E, 2 mM G, and ~1 nM 5′-labeled CCCUCU or CCCUCC (50 mM Na MES (pH 6.6) and 10 mM $MgCl_2$, 50 °C). The products were identified as P″ (CCCU) and P′ (CCCUC; seen in longer exposures) by comparison with the ladder generated by partial digestion with P1 nuclease, and their ratio was determined to be P″/P′ = ~20 and ~15 for the U- and C-containing oligonucleotides, respectively, by PhosphorImager quantitation of analogous data. (For the CCCUCU reaction, there is an extra band that is observed at early times and migrates slightly slower than P″. This product accounts for less than 1% of the total material and thus does not significantly contribute to the results obtained; it was not identified.) (B) Dependence of the rate constant for miscleavage of CCCUCU (closed symbols) and CCCUCC (open symbols) on [E] in single-turnover reactions with subsaturating G (25 μM) and ≤1 nM 5′-labeled oligonucleotide (50 mM Na MES (pH 6.6) and 10 mM $MgCl_2$ at 50 °C). The lines represent nonlinear least-squares fits to the data with $k_c^{U,app} = 2.2 \times 10^{-4}$ min^{-1} and $k_c^{C,app} = 0.011$ min^{-1} (for the U and C oligonucleotides, respectively), $K_m^U = 1.5$ nM and $K_m^C = 24$ nM, and $(k_{cat}/K_m)^{U,app} = 1.5 \times 10^5$ M^{-1} min^{-1} and $(k_{cat}/K_m)^{C,app} = 4.6 \times 10^5$ M^{-1} min^{-1}. Different symbols represent results from independent experiments. Analogous experiments with saturating G (2 mM) instead of 25 μM G gave the following values: $k_c^U = 2.6 \times 10^{-3}$ min^{-1} and $k_c^C = 0.21$ min^{-1}, $K_m^U = 2.8$ nM and $K_m^C = 31$ nM, and $(k_{cat}/K_m)^U = 9.3 \times 10^5$ M^{-1} min^{-1} and $(k_{cat}/K_m)^C = 6.8 \times 10^6$ M^{-1} min^{-1} (data not shown). Evidence that the K_m values represent actual dissociation constants is described in the Materials and Methods. (C) Stability of duplexes between CCCUCU (closed symbols) or CCCUCC (open symbols) and the IGS analog, GGAGGG. Single-turnover reactions of ~1 nM 5′-end-labeled CCCUCU or CCCUCC with GA were catalyzed by 5 nM E in the presence of 50 mM Na MES (pH 6.6) and 10 mM $MgCl_2$ at 50 °C in the presence of varying concentrations of IGS′ (GGAGGG). The rate constants were normalized such that $k_{norm} = 1$ in the absence of IGS′ inhibitor. Different symbols represent results from independent experiments. The lines represent nonlinear least-squares fits to the data and give $K_I = 7.0$ and 1.6 μM for CCCUCU and CCCUCC, respectively. Control experiments established that K_I equals K_d for the respective duplexes (see Materials and Methods).

namic reaction framework for the L-21ScaI ribozyme (Figure 1B) to discern the contribution of the conserved G·U pair in individual reaction steps. The first section of the Results describes the miscleavage of 5′-exon analogs by this ribozyme as an initial probe of the role of the G·U pair in 5′-splice site selection.

Effects of Changes in the G·U Pair on Miscleavage Suggest That Tertiary Interactions with This Pair Help Position the 5′-Splice Site. There are tertiary interactions involving the 2′-hydroxyl groups of both the 5′-exon analog and the IGS strand of the P1 duplex (Bevilacqua & Turner, 1991; Pyle & Cech, 1991; Pyle et al., 1992; Herschlag et al., 1993a; Strobel & Cech, 1993). Consistent with such tertiary interactions, 5′-exon analogs bind the ribozyme in two steps: there is initial formation of an open complex, $(E{\cdot}P)_o$, in which the oligonucleotide appears to be held solely by base-pairing interactions, followed by docking into tertiary interactions to form a closed complex, $(E{\cdot}P)_c$ (Figure 2; Bevilacqua et al., 1992; Herschlag, 1992; G. J. Narlikar and D. Herschlag, manuscript in preparation). Miscleavage of 5′-exon analogs to give shortened oligonucleotides can occur when the P1 duplex docks into its tertiary interactions in a different register, as depicted in Figure 2 (Herschlag, 1992; Strobel & Cech, 1994).

Figure 3A shows a comparison of the miscleavage of the normal 5′-exon analog, CCCUCU, with CCCUCC, in which the U of the G·U pair in the P1 duplex has been replaced with C. The products and their ratios are the same, but CCCUCC is miscleaved 50−80-fold faster than CCCUCU in single-turnover reactions with saturating ribozyme (Figure 3B, k_c and k_c^{app}). The slower miscleavage of the U-containing oligonucleotide suggests that there is a favorable interaction of the G·U pair that contributes 50−80-fold to tertiary binding interactions, but only in the correct docking mode, $(E{\cdot}P)_c$ (Figure 2). This would then reduce misdocking and subsequent miscleavage of the U-containing oligonucleotide by 50−80-fold. The alternative explanations that the 50−80-fold faster miscleavage with C than with U arises from unfavorable interactions of the G·U pair in the incorrect docking mode or from unfavorable interactions of the G·C pair in the correct docking mode are countered by the observation that the 5′-exon analog with the terminal U removed, CCCUC_, is miscleaved at essentially the same rate as CCCUCC [k_c = 0.0022 and 0.0011 min^{-1} for CCCUCC and CCCUC_, respectively, with saturating E (100−1000 nM), 2 mM G, and 50 mM MES (pH 6.7), 30 °C].

Test of the Model That Tertiary Interactions with the G·U Pair Contribute to Proper P1 Docking. The following experiments were performed to test the model, derived from the miscleavage results, that tertiary interactions with the G·U pair contribute 50−80-fold to the correct docking of the P1

Scheme 1

$$E^G + S \xrightarrow{k_{cat}/K_m} E + P + GA_5$$

$$\pm IGS' \quad \Big\| \; K_d^{P1}$$

$$IGS' \cdot S$$

duplex. The binding of CCCUCU and CCCUCC to the ribozyme was first compared by determining the dependence of miscleavage on [E] (Figure 3B). The 16-fold stronger binding of CCCUCU than CCCUCC to the ribozyme is remarkable given that the G·C pair is predicted to be *more* stable than the G·U pair in a duplex (Freier et al., 1986). This confirms previous results from equilibrium gel binding experiments (A. Pyle and T. R. Cech, personal communication). The stronger binding of the U-containing oligonucleotide thus suggested that there are tertiary interactions of the ribozyme with the G·U pair that more than offset its weaker duplex stability.

To confirm that the G·U pair does indeed contribute less to duplex stability in the context of the P1 duplex and to quantitate this effect, the stabilities of model P1 duplexes between CCCUCU or CCCUCC and the IGS analog, GGAGGG (IGS′), were measured. The technique of substrate inhibition was employed, which is outlined in Scheme 1 and will be described elsewhere in detail. Briefly, the oligonucleotide IGS′ is used as a competitive inhibitor of the ribozyme reaction. But instead of inhibiting the reaction by binding to the ribozyme and blocking its active site, IGS′ binds to the oligonucleotide substrate (IGS′·S), thereby preventing cleavage of the substrate by the ribozyme. Since binding in this case is simply duplex formation, substrate inhibition allows the determination of duplex stabilities.

The results from such determinations of duplex stabilities for CCCUCU and CCCUCC with IGS′ are shown in Figure 3C. The C-containing oligonucleotide bound 4.5-fold stronger than the U-containing oligonucleotide. This binding difference of 0.96 kcal/mol, measured at 50 °C (10 mM $MgCl_2$), agrees well with the difference of 1.0 kcal/mol predicted from nearest-neighbor rules (1 M NaCl and 37 °C; Freier et al., 1986). There is also reasonable agreement between the observed and predicted values for the individual duplex stabilities. Furthermore, the 3−6-fold difference observed in k_{cat}/K_m values for miscleavage of CCCUCU and CCCUCC is similar to the difference in duplex stability (Figure 3B and data not shown). This suggests that the G·U and G·C pairs behave as simple base pairs when P1 is not docked in the correct register [e.g., in $(E{\cdot}P)_c''$ for miscleavage; Figure 2]. This is in contrast to the additional tertiary stabilization obtained from the G·U pair when P1 is correctly docked [$(E{\cdot}P)_c$] and further supports the conclusion that the G·U pair provides an extra contribution to binding only when it is positioned correctly in the active site.

In summary, the duplex with the U-containing oligonucleotide is less stable than that with the C-containing oligonucleotide, as predicted. The 16-fold stronger binding of CCCUCU than CCCUCC to the ribozyme combined with the 4.5-fold weaker intrinsic duplex stability of the U-containing oligonucleotide suggests that tertiary interactions with the G·U pair contribute ~70-fold (16 × 4.5 = 72) to the binding of CCCUCU. This is in good agreement with the 50−80-fold effect obtained in the miscleavage assay. Thus, tertiary interactions with the conserved G·U pair are

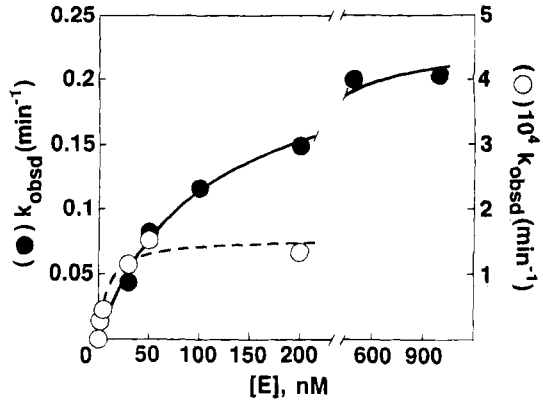


FIGURE 4: Comparison of reactions of −1r,dSU (closed symbols) and −1r,dSC (open symbols): Single-turnover reactions of ~1 nM 5′-end-labeled oligonucleotide substrate with varying [E] and 2 mM G in 50 mM Na MES (pH 6.7) and 10 mM $MgCl_2$ at 30 °C. Note the different axes. The lines represent nonlinear least-squares fits to the data and give k_c = 0.23 and 1.5 × 10^{-4} min^{-1} and K_m = 101 and 9.4 nM for −1r,dSU and −1r,dSC, respectively. Evidence that the observed values of K_m represent equilibrium dissociation constants is described in the Materials and Methods.

used for correct positioning of the P1 duplex, thereby contributing to selection of the correct 5′-splice site.

Overall Effect of the G·U Pair Determined in the Deoxyribose Background. The endonucleolytic cleavage of −1r,dSU and −1r,dSC (eq 1) is compared in Figure 4 (note the different axes). These substrates contain a single 2′-hydroxyl group at the cleavage site, which allows for a faster rate of chemical cleavage (Herschlag et al., 1993b), while the deoxyribose background allows for the probing of several reaction steps at once, as described here.

The maximal rate of the single-turnover reaction, k_c, with the U-containing substrate was 1200-fold faster than that with the C-containing substrate (Figure 4 and data not shown). For these substrates, k_c represents the reaction beginning with (1) S bound in the *open* complex and includes (2) the docking step and (3) the chemical step, as depicted in Figure 5A. Evidence for the involvement of these steps is as follows.

(1) Initial Binding in the Open Complex. The C-containing substrate binds to the ribozyme ~10-fold stronger than the U-containing substrate (K_d = 8 and 90 nM for the C- and U-containing oligonucleotide, respectively; Figure 4 and data not shown). This is consistent with the ~25-fold stronger duplex formed between −1r,dSC and the IGS analog, IGS′, than that between −1r,dSU and IGS′ (K_d = 30 and 800 nM, respectively, determined by substrate inhibition at 30 °C).[2] In addition, −1r,dSU binds to a different IGS analog, GGAGGGA, with the same affinity as it binds the ribozyme, suggesting that this oligonucleotide binds to the ribozyme as a simple duplex without additional tertiary interactions (G. J. Narlikar and D. Herschlag, manuscript in preparation; this IGS analog contains the 3′-A of the ribozyme's $J_{1/2}$ sequence, which allows the same potential stacking interactions in the model duplex as in the ribozyme's open complex). Further, the absence of coupled binding between G and −1r,dSU or dSU suggests that substrates in the deoxyribose background bind in the open complex (McCo-

[2] An ~4-fold larger difference in the stability of duplexes with the C- and U-containing substrates than that of duplexes with the C- and U-containing products may result from more favorable stacking interactions of the 3′-A of the substrate with a G·C Watson−Crick pair relative to a G·U wobble pair (G. J. Narlikar and D. Herschlag, unpublished results).

Table 1: Effects of Cleavage Site Base Identity on Binding and Reactivity of Substrates in the Deoxyribose Background[a]

	k_c		K_m		$(k_{cat}/K_m)^S$	
substrate	(min^{-1})	relative	(nM)	relative	(M^{-1} min^{-1})	relative
−1r,dSU	2.1×10^{-1}	(1)	85	(1)	2.5×10^6	(1)
−1r,dSC	1.7×10^{-4}	1/1200	8.2	1/11	2.1×10^4	1/120
−1r,dSA	2.6×10^{-4}	1/800	360	4.2	7.2×10^2	1/3500
−1r,dSG	2.3×10^{-5}	1/9000	470	5.5	4.9×10^1	1/5100

[a] Reactions with 50 mM Na MES (pH 6.7), 2 mM G, and 10 mM $MgCl_2$ at 30 °C.

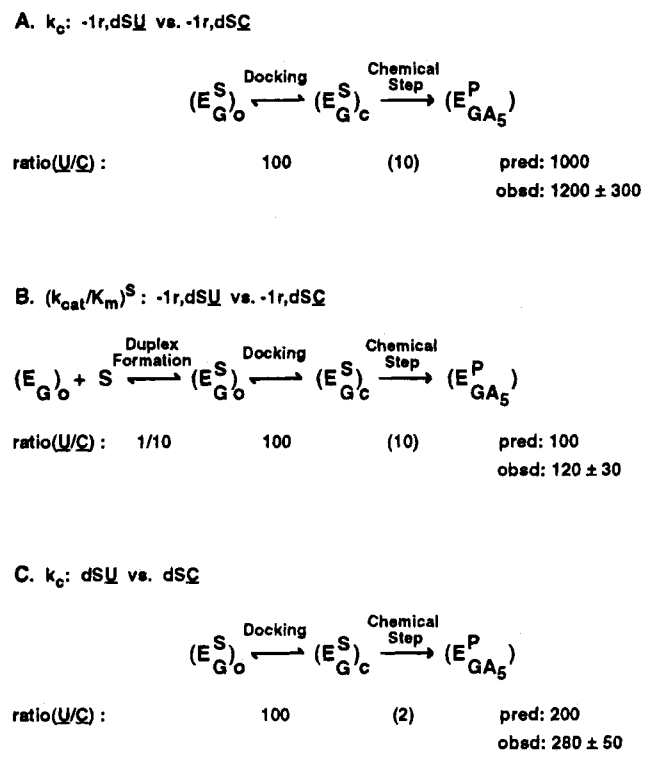


FIGURE 5: Attribution of rate effects in individual reaction steps from the comparison of U- and C-containing oligonucleotide substrates in the deoxyribose background. Comparison of k_c (A) and $(k_{cat}/K_m)^S$ (B) for reactions of −1r,dSU vs −1r,dSC, and comparison of k_c for dSU vs dSC (C). The observed values were obtained in 50 mM Na MES (pH 6.7) and 10 mM $MgCl_2$ at 30 °C from the data of Figure 4 and additional data not shown. The values used in the prediction for individual steps were obtained as follows. The value of 1/10 for duplex formation in B is the ratio obtained for binding of −1r,dSU and −1r,dSC to the ribozyme. The value of 100 for docking in A, B, and C was obtained as described in the text. The values of 10 in the chemical step for −1r,dS (A and B) and 2 in the chemical step for dS (C) were obtained from the data of this figure to give good agreement between the overall observed and predicted values in this figure and in Figure 8 and are shown in parentheses in order to indicate that they were not obtained independently.

nnell et al., 1993; D. Herschlag and P. Legault, unpublished results). Additionalevidence for and characterization of the binding in the open complex will be presented elsewhere.

(2) Docking Is Required for Cleavage. The 2′-hydroxyl group at position −3 of the substrate is involved in a tertiary interaction with the ribozyme (Bevilacqua and Turner, 1991; Pyle & Cech, 1991; Pyle et al., 1992). This hydroxyl group contributes to binding, but provides no further contribution to the chemical step (Herschlag et al., 1993a). The strong inhibitory effect on the cleavage rate upon substitution of a 2′-methoxy group at this position suggests that disruption of docking slows the reaction (Herschlag et al., 1993a; G. J. Narlikar and D. Herschlag, manuscript in preparation).

(3) The Chemical Cleavage Step Is Rate-Limiting. The log-linear pH dependence for reaction of these substrates (data not shown) and the rate decrease upon substitution of sulfur for the *pro-*R_P nonbridging phosphoryl oxygen at the cleavage site of −1r,dSU (Herschlag et al., 1993b; D. S. Knitt and D. Herschlag, manuscript in preparation) strongly suggest that the actual cleavage step is rate-limiting, as discussed previously (Herschlag et al., 1993b; Herschlag & Khosla, 1994).

The ~100-fold stronger tertiary interactions with G·U than with G·C described in the previous sections[3] can account for most of the 1200-fold larger value of k_c for −1r,dSU than that for −1r,dSC. As outlined in Figure 5A, the remainder can be accounted for by a ~10-fold effect on the chemical step. A 10-fold effect on the chemical step is also obtained from comparison of the $(k_{cat}/K_m)^S$ values for these substrates (Figure 5B).

Table 1 summarizes the effect of the four natural base substitutions at position −1. The A- and G-containing substrates bind ~5-fold weaker than the U-containing substrates and ~50-fold weaker than the C-containing substrate, which is in reasonable agreement with the destabilization predicted for the introduction of a mismatch into an RNA/RNA duplex (Freier et al., 1986). The similar effects of the C-containing substrate, which can form a Watson−Crick base pair with the IGS, and the A- and G-containing substrates on the docking and chemical steps (Table 1, k_c) provide no indication of large inhibitory effects from unfavorable interactions of non-G·U pairs at the active site. However, the ~10-fold different overall effect of the A and G substitutions suggests that local interactions of or with the duplex can affect the docking or chemical steps, even in the absence of the conserved G·U pair.

The overall difference in k_c for reactions of dSU and dSC is ~280-fold, which is surprising given that there is a 1200-fold difference when there is a ribose residue at position −1 (Figure 5A,C; data not shown). This difference suggests that the G·U pair contributes ~5-fold more in the chemical step when there is a 2′-hydroxyl on the U residue. An alternative explanation is that this 2′-hydroxyl preferentially contributes to formation of the closed complex with a G·U pair rather than contributing to the actual chemical step. However, this alternative is rendered unlikely by the observation that the ~3-fold weaker binding in the closed complex upon replacement of the 2′-hydroxyl group with a hydrogen atom at

[3] A value of 100 for the effect of the G·U pair on the docking of S is used throughout rather than the value of 50−80 observed for the effect on the docking of P. This rounded value is adopted because of limits in the accuracy of this determination (for example, there are differences in the contribution from stacking of the 3′-A overhang in model P1 duplexes with SU and SC that may affect the docking equilibrium) and to give overall consistency for the predicted vs observed values in Figures 5 and 8. It should be emphasized that there may be errors in these estimates on the order of a factor of 2 due both to experimental uncertainty and to a small amount of nonindependence from multiple substitutions; however, such small uncertainties do not affect the conclusions herein.

position −1 is the same as that observed in simple duplexes [measurements of duplex stability: $K_d^{IGS'}$(rSU) = 1.0 μM and $K_d^{IGS'}$(−1d,rSU) = 2.5 μM; 50 mM Na MES (pH 6.6) and 10 mM $MgCl_2$, 50 °C; see Herschlag et al. (1993a) and Bevilacqua and Turner (1991) for ribozyme data and additional duplex data].

Additional experiments are presented below that were designed to isolate the chemical step and to further test the differential effect of the G·U pair when there is a 2′-hydroxyl vs a 2′-hydrogen at position −1. In the course of these experiments, complexities were observed that have led to additional insights concerning the G·U pair. The experimental support for the cooperative effect in the chemical step is described after these additional insights, and a model to account for this cooperative energetic effect between the 2′-hydroxyl of U(−1) and the G·U pair is described in the Discussion.

Unusual pH Dependencies in the Ribose Background Provide Evidence for a G·C⁺ Wobble Pair. The U- and C-containing oligonucleotide substrates were compared in the ribose background in an attempt to isolate the chemical cleavage step. An initial comparison of $(k_{cat}/K_m)^G$, the second-order rate constant for the reaction of E·S and G, gave a 50-fold difference at pH 6.6 (rSU vs rSC). Previous results have established that the chemical step is rate-limiting for the reaction of the U-containing substrate under these conditions (Herschlag & Khosla, 1994), so that our initial interpretation of this result was that the G·U pair enhanced the rate of the chemical step by 50-fold. However, this is larger than the value of 10 obtained in the deoxyribose background. The results described below indicate that two complications account for this discrepancy, while suggesting additional features of the reaction: (1) rSC binds predominantly in the open complex, so that there is a barrier for formation of the closed complex (see Figure 2) in addition to the barrier for the chemical step, and (2) a G·C⁺ wobble pair is formed in the P1 duplex with rSC and contributes to the reaction of this substrate, even at neutral pH. Evidence for the G·C⁺ wobble pair is described first.

The chemical step in the ribozyme reaction requires the loss of a proton and is therefore log-linear with pH, as previously determined for rSU, −1r,dSU, and dSU (Herschlag & Khosla, 1994; D. S. Knitt and D. Herschlag, manuscript in preparation). The pH dependence of $(k_{cat}/K_m)^G$ for rSC was thus obtained and compared to that for rSU to determine whether its reaction was limited by the actual chemical step (Figure 6). The pH dependence for rSC was unexpectedly complex, with log-linear behavior at low pH, leveling with an apparent pK_a of ~5.5, and a second region of log-linear behavior above pH 8. Also unexpected were the similar values of $(k_{cat}/K_m)^G$ for the two substrates at the low pH values.

To determine the cause of the rSC pH dependence, possible contributions from a conformational step were considered. There is precedence for a rate-limiting conformational step in the reaction of rSU above pH 7, which is responsible for the leveling off in its pH−rate profile (Figure 6; Herschlag & Khosla, 1994). Thus, a thio effect, which distinguishes between the chemical and conformational steps for rSU (Herschlag & Khosla, 1994), was used to probe whether a chemical or conformational step is rate-limiting for the rSC reaction in each of the pH regimes. In this approach, one of the nonbridging phosphoryl oxygen atoms at the cleavage site is replaced by sulfur. Replacement of

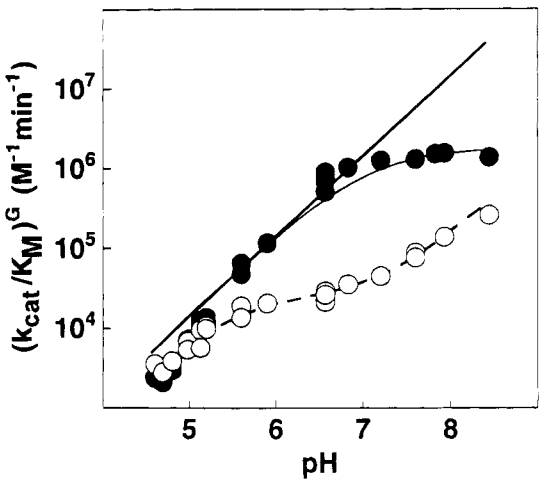


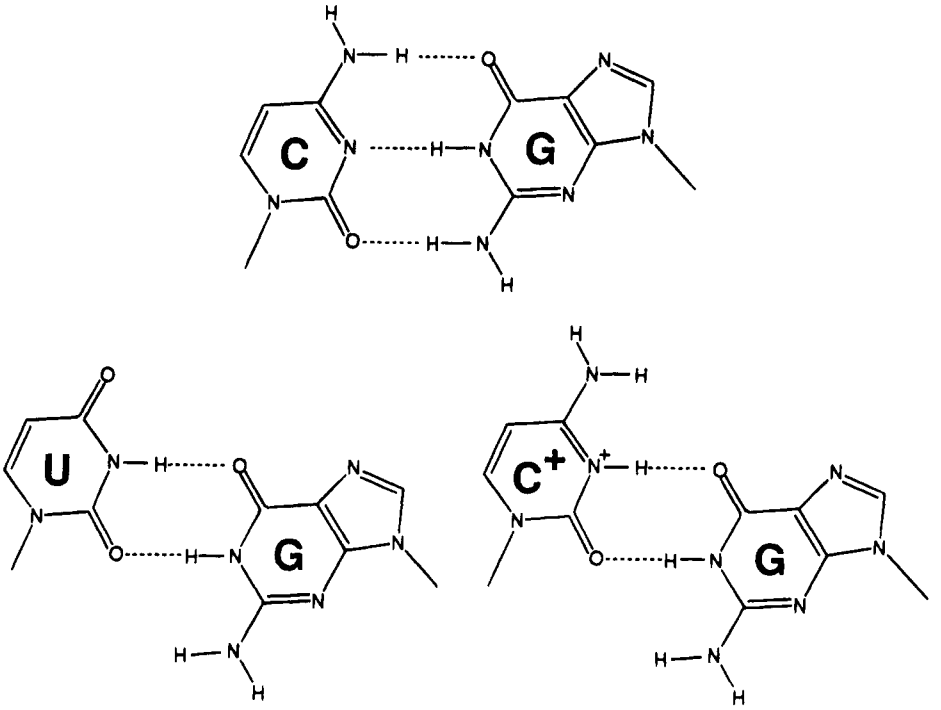
FIGURE 6: Comparison of the pH dependence of $(k_{cat}/K_m)^G$ for rSU (closed symbols) and rSC (open symbols) (10 mM $MgCl_2$, 50 °C). The fits are nonlinear least-squares fits to the data according to the model of Scheme 2 and eqs 2−7. The following values were obtained from the fits: $k_{G\cdot C}$ = 0.0012 min^{-1}, $k_{G\cdot C^+}$ = 0.08 min^{-1}, $k_{G\cdot U}$ = 0.13, and p$K_{G\cdot C}$ = 5.5. The leveling off at pH >7 in the rSU reaction arises from a rate-limiting conformational step (Herschlag & Khosla, 1994) and was therefore not used in comparisons of rates of the chemical step. Rather, the fit corresponding to the straight line was used because the continued log-linear increase in the rate of the −1d,rSU reaction beyond pH 8 suggests that the actual chemical step remains pH dependent above pH 7 (Herschlag & Khosla 1994; Figure 7). There are systematic deviations in the data below pH 5.5 that presumably arise from additional titrations of ribozyme functional groups (D. S. Knitt and D. Herschlag, manuscript in preparation).

the *pro-*R_P oxygen has a modest effect of 2−3-fold on the chemical step in the ribozyme reaction, which is similar to the effect of 4−11-fold observed in model reactions, but this replacement does not affect the binding and conformational steps in the reaction (Herschlag et al., 1991; Herschlag & Khosla, 1994). A thio effect of greater than 1.0 for the rSC reactions was observed over the entire the pH range [thio effects of 1.6, 1.8, and 1.5 were obtained at pH 5.0, 6.6, and 7.6, respectively, for $(k_{cat}/K_m)^G$, similar to the value of 2.3 obtained for rSU (Herschlag & Khosla, 1994; see Materials and Methods)]. This suggests that the chemical step is rate-limiting and that a rate-limiting conformational change does not account for the unusual pH dependence.

Protonation of N3 of C to allow the formation of a G·C⁺ wobble pair with the same geometry as a G·U wobble pair (Chart 1) can account for the pH dependence of Figure 6 and can also account for additional data described below. Formation of the G·C⁺ wobble pair at low pH accounts for the observation that rSC reacts as fast as rSU at low pH (Figure 6).[4] The reactions of both rSU and rSC are first-order in hydroxide ion concentration at low pH because a

[4] The value of $k_{G\cdot C^+}$ obtained from the fit in Figure 6 is 1.6-fold lower than that for $k_{G\cdot U}$, even though the actual rates are ~30% faster for the C-containing substrate at the lowest pH's. This is because of the nonideal behavior of the pH−rate profiles for both substrates below pH ~5.5, in which the decrease in rate with decreasing pH is steeper than was predicted from eq 2. This nonideality was not accounted for in the theoretical fits and the deviations lower the value of $k_{G\cdot C^+}$ obtained, whereas the value of $k_{G\cdot U}$ is not lowered because it was obtained from a fit of only the data above pH 5.5. The deviations arise presumably because of additional titratable groups on the ribozyme that affect activity (D. S. Knitt and D. Herschlag, manuscript in preparation), but do not affect the conclusions herein. The same overall conclusions are drawn from the reactions of −1d,rSU and −1d,rSC with saturating G, where these deviations are not observed (Figure 7).

Chart 1



proton is lost in the reaction prior to the chemical step (Herschlag & Khosla 1994; D. S. Knitt and D. Herschlag, manuscript in preparation), as depicted in Scheme 2 and eqs 2−7. The rSC reaction levels off at intermediate pH values because the predominant bound form is the unprotonated G·C Watson−Crick pair ($E^{G\cdot C}$; Scheme 2), whereas the predominant reactive form is the protonated G·C$^+$ wobble pair ($E^{G\cdot C^+}$). Thus, for reaction to occur, the G·C pair must gain a proton while another functional group must lose a proton, so that there is no net gain or loss of a proton and no pH dependence in this region. At the higher pH values, the reaction again becomes log-linear with pH because the concentration of the more reactive G·C$^+$ species becomes low enough that the slower reaction of the G·C species predominates.

$$k_{obsd} = \{(k_{cat}/K_m)^{G}_{G\cdot C}(K_{G\cdot C}/(K_{G\cdot C} + [H^+])) + (k_{cat}/K_m)^{G}_{G\cdot C^+}([H^+]/(K_{G\cdot C} + [H^+]))\}[G] \quad (2)$$

$$K_{G\cdot C} = [H^+][E^{G\cdot C}]/[E^{G\cdot C^+}] \quad (3)$$

$$pK_{G\cdot C} = -\log K_{G\cdot C} \quad (4)$$

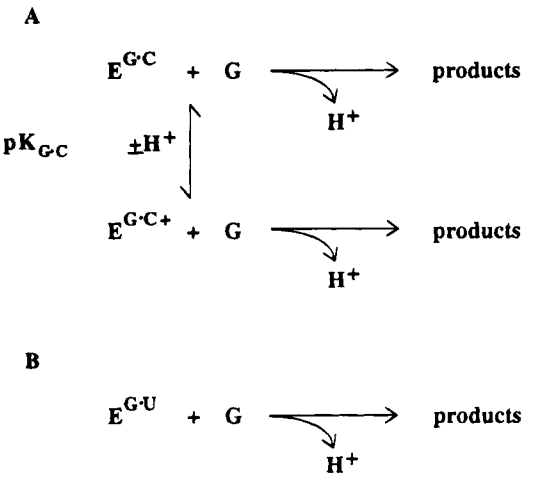
$$(k_{cat}/K_m)^{G}_{G\cdot C} = k_{G\cdot C}/[H^+] \quad (5)$$

$$(k_{cat}/K_m)^{G}_{G\cdot C^+} = k_{G\cdot C^+}/[H^+] \quad (6)$$

$$(k_{cat}/K_m)^{G}_{G\cdot U} = k_{G\cdot U}/[H^+] \quad (7)$$

The rSC substrate in its unprotonated form reacts ~100-fold slower than rSU ($k_{G\cdot C}$ vs $k_{G\cdot U}$; eqs 5 and 7 and Figure 6). This difference is larger than the estimate of 10 in the previous section from the reactions in the deoxyribose background because there is an additional barrier for rSC that is not present for rSU: rSU is bound stably in the closed complex, whereas rSC is bound predominantly in the open complex and thus must dock into the closed complex prior to reaction, as described in the following section (Figure 2). This complication did not interfere with the −1r,dSU and −1r,dSC comparison because both are bound predominantly in the open complex (see above).[5]

Scheme 2



Several additional observations support the model in which a G·C$^+$ wobble pair is formed at low pH, with its formation favored by tertiary interactions with the ribozyme.

(1) The formation of a G·C$^+$ wobble pair at low pH accounts for the otherwise surprising observation that although there is a small amount of P′ (CCCUC) relative to P″ (CCCU) formed from CCCUCU or CCCUCC at high pH, similar amounts of P′ and P″ are formed at low pH (P″/P′ = 20 and 2 at pH 6.6 and 5.0, respectively, for both CCCUCU and CCCUCC; data not shown). Cleavage to give P″ occurs after an A·U pair of the P1 duplex, whereas cleavage to give P′ occurs after a G·C pair. Thus, the ability of this G·C pair to form a protonated wobble pair would be

[5] It should be noted that the formation of a G·C$^+$ wobble pair did not interfere with the determinations in the deoxyribose background as the reactions for both the U- and C-containing oligonucleotide substrates were log-linear with pH in the region above pH 7 used in the comparisons of Figure 5.

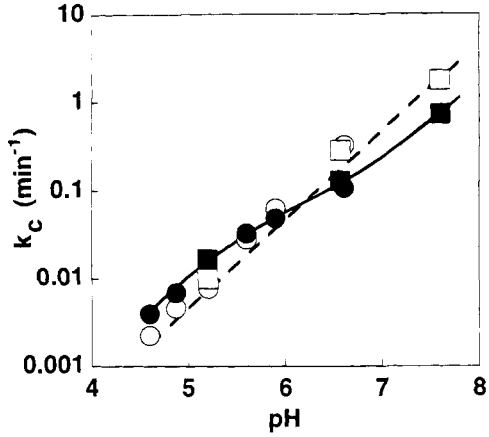


FIGURE 7: Comparison of the pH dependence of k_c for −1d,rSU (open symbols) and −1d,rSC (closed symbols). Reactions were carried out as described in the Materials and Methods with 2.8–4.3 mM G at 50 °C in 10 mM $MgCl_2$. The fits are nonlinear least-squares fits to the data according to the model of Scheme 2 and eqs 2–7, except that the reactions were followed with near-saturating G so that k_c rather than $(k_{cat}/K_m)^G$ was followed. The following values were obtained from the fits: $k_{G\cdot C} = 1.7 \times 10^{-8}$ M min^{-1}, $k_{G\cdot C^+} = 1.1 \times 10^{-7}$ M min^{-1}, $k_{G\cdot U} = 4.5 \times 10^{-8}$ M min^{-1}, and $pK_{G\cdot C} = 5.8$. (These rate constants have units of M min^{-1} rather than min^{-1} as in Figure 6 because G was saturating for the reactions in this figure.) Different symbols represent independent experiments.

expected to favor this misdocking mode and increase the rate of miscleavage to P′ at low pH.

(2) The ratio of miscleavage rates for CCCUCC relative to CCCUCU decreases from 80 at pH 6.6 to 40 at pH 5.6 and to 7 at pH 4.6. This is consistent with inhibition of miscleavage of the C-containing oligonucleotide due to tertiary interactions with the protonated G·C pair that favor correct docking and thereby disfavor incorrect docking as the pH is lowered.

(3) The weaker binding of CCCUCC than that of CCCUCU was less pronounced at low pH (4.6) than at neutral pH (data not shown), which is consistent with tertiary interactions stabilizing a protonated G·C$^+$ wobble pair in the P1 duplex with CCCUCC.

(4) There is an observed pK_a of ~6 in the pH dependence of the cleavage step for −1d,rSC, which is not observed in the pH dependence for −1d,rSU (Figure 7). This suggests that the tertiary interactions with the ribozyme perturb the pK_a of N3 of C by ~1.5 units from its free pK_a of ~4.5 (Jencks & Regenstein, 1976). This result is even more striking when it is considered that the formation of a Watson–Crick G·C pair is expected to lower the pK_a of C by about 1 unit, as a G·C pair is ~10-fold more stable than a G·U wobble pair in a simple duplex (see above).

Binding of the C-Containing Substrate in the Open Complex Resolves Remaining Paradoxes: A Cooperative Interaction in the Chemical Step between the G·U Pair and the 2′-Hydroxyl of the U. The cleavage of −1d,rSU from the E·S·G ternary complex is 3-fold faster than the cleavage of −1d,rSC (Figure 7, higher pH). This is in reasonable agreement with the 2-fold effect suggested earlier from the comparison of dSU and dSC (see also the following). At low pH, the C-containing substrate reacts 2-fold *faster* than the U-containing substrate (Figure 7), which is consistent with a similar rate of reaction for a G·U and G·C$^+$ wobble pair at the cleavage site. As for rSC, the thio effect for −1d,rSC is maintained over the entire pH range, suggesting that the chemical step remains rate-limiting [thio effects of 1.8, 2.3, and 2.0 were obtained at pH 5.0, 6.6, and 7.6, respectively, for k_c, which is similar to the value of 2.8 obtained for −1d,rSU (Herschlag & Khosla, 1994); see Materials and Methods].

There are two remaining paradoxes concerning the comparison of the chemical step for U- and C-containing substrates with 2′-hydroxyl vs 2′-hydrogen substituents at position −1.

(1) As stated earlier, the apparent effect of the G·U pair in the chemical step (relative to the G·C pair) with the 2′-hydroxyl at position −1 is ~10-fold in the deoxyribose background, whereas it is ~100-fold in the ribose background (Figures 5 and 6). (2) Although there is a 2–3-fold effect of the G·U pair on k_c for dS and for −1d,rS, there is a 16-fold effect of the value of $(k_{cat}/K_m)^G$ for the reaction of −1d,-rS. (The relative effects were determined from values obtained at pH values ≥7 so that reaction of the G·C$^+$ wobble pair did not interfere; data not shown.) Previous data indicate that $(k_{cat}/K_m)^G$ for −1d,rSU is limited by the chemical step (Herschlag & Khosla, 1994).

These paradoxes are both resolved by a model in which the C-containing substrates bind the ribozyme in the open complex (Figure 2), even in the ribose background. This would provide an additional barrier for the reactions of rSC [1, previous paragraph; Figure 8A,B]. Further, because there is coupling between the binding of oligonucleotide substrates and G, but only when the substrates bind in the closed complex (McConnell et al., 1993), saturating G would help drive bound −1d,rSC into the closed complex, thereby reducing the 16-fold difference in $(k_{cat}/K_m)^G$ for −1d,rSU and −1d,rSC to the ~3-fold difference in k_c with saturating G [2, previous paragraph]. This difference is in reasonable quantitative agreement with the previously observed coupling of ~5-fold (McConnell et al., 1993; cf. Figure 8B,C, K_{dock}^C = 1/6 and 5/6 in the absence and presence of bound G, respectively; see following). Thus, it appears that coupled binding of G and S is maintained when U is replaced by C; the effect is not detected in the observed binding constant for G because the oligonucleotide substrate binds in the open rather than the closed complex, thereby masking the coupling that occurs only in the closed complex.

Binding of the C-containing oligonucleotide in the open vs closed complex was tested in the experiments described in the next section. The results strongly suggest that these oligonucleotides do indeed bind predominantly in the open complex, supporting the conclusion that there is a ~10-fold effect of the G·U pair on the chemical step with a 2′-hydroxyl group at position −1, but only a ~2-fold effect with a 2′-hydrogen at position −1, as depicted in Figure 8.

Evidence for Binding of the C-Containing Substrate in the Open Complex. The binding of G is thermodynamically coupled to the binding of the oligonucleotide substrates rSU and −1d,rSU, but not to the binding of −1r,dSU (McConnell et al., 1993). It was therefore suggested that coupled binding occurs when substrates in the ribose background are bound in the closed complex, but that coupling is absent for substrates in the deoxyribose background because they are bound in the open complex. The absence of coupled binding of G and −1d,rSC is then consistent with binding of this substrate in the open complex ($K_d^{E\cdot G} = 1.4 \pm 0.5$ mM and $K_d^{E\cdot S\cdot G} = 1.0 \pm 0.2$ mM). As a change in the geometry of the P1 duplex upon replacement of the G·U pair with a G·C pair could provide an alternative explanation for the loss of coupling, it was important to probe further whether the

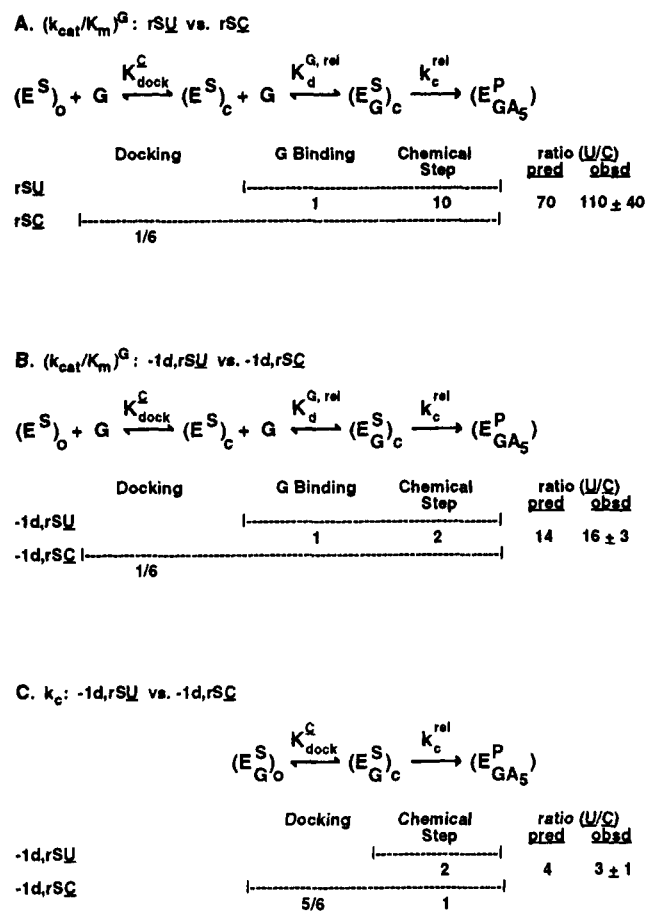


FIGURE 8: Summary of observed and predicted effects for replacing U with C in the ribose background. (A) Comparison of $(k_{cat}/K_m)^G$ for rSU and rSC. The values directly under each step represent the predicted values. The relative value in the chemical step, k_c^{rel} = 10, is from results with −1r,dSU and −1r,dSC (Figure 5). It is assumed that G binds the same to either E·S *closed* complex ($K_d^{G,rel}$ = 1; note that the observed binding of G is different because rSC is bound primarily in the open complex; see text). Results from previous studies indicate that rSU binds predominantly in the closed complex [e.g., Bevilacqua et al. (1992); Herschlag, 1992]. However, results described in the text suggest that rSC binds predominantly in the open complex, so that there is an additional barrier for its reaction. The value of K_{dock}^C = 1/6 (=[closed complex]/[open complex]) was selected to give good agreement with the observed difference. The observed difference of 110 listed in the figure was obtained from the ratio $k_{G\cdot U}/k_{G\cdot C}$ from Figure 6. (B) Comparison of $(k_{cat}/K_m)^G$ for −1d,rSU and −1d,rSC. The values directly under each step represent the predicted values, as described in A, except the relative value of k_c^{rel} = 2, which is from results with dSU and dSC (Figure 5). The same value of K_{dock}^C that was selected to give good agreement in part A also gives good agreement in this comparison, supporting the validity of this model. (C) Comparison of k_c for −1d,rSU and −1d,rSC. The values directly under each step represent the predicted values as described in A, except for K_{dock}^C = 5/6, which is 5-fold greater than K_{dock}^C = 1/6 used in part A to account for the 5-fold coupling in binding between G and oligonucleotide substrates bound in the closed complex (McConnell et al., 1993). Use of this factor gives reasonable agreement between the predicted and observed values. The observed value listed in the figure is from $k_{G\cdot U}/k_{G\cdot C}$ from Figure 7.

C-containing substrate binds in the open or closed complex. The following three independent results strongly support the conclusion that the C-containing oligonucleotides bind in the open complex.

(1) Specific 2′-hydroxyl groups of the P1 duplex are involved in tertiary interactions with the ribozyme and are expected to be present in the closed complex, but not in the open complex (Figure 2; Bevilacqua & Turner, 1991; Pyle & Cech, 1991; Pyle et al., 1992; Herschlag et al., 1993a; Strobel & Cech, 1993). When the 2′-hydroxyl group at position −3 (eq 1) is converted to a methoxy group, there is a large inhibitory effect on the cleavage reaction (Herschlag et al., 1993a), suggesting that a 2′-methoxy substituent at this position may act as a doorstop to inhibit docking into the closed complex. Thus, this substitution should have a large effect on the affinity for an oligonucleotide bound in the closed complex, but no significant effect on the affinity for an olignucleotide bound in the open complex. As expected, −3m,rSU binds ~10-fold more weakly than rSU, whereas there is no significant effect on duplex stability (G. J. Narlikar and D. Herschlag, manuscript in preparation). In addition, the affinity of −3m,rSU for the ribozyme and an IGS analog, GGAGGGA, is the same within 2-fold, suggesting that the methoxy-containing substrate is indeed bound in the open complex (G. J. Narlikar and D. Herschlag, manuscript in preparation). In contrast to the weakened binding of the U-containing oligonucleotide, introduction of a 2′-methoxy substituent in the C-containing oligonucleotide had no significant effect on binding (Figure 9). Control experiments indicated that this 2′-methoxy substitution had no significant effect on duplex stability. Thus, these results suggest that rSC binds predominantly in the open complex.

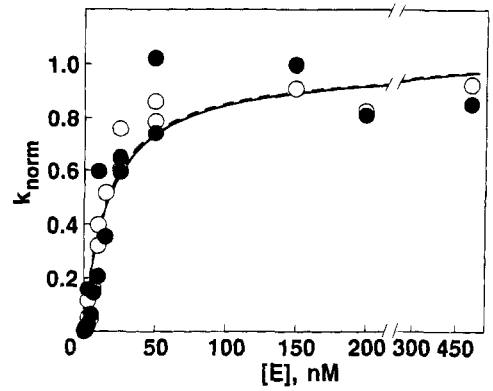


FIGURE 9: Comparison of the binding of rSC (○) and −3m,rSC (●) as a test of whether rSC binds in the open complex. Reactions were carried out in 50 mM Na MES (pH 6.6) and 10 mM $MgCl_2$ in the absence of G (hydrolysis reaction) at 50 °C with varying ribozyme concentrations. The lines represent nonlinear least-squares fits to the data and give K_m = 17 and 18 nM for rSC and −3m,rSC, respectively. The rate constants have been normalized by dividing each value of k_{obsd} by the value k_{max} = 0.061 and 0.0013 min^{-1} for rSC and −3m,rSC, respectively, which are the rate constants at infinite ribozyme obtained from the fits. Evidence that the K_m values represent actual dissociation constants is described in the Materials and Methods.

(2) The oligonucleotide substrate rSC binds to the ribozyme and the IGS analog, GGAGGGA, with the same affinity, within 2-fold ($K_d^{IGS'A}$ = 6 ± 2 nM and K_d^E = 13 ± 5 nM; 50 mM Na MES (pH 6.6) and 10 mM $MgCl_2$, 50 °C), suggesting that the C-containing substrate binds in the open complex.

(3) Comparison of the binding of a series of oligonucleotide substrates to the ribozyme and to an IGS analog, GGAGGGA, suggests that the closed complex with the rSU (and −1d,rSU) substrate is ~25-fold (2.1 kcal/mol; 50 °C) more stable than the open complex (G. J. Narlikar and D. Herschlag, unpublished results). This predicts that any perturbation of the P1 duplex that destabilizes the closed complex by more than ~25-fold relative to the open complex should result in binding that is predominantly in the open complex. Thus, the 100-fold destabilization of tertiary interactions upon replacement of the G·U pair with a G·C

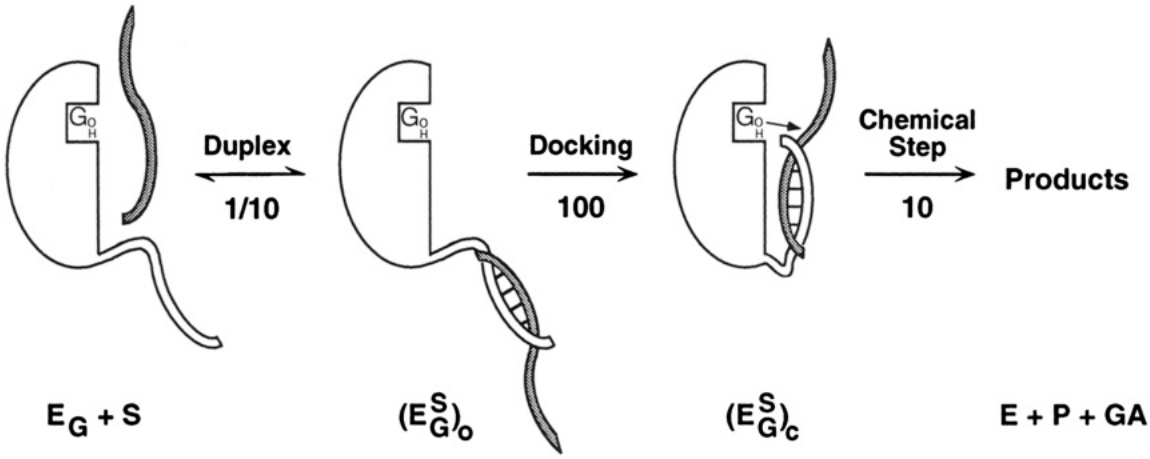


FIGURE 10: Summary of the contribution of the conserved G·U pair in individual reaction steps. Values are relative to a G·C pair.

pair is predicted to result in binding of the open complex that is favored by ~4-fold over the closed complex. This is in reasonable agreement with the value of K_{dock} = 1/6 {= [closed complex]/[open complex]} that was required to account for the overall differences in rates of reaction observed in Figure 8 ($K_{dock} \approx 25$ for −1d,rSU, so that 100-fold destabilization would give $K_{dock}^{C} \approx 25/100 = 1/4$).

A 2′-Methoxy Substituent at Position −1 also Destabilizes Docking. It was previously shown that the substitution of a methoxy group for the 2′-hydroxyl of U(−1) greatly weakens binding to the ribozyme [K_d^E(−1m,rSU) = 100 nM and K_d^E-(rSU) = 6 nM; Herschlag et al., 1993b; G. J. Narlikar and D. Herschlag, unpublished results]. Since the effect of this substitution on duplex stability was not known, this destabilization could not be ascribed with certainty to a disruption of tertiary interactions. In addition, substitution of a hydrogen atom for this 2′-hydroxyl group has only a ~3-fold effect on binding to the ribozyme, the same effect observed in model duplexes, thus providing no indication of a tertiary interaction with this 2′-hydroxyl [see above and Bevilacqua and Turner (1991); Pyle & Cech, 1991; Herschlag et al., 1993b].

The effect of the 2′-methoxy substitution on duplex stability was therefore measured. There is no significant destabilization of the model P1 duplex upon replacing the 2′-hydroxyl at position −1 with a 2′-methoxy substituent [$K_d^{IGS'}$(−1m,rSU) = 1.8 μM and $K_d^{IGS'}$(rSU) = 1.0 μM; 50 mM Na MES (pH 6.6) and 10 mM $MgCl_2$, 50 °C]. Thus, the methoxy substituent destabilizes tertiary interactions with the ribozyme even though the hydroxyl that it replaces does not appear to be involved in an interaction that contributes to tertiary stabilization (see Discussion).

The dissociation constant of 100 nM for the E·(−1m,rSU) complex is within ~3-fold of that expected for binding in the open complex (G. J. Narlikar and D. Herschlag, manuscript in preparation). Further, −1m,rSU and rSC undergo miscleavage at the same rate of 2−3 min^{-1} [Herschlag et al. (1993b) and data not shown], supporting the conclusion that both are bound in the open complex and the view that an increased rate of miscleavage can be related directly to destabilization of the tertiary interactions involved in aligning the P1 duplex.

DISCUSSION

Figure 10 summarizes the effects of replacing the conserved G·U pair with a G·C pair on individual steps of the Tetrahymena ribozyme reaction. The ~100-fold effect of docking the P1 duplex into its tertiary interactions with the ribozyme's core was demonstrated independently by the effect of the U to C change on miscleavage of 5′-exon analogs and by the effect of this change on the binding of oligonucleotides to the ribozyme relative to the effect on the stability of model P1 duplexes. The same effect on miscleavage was obtained upon removal of the U, suggesting that there is a favorable interaction with the G·U pair rather than an unfavorable interaction introduced with the G·C pair. Although some of the data used to determine the effect of the U to C change on the chemical step are complex, the same values are obtained from independent experimental approaches (see Figures 5 and 8, observed vs predicted values). These complexities have been attributed to binding of the C-containing substrates in the open complex and formation of a G·C^+ wobble pair (Chart 1), with several independent experiments providing strong support for these interpretations. The small discrepancies between observed and predicted values in Figures 5 and 8 of less than 2-fold may arise from experimental uncertainty, temperature effects, as experiments were performed at both 30 and 50 °C, and a small amount of interdependence of 2′-substituent effects. The 10-fold effect on the chemical step and a model to explain why this is 5-fold larger than the effect observed when there is a deoxyribose residue at position −1 are discussed below (see Contributions to the Chemical Step from a Directed Hydrogen Bond).

What Is the Role of the G·U Pair in Self-Splicing? It was previously shown that the G·U pair is important for 5′-splice site recognition in self-splicing and ribozyme reactions (Barfod & Cech, 1989; Doudna et al., 1989; Downs & Cech, 1994). The increased miscleavage observed herein upon replacement of the G·U pair with G·C and upon removal of the U residue supports the importance of the G·U in splice site selection. The reaction steps responsible for this contribution of the G·U to correct splice site selection have been identified (Figure 10). The G·U pair contributes ~100-fold to tertiary interactions of the P1 duplex with the active site, but contributes only when the P1 duplex is docked in the correct register, with the 5′-splice site aligned for reaction (e.g., Figure 2). This then provides a 100-fold preference for choosing the U of the G·U pair as the 5′-splice site. There is an additional contribution of ~10-fold to 5′-splice site selection in the chemical step, as the bond following the U of the G·U wobble pair is cleaved 10-fold faster than the

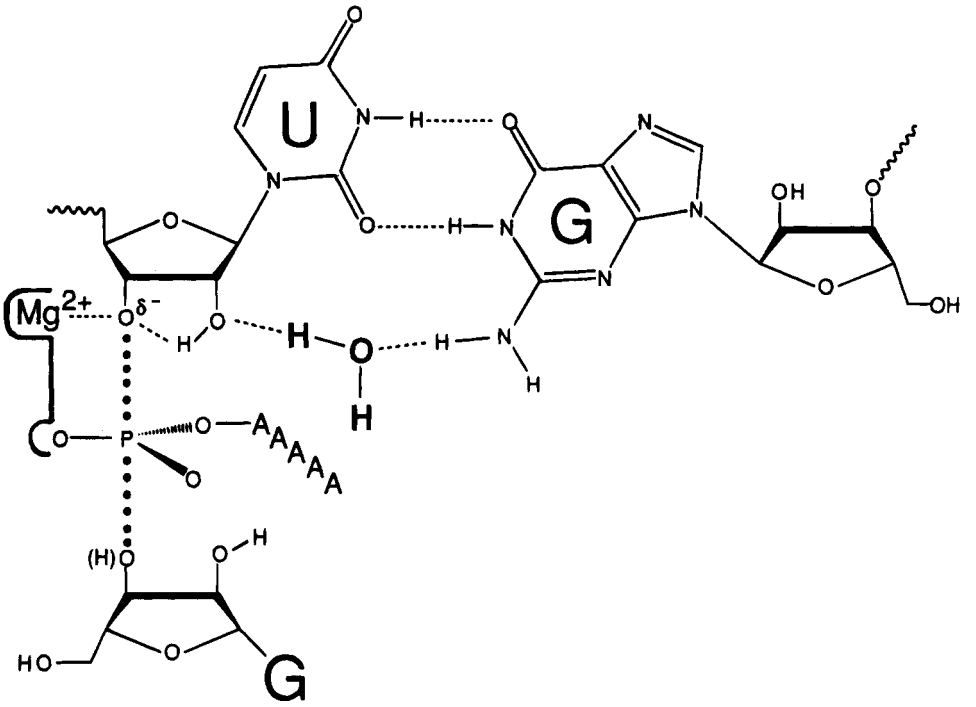


FIGURE 11: Transition state model for the effect of the G·U pair on the chemical step. A water molecule that bridges the exocyclic amino group of the G and the 2′-hydroxyl group of the U of the conserved G·U pair can orient the 2′-hydroxyl, thereby enhancing its ability to stabilize the neighboring incipient 3′-oxyanion in the transition state (Herschlag et al., 1993b). Evidence for an interaction of the 3′-oxygen atom with the Mg ion comes from changes in metal ion specificity upon replacement of the 3′-oxygen atom with sulfur (Piccirilli et al., 1993). The attack by G is shown as in-line because of the observed inversion of stereochemistry of the reactive phosphoryl group (McSwiggen & Cech, 1989; Rajagopal et al., 1989).

bond following the C of a G·C Watson–Crick pair.

The overall effect of ~1000-fold of the G·U pair on 5′-splice site selection is similar to the estimated fidelity of ~5000-fold for the L-21ScaI ribozyme reaction (Herschlag et al., 1993b). Had there been no other experimental data, we might have been tempted to conclude that the G·U pair was the only important recognition element in 5′-splice site selection. However, this is not the case. Previous work has indicated that the length of the P1 duplex, the J1/2 region that joins the P1 duplex to the remainder of the ribozyme, and residues at the base of the P2.1 duplex also have profound effects on the fidelity of 5′-splice site selection, and there is no reason to believe that all of the important positions have been identified (Doudna et al., 1989; Herschlag, 1992; Downs & Cech, 1994). Thus, there appear to be a number of interactions that are each important in aligning and docking the 5′-splice site. Taken together, the results suggest that mutations at positions not directly involved in docking of the G·U pair in the active site can nevertheless disrupt the alignment of the G·U pair with respect to its docking partner on the ribozyme, thereby compromising splicing fidelity [see also Downs and Cech (1994)].

It was previously suggested that the tertiary interactions with 2′-hydroxyl groups of the 5′-exon (and the IGS) may play a biological role in preventing the dissociation of the 5′-exon, as the 5′-exon is no longer covalently attached after the first step of splicing (Herschlag & Cech, 1990b). The G·U pair may aid splicing efficiency both by slowing dissociation of the 5′-exon via additional tertiary interactions and by speeding exon ligation via specific contributions in the chemical step.

It is interesting that the G·U pair, which is universally or nearly universally conserved, plays a smaller role in the chemical step than in binding and positioning. This is consistent with a biological requirement of the intron to turn over only once in self-splicing, but to carry out its single turnover with high accuracy and efficiency.

Contributions to the Chemical Step from a Directed Hydrogen Bond. The effect of the conserved G·U pair, relative to a G·C pair, is 5-fold greater with rU and rC than with dU and dC at this position (Figures 5 and 8). This indicates a cooperative energetic interaction of the G·U pair and the 2′-hydroxyl and is consistent with a model in which a water molecule (or 2′-hydroxyl group) bridges the exocyclic amino group of a G·U wobble pair and the 2′-hydroxyl group of the U residue (Figure 11). A bridging water molecule was observed in an X-ray crystallographic structure of a model duplex containing a G·U wobble pair (Holbrook et al., 1991) and can account for the rate enhancement as follows.

Previous analysis of a linear free energy relationship for the ribozyme reaction suggested that there is a development of negative charge on the 3′-oxygen of U(−1) in the transition state and that the neighboring 2′-hydroxyl group of this U residue helps to stabilize this incipient oxyanion via an intramolecular hydrogen bond (Herschlag et al., 1993b). Directing the 2′-hydroxyl group via a hydrogen bond with the bridging water would be expected to enhance the ability of the 2′-hydroxyl group to stabilize the incipient 3′-oxyanion by reducing the entropic barrier for formation of this hydrogen bond; the oriented 2′-hydroxyl can more efficiently donate a hydrogen bond to the neighboring 3′-oxygen.

The 5-fold synergistic effect from the G·U wobble pair and the 2′-hydroxyl group of U is similar to the entropy lost upon fixing a single bond rotation (Page & Jencks, 1971; Page, 1977). However, the entropy loss in hydrogen bond formation may be less than this value (Kirby, 1980). It is tempting to invoke polarization of the 2′-hydroxyl of the U

by the bound water molecule to render the 2′-hydroxyl a stronger hydrogen bond donor. Such a second-order effect on the hydrogen-bonding ability of the 2′-hydroxyl group may not differ significantly from the polarizing effect of a hydrogen bond donated from a water molecule in bulk solution. However, hydrogen bond donation from water may be inhibited by the 2′-hydrogen atom within the active site. In general, it is difficult to distinguish the removal of favorable interactions from the introduction of unfavorable interactions.

The 3-fold weaker binding of oligonucleotides to the ribozyme upon removal of the 2′-hydroxyl of U(−1) (i.e., upon changing the sugar residue from ribose to deoxyribose) is the same as the effect on the stability of model P1 duplexes [Herschlag et al. (1993b) and Results]. This is consistent with the model of Figure 11, as directing the hydrogen bond may provide a significant free energy contribution in the transition state, where a strong hydrogen bond can be formed with the partially anionic 3′-oxygen atom, but not in the ground state, where this bridging phosphoryl oxygen atom is not a strong hydrogen bond acceptor.

It is also of interest to note that substitution of the 2′-hydroxyl group of C in a G·C pair by a hydrogen atom destabilizes a model P1 duplex by ~2-fold (G. J. Narlikar and D. Herschlag, unpublished results), similar to the ~3-fold destabilization observed upon substitution of the 2′-hydroxyl group of U in a G·U pair of a model duplex. Thus, there is no indication of strong stabilization of the wobble pair from the bridging water molecule, even though it is observed as a bound water in the structure. This underscores the difficulty in using structural inspection alone to understand energetics.

Tertiary Interactions Direct the Formation of a G·C$^+$ Wobble Pair. The formation of a G·C$^+$ wobble pair with the same geometry as a G·U wobble (Chart 1) in the ribozyme active site accounts for several observations described in the Results. Indeed, if the ribozyme specifically recognizes a wobble pair at the 5′-splice site, then the ribozyme would be *expected* to drive the protonation of C and formation of a G·C$^+$ wobble pair, provided there are no major groove tertiary interactions with the U of the G·U wobble that would be lost with the G·C$^+$ wobble (see A Physical Picture of the G·U Interactions).

It is often assumed that RNA primary structure determines secondary structure, which in turn determines tertiary structure. This hierarchy sometimes can be observed experimentally, which is consistent with the high stability of RNA secondary structure [e.g., Banerjee et al. (1993); D. Herschlag, manuscript in preparation]. Computer algorithms to predict RNA secondary structure can sometimes give the observed RNA secondary structure, even though they cannot account for tertiary interactions due to the dearth of understanding of these interactions; this is also consistent with a hierarchy in which secondary structure determines tertiary structure.

In contrast, the results herein provide an example of RNA secondary structure being influenced by tertiary interactions. This might be likened to examples of identical protein sequences that have been found in either α-helices or β-sheets in different proteins (Cohen et al., 1993). It is not difficult to imagine an analogous RNA example in which tertiary interactions dictate the formation of one potential duplex rather than a competing duplex that would have been more stable in isolation. Indeed, it has been suggested that the tRNA tertiary structure can overcome the formation of an alternative non-cloverleaf secondary structure that is preferred under some conditions (Cole et al., 1972), and the formation of a base triplet in the core of the sunY group I intron appears to induce a nucleotide bulge in the secondary structure of mutant introns (Green & Szostak, 1994).

Given the ~1000-fold contribution of the phylogenetically conserved G·U pair in the Tetrahymena ribozyme reaction, it is surprising that even two of the >100 sequenced group I introns may lack this conserved pair and contain a G·C pair instead (Michel & Westhof, 1990). If further analysis substantiates the occurrence of these 5′-splice site G·C pairs, it will be interesting to qualitatively and quantitatively compare the strategies used by these introns in order to understand how G·C-containing introns can ensure accurate and efficient splicing.

A Physical Picture of the G·U Interactions. The ability of C to replace the U of the G·U pair at low pH, where a wobble pair can form (Chart 1), supports the previous conclusion that the G·U forms a wobble pair (Doudna et al., 1989; Green et al., 1991). In addition, oligonucleotide substrates containing dT at position −1 are as reactive as those with dU under conditions where the binding and chemical steps both contribute to reactivity (Herschlag et al., 1993b) and the G·C$^+$ pair is as reactive as the G·U pair. This absence of an effect from the addition of a methyl group to the C5 of U and replacement of the C4 carbonyl oxygen atom with an amino group suggests that there are no major groove interactions with the U residue of the G·U wobble.

Replacement of the 2′-hydroxyl of U with a 2′-methoxy substituent disrupts docking of the P1 duplex. As the replacement of this 2′-hydroxyl group with a hydrogen atom has no effect on docking, this disruption suggests that there is an unfavorable interaction between the methoxy substituent and a ribozyme functional group positioned in the minor groove of P1. A recent investigation of an extensive set of wobble pair analogs has indicated that the exocyclic amino group of the G specifically contributes to tertiary binding (S. A. Strobel and T. R. Cech, personal communication). Thus, the 2′-methoxy substituent may directly or indirectly interfere with the minor groove interaction with this amino group. Further, this clash indicates that there is a significant precision of packing within the active site. Finally, the ability of a 2′-methoxy substitution to have a profound energetic effect at a position that does not distinguish between a hydroxyl group and a hydrogen atom underscores the limited usefulness of the methoxy substituent in distinguishing whether a 2′-hydroxyl acts as a hydrogen bond donor or acceptor.

Promiscuous RNA/RNA Interactions. The similarity of base pairs within an RNA duplex can give rise to promiscuous RNA/RNA interactions. For example, the P1 duplex of the Tetrahymena ribozyme can dock into tertiary interactions in different registers, with different residues positioned in the cleavage site (Herschlag, 1992; Strobel & Cech, 1994). This misdocking leads to miscleavage and infidelity in 5′-splice site selection. The ribozyme distinguishes between these different binding registers in part by recognizing the difference between a G·U wobble pair and Watson−Crick pairs. We have added 2′-methoxy substituents at different positions along the P1 helix to further direct the docking register. This is possible because the ribozyme uses tertiary interactions with 2′-hydroxyl groups within the P1 duplex in docking. Thus, replacement of the 2′-hydroxyl group at

position −3, a normal position of tertiary interaction, with a 2′-methoxy residue hinders correct docking and enhances miscleavage (Results and G. J. Narlikar and D. Herschlag, manuscript in preparation). Conversely, substitution of methoxy groups for the 2′-hydroxyl groups at positions −6, −5, and −4, which do not normally participate in tertiary docking interactions, shuts down aberrant cleavage without affecting the normal reaction. The −3 methoxy-substituted oligonucleotides have been practically useful in providing a doorstop to prevent docking, thereby allowing the measurement of binding affinities in the open complex (Results and G. J. Narlikar and D. Herschlag, manuscript in preparation). The −6 to −4 methoxy-substituted oligonucleotides have been useful in shutting down miscleavage, thereby allowing measurement of cleavage reactions at the normal site that are very slow. Perhaps this is one of the reasons that nature has utilized modified sugars and bases in ribosomal and other RNAs. We expect that chimeric oligonucleotides will find analogous uses in the study of other ribozymes and RNAs.

ACKNOWLEDGMENT

We thank Mala Khosla for technical and other assistance throughout this project; Sylvester Muthini (Clontech), Al Smith, Elizabeth Zuo (PAN Facility, Beckman Center, Stanford), and Vidhya Gopalakrishnan for oligonucleotide synthesis; M. Chung and R. Waring for helpful discussions; Anna Pyle for the communication of results of equilibrium gel binding studies prior to the initiation of this work; T. Cech and S. Strobel for permission to cite unpublished work; and the Joyce lab for helpful comments on the manuscript.

REFERENCES

Banerjee, A. R., Jaeger, J. A., & Turner, D. H. (1993) *Biochemistry 32*, 153−163.

Barfod, E. T., & Cech, T. R. (1989) *Mol. Cell. Biol. 9*, 3657−3666.

Bevilacqua, P. C., & Turner, D. H. (1991) *Biochemistry 30*, 10632−10640.

Bevilacqua, P. C., Kierzek, R., Johnson, K. A., & Turner, D. H. (1992) *Science 258*, 1355−1358.

Cech, T. R. (1990) *Annu. Rev. Biochem. 59*, 543−568.

Cohen, B. I., Presnell, S. R., & Cohen, F. E. (1993) *Protein Sci. 2*, 2134−2145.

Cole, P. E., Yang, S. K., & Crothers, D. M. (1972) *Biochemistry 11*, 4358−4368.

Doudna, J. A., Cormack, B. P., & Szostak, J. W. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7402−7406.

Downs, W. D., & Cech, T. R. (1994) *Genes Dev. 8*, 1198−1211.

Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373−9377.

Good, N. E., Winget, D., Winter, W., Connolly, T. N., Izawa, S., & Singh, R. M. M. (1966) *Biochemistry 2*, 467−477.

Green, R., & Szostak, J. W. (1994) *J. Mol. Biol. 235*, 140−155.

Green, R., Szostak, J. W., Benner, S. A., Rich, A., & Usman, N. (1991) *Nucleic Acids Res. 19*, 4161−4166.

Gueffroy, D. E., Ed. (1978) *Buffers. A guide for the preparation and use of buffers in biological systems*, Calbiochem-Behring Corp., La Jolla, CA.

Herschlag, D. (1992) *Biochemistry 31*, 1386−1399.

Herschlag, D., & Cech, T. R. (1990a) *Biochemistry 29*, 10159−10171.

Herschlag, D., & Cech, T. R. (1990b) *Biochemistry 29*, 10172−10180.

Herschlag, D., & Khosla, M. (1994) *Biochemistry 33*, 5291−5297.

Herschlag, D., Piccirilli, J. A., & Cech, T. R. (1991) *Biochemistry 30*, 4844−54.

Herschlag, D., Eckstein, F., & Cech, T. R. (1993a) *Biochemistry 32*, 8299−8311.

Herschlag, D., Eckstein, F., & Cech, T. R. (1993b) *Biochemistry 32*, 8312−8321.

Holbrook, S. R., Cheong, C., Tinoco, I., & Kim, S. H. (1991) *Nature 353*, 579−581.

Hou, Y. M., & Schimmel, P. (1988) *Nature 333*, 140−145.

Jencks, W. P., & Regenstein, J. (1976) Ionization Constants of Acids and Bases, *Handbook of Biochemistry and Molecular Biology*, pp 305−351, CRC Press, Cleveland, OH.

Kirby, A. J. (1980) *Adv. Phys. Org. Chem. 17*, 183−278.

McConnell, T. S., Cech, T. R., & Herschlag, D. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8362−8366.

McSwiggen, J. A., & Cech, T. R. (1989) *Science 244*, 679−683.

Michel, F., & Westhof, E. (1990) *J. Mol. Biol. 216*, 585−610.

Michel, F., Hanna, M., Green, R., Bartel, D. P., & Szostak, J. W. (1989) *Nature 342*, 391−395.

Neddermann, P., & Jiricny, J. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 1642−1646.

Page, M. I. (1977) *Angew. Chem., Int. Ed. Engl. 16*, 449−459.

Page, M. I., & Jencks, W. P. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 1678−1683.

Piccirilli, J. A., Vyle, J. S., Caruthers, M. H., & Cech, T. R. (1993) *Nature 361*, 85−88.

Pyle, A. M., & Cech, T. R. (1991) *Nature 350*, 628−31.

Pyle, A. M., Murphy, F. L., & Cech, T. R. (1992) *Nature 358*, 123−128.

Rajagopal, J., Doudna, J. A., & Szostak, J. W. (1989) *Science 244*, 692−694.

Strobel, S. A., & Cech, T. R. (1993) *Biochemistry 32*, 13593−13604.

Strobel, S. A., & Cech, T. R. (1994) *Nature Struct. Biol. 1*, 13−17.

Zaug, A. J., Grosshans, C. A., & Cech, T. R. (1988) *Biochemistry 27*, 8924−8931.