Exocyclic Amine of the Conserved G·U Pair at the Cleavage Site of the *Tetrahymena* Ribozyme Contributes to 5′-Splice Site Selection and Transition State Stabilization†

Scott A. Strobel‡ and Thomas R. Cech*

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

Received September 19, 1995; Revised Manuscript Received November 27, 1995⊗

ABSTRACT: A phylogenetically conserved guanine·uracil (G·U) pair defines the 5′-exon/intron boundary of precursor RNAs containing group I introns. In this wobble base pair, the G forms two hydrogen bonds with U in a base pairing geometry shifted from that of a canonical Watson−Crick pair. On the basis of thermodynamic measurements of synthetic base pair analogs (inosine, diaminopurine riboside, guanosine, or adenosine paired with U, C, or isocytidine) in place of the G·U pair, we have previously reported that the N_2 exocyclic amine of the G is important for docking the 5′-exon into the active site of the *Tetrahymena* ribozyme [Strobel, S. A., & Cech, T. R. (1995) *Science 267*, 675−679]. Here we describe kinetic characterization of ribozyme−substrate combinations containing the same series of analogs. By measuring the rate constants of 5′-exon miscleavage (cleavage at incorrect phosphates), we demonstrate that the 5′-exon/intron boundary is primarily defined by the exocyclic amine of the G. The amine makes its contribution (2.5 kcal·mol^{-1}) in the context of all three wobble pairs tested but fails to make a significant contribution (<0.8 kcal·mol^{-1}) when presented in a Watson−Crick base pairing geometry. We also demonstrate that the exocyclic amine makes a modest contribution to chemical transition state stabilization (1.0 kcal·mol^{-1} relative to an inosine−U pair). The majority of this transition state contribution (0.7 kcal·mol^{-1}) is independent of that contributed by the 2′-hydroxyl of the neighboring U. This argues against the model in which substantial transition state stabilization is derived from a water molecule bridging between the exocyclic amine of G and the 2′-hydroxyl of U. Instead it suggests that the tertiary interaction between the exocyclic amine and its hydrogen bonding partner in the active site is slightly improved during the chemical transition. We conclude that the exocyclic amine of G is the primary contributor to many characteristics of reactivity that have been ascribed to the conserved G·U pair, including stabilization of the chemical transition state and definition of the 5′-exon/intron boundary.

The *Tetrahymena* group I intron accurately defines a single phosphate as the 5′-exon/intron boundary and catalyzes a transesterification reaction at that phosphate with a rate enhancement of 10^{11} (Cech, 1987). The basis for the reaction specificity is defined in part by the conserved secondary structure of the ribozyme. The 3′-terminus of the 5′-exon base pairs to a complementary sequence within the intron (the internal guide sequence or IGS) to form the splice site helix (also called P1 helix) (Been & Cech, 1986; Waring et al., 1986; Zaug et al., 1986; Figure 1). A G·U wobble base pair formed between a U at the 3′-end of the exon and a G located in the IGS is phylogenetically conserved at the exon/intron boundary (Cech, 1988; Michel & Westhof, 1990). The site of reactivity can be moved one or two positions on either side of the normal 5′-splice site if the G·U is moved to these positions (Doudna et al., 1989). This demonstrates that the G·U serves as a primary determinant of 5′-splice site selection.

Mutagenesis of the G·U pair to any other base pair significantly reduces the enzymatic activity (Doudna et al., 1989). Among all the natural base combinations the A·C pair retains the greatest reactivity, though it was not clear which step in the reaction pathway was rate-limiting under the conditions studied. Because A and C can base pair in a wobble configuration, this suggests the conformation of the base pair might be important for reaction site recognition and/or catalysis during the first step of splicing. Among the base pairs, only the G·U to G·C mutation has been carefully characterized (Pyle et al., 1994; Knitt et al., 1994). This mutation, which converts the wobble into a Watson−Crick pair, substantially impairs the ribozyme's ability to define the exon/intron boundary as judged by increased cleavage at incorrect sites. It also reduces the ribozyme's reactivity at the proper site by two orders of magnitude. In contrast, the requirement for a G·U pair at the reaction site is relaxed for the second step of splicing (Barfod & Cech, 1989).

The splice site helix is held into the ribozyme catalytic core by at least four tertiary interactions between the RNA duplex and conserved sequence elements within the ribozyme active site. These include sequence nonspecific contacts to three 2′-hydroxyls distributed on both strands of the splice site helix (Bevilacqua & Turner, 1991; Pyle & Cech, 1991; Strobel & Cech, 1993) and a sequence-specific tertiary contact to the exocyclic amine of the conserved G·U pair

† S.A S was a fellow of the Life Sciences Research Foundation sponsored by the Howard Hughes Medical Institute. T.R.C. is an Investigator of the Howard Hughes Medical Institute and an American Cancer Society Professor. We thank the W. M. Keck Foundation for the support of RNA science on the Boulder campus.

‡ Present address: Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520-8024.

⊗ Abstract published in *Advance ACS Abstracts,* January 15, 1996.

0006-2960/96/0435-1201$12.00/0

(Strobel & Cech, 1995). These functional groups are thought to hydrogen bond to groups in the core of the ribozyme (Pyle et al., 1992; Michel & Westhof, 1994). The thermodynamic role of the exocyclic amine in helix docking was defined using a series of natural and unnatural base pair mutations for the G·U pair that systematically altered each of the functional groups (Strobel et al., 1994). These included wobble base pairs I·U (inosine·U), A·C, A·^{Me}iC[1] (A·5-methyl-isocytidine), DAP·C (diaminopurine riboside·C), and DAP·^{Me}iC (Figure 2). They also included the Watson−Crick pairs G-C, I-C, DAP-U, and A-U (Figure 3). By systematically changing each functional group in the context of a wobble base pair configuration, the role of the exocyclic amine in helix docking was quantitated.

While the previous work focused on the thermodynamics of 5′-exon binding, here we characterize the kinetics of each of the G·U base pair mutations. We find that, in addition to its role in helix docking, the exocyclic amine of G is the primary contributor to many characteristics of reactivity that have been ascribed to the G·U pair, including stabilization of the chemical transition state and definition of the 5′-exon/intron boundary.

MATERIALS AND METHODS

Ribozymes and Substrates. The L-21 *ScaI* form of the *Tetrahymena* ribozyme (Zaug et al., 1988) with nucleotide substitutions X (Figure 1) in place of the conserved G at position 22 were prepared semisynthetically using T4 DNA ligase (Moore & Sharp, 1992) as described (Strobel & Cech, 1993, 1995). In this study X represents guanosine (G), inosine (I), adenosine (A), or diaminopurine riboside (2-aminoadenosine, DAP). The inosine phosphoramidite was purchased from ChemGenes, the diaminopurine riboside (2-aminoadenosine) phosphoramidite was synthesized as described (Strobel et al., 1994), and the N-acetyl protected phosphoramidite of C was purchased from Pharmacia.

Oligonucleotide substrates Y(−1)S of sequence 5′-GGCCCUCYAAAAA-3′ and products Y(−1)P of sequence 5′-GGCCYUCY-3′ were prepared by automated oligonucleotide synthesis. Y represents cytidine (C), uridine (U), or 5-methylisocytidine (^{Me}iC).

Synthesis of 5-Methylisocytidine Phosphoramidite. Preparation of the ^{Me}iC(−1)S oligonucleotide required synthesis of the ^{Me}iC phosphoramidite. Previous efforts to synthesize the phosphoramidite were unsuccessful because the 2′-silylation product of the *N*-benzoyl protected ^{Me}iC base could not be separated from the 3′ regioisomer (Strobel et al., 1994). Additionally, about half of the *N*-benzoyl-^{Me}iC underwent transamination to *N*-methyl-^{Me}iC when the oligonucleotide was deprotected with methyl amine in ammonium hydroxide (AMA). These difficulties were overcome using an acetyl protecting group for the amine.

[1] Abbreviations: ^{Me}iC, 5-methylisocytidine; DAP, 2,6-diaminopurine riboside (synonymous with 2-aminoadenosine); I, inosine; DMT, 4,4′-dimethoxytrityl; DMTCl, 4,4′-dimethoxytrityl chloride; TBDMS, *tert*-butyldimethylsilyl; TBDMSCl, *tert*-butyldimethylsilyl chloride; AMA, an equal mixture of 40% $MeNH_{2(aq)}$ and $NH_4OH_{(conc)}$; PAGE, polyacrylamide gel electrophoresis; IGS, internal guide sequence of the *Tetrahymena* ribozyme; GMP, guanosine 5′-monophosphate; *X22*, *Tetrahymena* L-21 *Sca*I ribozymes with nucleoside X (X = G, I, A, or DAP) at position 22; Y(−1)S, oligonucleotide substrates (5′-GGCCCUCYAAAAA-3′) for ribozyme-catalyzed cleavage (Y = U, C, or ^{Me}iC); $\overline{Y}$(−1)P, oligonucleotide reaction products analogous to the 5′-exon ($\overline{5'}$-GGCCCUCY-3′); P″, major product of ribozyme infidelity (5′-GGCCCU-3′); GMP or pG, 5′-guanosine monophosphate.

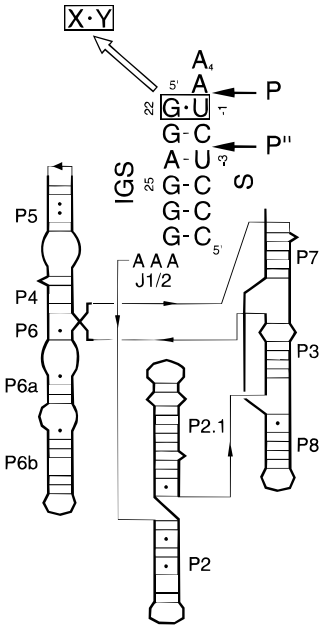

FIGURE 1: Diagram of the 5′-splice site helix and the catalytic core of the group I intron [modified from Salvo and Belfort (1992) and Cech et al. (1994)]. The splice site helix is shown in capital letters with the G·U pair changed to X·Y. The numbering of specific bases is shown on both sides of the helix. The sites of the primary cleavage, giving product P, and of miscleavage, giving product P″, are shown with arrows to the right of the helix. Other structural elements that contribute to 5′-splice site selection including J1/2 and P2/P2.1 are also shown. The P5abc extension contiguous with P5, and the P9 extension following P7 are present in these experiments, but they are not depicted in the figure. Helices P3−P8 constitute the catalytic core of the ribozyme.

The 5-methyl derivative of isocytidine was targeted for synthesis because in the DNA series hydrolytic deamination of isocytidine to 2′-deoxyuridine was reported to be a major side product upon base deprotection with concentrated ammonia (Switzer et al., 1993). However, substitution of a methyl group at the 5 position was reported to significantly reduce the deamination side reaction (Tor & Dervan, 1993). Furthermore, studies with thymidine (5-methyl 2′-deoxyuridine) showed that the methyl group had no detrimental effect upon ribozyme reactivity (Herschlag et al., 1993).

The ^{Me}iC phosphoramidite was synthesized following the general procedures of Usman et al. (1987) and Scaringe et al. (1990) (Figure 4). Synthesis was carried out using reagent grade chemicals as received unless otherwise stated. Dry solvents were purchased in sure seal containers under N_2. 5-Methyl-isocytosine was synthesized from tri-*O*-benzoyl-1-*O*-acetyl-D-ribofuranose and thymine as previously described (Kimura et al., 1980). NMR spectra were recorded on a Varian XL-400 instrument operating at 400 MHz (^{1}H) and 121 MHz (^{31}P). Chemical shifts are reported in ppm relative to the solvent residual signal. FAB mass spectra were recorded on a VG 7070 EQ-HF mass spectrometer. Thin layer chromatography (TLC) was performed on silica gel 60 F-254 precoated plates. Flash column chromatography was carried out using silica gel grade 62 (60−200 mesh).

*N^2-Acetyl-5-methylriboisocytidine (**1**).* A cold suspension of 5-methylisocytidine (1.94 g, 7.6 mmol) in dry pyridine (100 mL) was treated dropwise with acetic anhydride (8.6 mL, 91.2 mmol, 12 equiv). The reaction mixture was allowed to warm to room temperature and stirred for 3 h. The reaction was then chilled, quenched by the addition of MeOH (20 mL), and concentrated to an oil. The oil was

redissolved in CH_2Cl_2 (100 mL), washed with $NaHCO_3$ (2 × 50 mL) and water (2 × 50 mL), dried (Na_2SO_4), and concentrated *in vacuo*. The residue was redissolved in MeOH (50 mL) and reacted with NaOMe (3.69 g, 68.4 mmol, 9 equiv) at room temperature for 5 min. The reaction was quenched with Dowex resin (50 × 8−200 Py^+ form), filtered, washed with pyridine (100 mL) and methanol (50 mL), and concentrated to an oil. The residue was purified by flash column chromatography (5−10% MeOH/CH_2Cl_2) to afford 1.92 g (85% yield) of nucleoside **1** as a white foam. TLC (5% MeOH/CH_2Cl_2) $R_f = 0.09$. ^{1}H NMR (DMSO-d_6) δ 13.16 (s, 1H, NHCO, D_2O ex), 8.09 (s, 1H, H6), 6.20 (d, 1H, H1′), 5.39 (d, 1H, OH2′, D_2O ex), 5.22 (t, 1H, OH5′, D_2O ex), 5.09 (d, 1H, OH3′, D_2O ex), 4.07 (dd, 1H, H2′), 4.00 (dd, 1H, H3′), 3.88 (unresolved m, 1H, H4′), 3.72−3.57 (m, 2H, H5′), 2.08 (s, 3H, $COCH_3$) 1.84 (s, 3H, CH_3). MS (positive-ion FAB) calcd for $C_{12}N_3O_6H_{17}$ $[MH]^+$ 300, found 300.

*5′-O-(4,4′-Dimethoxytrityl)-N²-acetyl-5-methylriboisocytidine (**2**).* Nucleoside **1** (1.92 g, 6.4 mmol) was dissolved in dry pyridine (50 mL), and DMTCl (2.4 g, 7.1 mmol, 1.1 equiv) was added to the reaction in three equal portions over a period of 90 min. The reaction was greater than 90% complete after stirring overnight. The unreacted DMTCl was quenched by the addition of MeOH (10 mL), and the solution was concentrated to an oil. The reaction was dissolved in CH_2Cl_2, extracted twice with $NaHCO_3$ (sat) and once with water, dried (Na_2SO_4), and reconcentrated. Purification of the oil by flash column chromatography (0−3% MeOH/CH_2Cl_2) resulted in 2.8 g (73% yield) of nucleoside **2** as a light yellow foam. TLC (5% MeOH/CH_2Cl_2) $R_f = 0.45$. ^{1}H NMR (DMSO-d_6) δ 13.13 (s, 1H, NHCO), 7.72 (s, 1H, H6), 7.39−7.15 (m, 9H, DMT), 6.89 (d, 4H, DMT), 6.20 (s, 1H, H1′), 5.51 (d, 1H, OH2′, D_2O ex), 5.18 (d, 1H, OH3′, D_2O ex), 4.22 (t, 1H, H2′), 4.14 (dd, 1H, H3′), 4.02 (unresolved m, 1H, H4′), 3.73 (s, 6H, OCH_3), 3.31−3.22 (m, 2H, H5′), 2.08 (s, 3H, $COCH_3$), 1.43 (s, 3H, CH_3). MS (positive-ion FAB) calcd for $C_{33}N_3O_8H_{35}$ $[MH]^+$ 602, found 602.

*5′-O-(4,4′-Dimethoxytrityl)-2′-O-(tert-butyldimethylsilyl)-N²-acetyl-5-methylriboisocytidine (**3**).* Compound **2** (2.4 g, 4.0 mmol) was coevaporated with pyridine (3 × 50 mL), dissolved in dry THF (40 mL), and combined with $AgNO_3$ (0.81 g, 4.8 mmol, 1.2 equiv) and TBDMSCl (0.78 g, 5.2 mmol, 1.3 equiv), whereupon the solution became instantly turbid. No product was observed after reacting overnight. Triethylamine (2.8 mL, 20 mmol, 5.0 equiv) and TBDMS trifluoromethanesulfonate (1.0 mL, 4.0 mmol, 1.0 equiv) were added to the reaction mixture and stirred an additional 8 h. After the addition of TBDMS trifluoromethanesulfonate, more than 70% of **2** reacted, but the reaction proceeded without regioselectivity. The precipitate was removed by filtration through Celite into a stirring solution of $NaHCO_3$ (sat), extracted into CH_2Cl_2, and the organic phase washed with $NaHCO_3$ (2 × 50 mL), brine (2 × 50 mL), and water (2 × 50 mL), dried (Na_2SO_4), and concentrated. The resulting oil was purified by flash column chromatography (20−25% EtOAc in hexane) to afford 0.63 g (22% yield) of **3** as a white foam. Unlike the benzoyl protected compound (Strobel et al., 1994), the acetyl protected regioisomers were readily separated by chromatography. 2′−3′ disilylated product TLC (1:1 EtOAc/hexane) $R_f = 0.72$, 2′-silylated product $R_f = 0.62$, 3′-silylated product $R_f = 0.57$. ^{1}H NMR (DMSO-d_6) δ 13.20 (s, 1H, NHCO), 7.76 (s, 1H, H6), 7.40−7.24 (m, 9H, DMT), 6.89 (d, 4H, DMT), 6.28 (s, 1H, H1′), 5.21 (d, 1H, OH3′, D_2O ex), 4.30 (t, 1H, H2′), 4.13 (dd, 1H, H3′), 4.06 (unresolved m, 1H, H4′), 3.73 (s, 6H, OCH_3), 3.34−3.25 (m, 2H, H5′), 2.07 (s, 3H, $COCH_3$), 1.39 (s, 3H, CH_3), 0.85 (s, 9H, *tert*-butyl), 0.06 (s, 3H, $SiCH_3$), 0.04 (s, 3H, $SiCH_3$). MS (negative-ion FAB) calcd for $C_{39}N_3O_8H_{49}Si$ $[M]^-$ 714, found 714.

*5′-O-(4,4′-Dimethoxytrityl)-2′-O-(tert-butyldimethylsilyl)-N²-acetyl-5-methylriboisocytidine-3′-O-β-cyanoethyldiisopropylaminophosphoramidite (**4**).* Nucleoside **3** (0.63 g, 0.88 mmol) was dissolved in cold CH_2Cl_2 (10 mL). In a separate flask, 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.30 mL, 1.32 mmol, 1.5 equiv), diisopropylethylamine (0.31 mL, 1.76 mmol, 2.0 equiv), and *N*-methyl imidazole (0.035 mL, 0.44 mmol, 0.5 equiv) were added to an additional portion of dry CH_2Cl_2 (10 mL), cooled to 0 °C, and added dropwise to the solution of nucleoside **3**. The reaction was warmed slowly to room temperature and stirred for 1 h. NMR of the H1′ proton showed complete conversion to phosphoramidite **4**. The reaction mixture was concentrated and the residue purified by flash column chromatography (1% triethylamine in 1:1 EtOAc/hexane) to afford 0.70 g (87% yield) of phosphoramidite **4** as a white foam. TLC (1:1 EtOAc/hexane) of both regioisomers same as starting material, $R_f = 0.62$. ^{31}P NMR ($CDCl_3$) δ 152.75 and 150.91 (two phosphoramidite diastereomers).

Oligonucleotide Synthesis, Deprotection, and Characterization. Oligonucleotides were synthesized on an Applied Biosystems 394 synthesizer. The coupling efficiency of the ^{Me}iC phosphoramidite was comparable to that obtained with commercially available phosphoramidites (98−100%). All oligonucleotides were prepared using *N*-acetyl C phosphoramidite and deprotected using 1.5 mL of AMA at 70 °C for 30 min (Reddy et al., 1994; Wincott et al., 1995). Oligonucleotide purification and characterization followed the procedure previously described (Strobel et al., 1994). Base composition analysis was performed on a model ^{Me}iC oligonucleotide. No transamination to *N*-methyl-^{Me}iC was detected-and less than 3% of the nucleoside underwent hydrolytic deamination to 5-methyluridine.

Oligonucleotide substrates with a phosphorothioate substitution at the scissile phosphate (CCCUC$\underline{Y}_{Rp}$A) were synthesized using Beaucage reagent (Glen Research) in place of iodine for the oxidation step during coupling of phosphoramidite Y, where Y is U, C, or ^{Me}iC. This produced a diastereomeric mixture of R_P and S_P isomers. The R_P isomer was purified by semipreparative reverse-phase HPLC as described (Slim & Gait, 1991).

Kinetic Measurements of Ribozyme-Catalyzed Cleavage Reactions. Kinetics methods followed those previously described (Herschlag & Cech, 1990; Herschlag et al., 1991; McConnell et al., 1993). The L-21 *Sca*I ribozyme or its variants (abbreviated *X22* where X is G, A, I, or DAP) were incubated in reaction buffer (50 mM HEPES pH 7.0, 10 mM $MgCl_2$) at 50 °C for 20 min to create a homogeneously folded ribozyme population prior to cooling to 30 °C for reaction. Reactions were initiated by the addition of oligonucleotide substrate and GMP which had been previously equilibrated to 30 °C in reaction buffer. Portions of the reaction were removed at various times and quenched on ice with 2 volumes of 7.5 M urea, 20 mM EDTA, 0.1% xylene cyanol, and 0.1% bromophenol blue. Reaction products were resolved from substrate by denaturing 20% PAGE, and the fraction reacted was quantitated with a PhosphorImager (Molecular Dynamics).

Cleavage of Y(−1)P (GGCCCUCY) at −3U to produce P″ (GGCCCU) was assayed under single-turnover conditions with saturating ribozyme and GMP (Herschlag, 1992). Rate constants for P″ production ($k_{cat}^{P''}$) were measured at 30 °C with 5 mM GMP using a trace amount (≈0.2 nM) of ^{32}P-radiolabeled Y(−1)P. The ribozyme concentration was at least 50-fold above the equilibrium dissociation constant for product binding (Strobel & Cech, 1995). These concentrations were as follows: 100 nM for G·U, I·U, DAP·C, DAP·^{Me}iC, and G-C; 500 nM for A·^{Me}iC; 1 μM for I-C, DAP-U, and A-U; and 1.5 μM for A·C. The ribozyme and GMP concentrations were doubled with no appreciable increase in the reaction rate, demonstrating that in all cases the reactions were saturated with respect to ribozyme and GMP concentration. At 30 °C, very little cleavage at −2C to produce the P′ product (GGCCCUC) was detected for any of the ribozyme/product combinations. Therefore, reported values reflect the rate constant for conversion of Y(−1)P into P″. For ribozyme/product combinations that reacted very slowly ($<10^{-4}$ min^{-1}), the reported values reflect the rate of the first 10−20% of the reaction. For faster reacting ribozyme/product combinations ($>10^{-4}$ min^{-1}), the reaction was monitored through at least four half-lives and obeyed first-order kinetics.

RESULTS

The Exocyclic Amine of the Conserved G·U Pair Helps Define the 5′-Exon/Intron Boundary. Infidelity has been defined as the propensity of the ribozyme to cut at sites other than −1U, the proper cleavage site of the first step in splicing (Herschlag, 1992). Previous studies have demonstrated that disruption of 2′-hydroxyls important for docking of the splice site helix into the ribozyme active site results in a 30−80-fold higher rate of miscleavage (Herschlag, 1992; Strobel & Cech, 1994). Infidelity is caused by translocation of the splice site helix into alternative binding registers. This places alternative phosphodiesters into the ribozyme active site, resulting in miscleavage. The rates of infidelity provide a measure of a mutant's ability to properly define the 5′-exon/intron boundary for the first step of splicing.

Rates of cleavage infidelity were first measured at 30 °C for the ribozyme containing the G·U pair at the cleavage site (Table 1). *G22* miscuts the U(−1)P oligonucleotide (the wild-type analog of the 5′-exon) at −3U to produce the P″ product (GGCCCU) at a rate of 1.4×10^{-5} min^{-1}. When compared to the maximal rate of cleavage at the G·U pair under these conditions (13 min^{-1}), this rate reflects approximately a million-fold (10^6) preference for the 5′-exon/intron boundary over any other site. This preference is sufficient to ensure that all of the 5′-exons are properly processed in the first step of splicing.

To characterize the molecular basis by which the G·U pair contributes to cleavage site fidelity, the functional groups of the G·U pair were systematically altered or deleted in the context of a wobble base pair configuration. For example, the I·U pair retains a wobble base pairing configuration but lacks the exocyclic amine normally present in the G·U pair (Figure 2). *I22* miscuts U(−1)P to produce P″ 60-fold faster than *G22* (Table 1). This strongly suggests that the ribozyme uses the exocyclic amine as a molecular handle to properly define the exon/intron boundary.

The A·C wobble pair has been reported to have the highest level of reactivity among all possible natural nucleotide pairing combinations (Doudna et al., 1989). Although the A·C can adopt a wobble pairing conformation, the functional groups that comprise the pair are substantially different than those of the G·U pair (Figure 2). *A22* miscuts C(−1)P 200-fold faster than the rate for *G22* cleavage of U(−1)P (Table 1). While this points to something being fundamentally amiss with the A·C pair, its rate of infidelity is only 4-fold greater than that of the I·U pair. Similarly, a ribozyme−substrate complex containing an A·^{Me}iC pair miscut 170-fold faster than for a G·U pair, but within 3-fold of the rate of the I·U pair. This indirectly implies that, despite their differences, I·U, A·C, and A·^{Me}iC pairs lack the same critical functional group, possibly the N_2 exocyclic amine of the purine base.

To test this possibility, we prepared a ribozyme with diaminopurine riboside (DAP) in place of the conserved G at position 22. Like A, DAP forms a wobble pair with C and ^{Me}iC, but, unlike A, it also has the N_2 exocyclic amine present in the original G·U wobble pair (Figure 2). If the exocyclic amine is the critical functional group for defining the exon/intron boundary, then DAP·C and DAP·^{Me}iC should restore cleavage fidelity approximately 60-fold compared to the rates for A·C and A·^{Me}iC, respectively. The difference in fidelity between DAP·C and A·C or DAP·^{Me}iC and A·^{Me}iC should be the same as the difference between G·U and I·U (Figure 5).

Miscleavage of C(−1)P by ribozyme *DAP22* was 50-fold slower than the A·C pair and within 5-fold of the original G·U rate (Table 1). *DAP22* miscut ^{Me}iC(−1)P at a slightly faster rate. While it is 10-fold more accurate than the A·^{Me}iC pair, it remains almost 20-fold less accurate than the rate for the G·U pair (Table 1). Although the restoration of reaction fidelity was incomplete in this case, it is clear that the exocyclic amine makes a significant contribution in the context of three different wobble pairs (Figure 5), providing strong evidence that, among the functional groups of the G·U

Table 1: Cleavage Infidelity for G·U Wobble Pair Substitutions (pH 7.0, 30 °C)[a]

base pair (X·Y)	infidelity k_{cat} (10^{-5} min^{-1})	infidelity $\Delta\Delta G^{\ddagger}$ [b] (kcal·mol^{-1})	tertiary $\Delta\Delta G^{\circ}$ [c] (kcal·mol^{-1})
Wobble Pairs with Exocyclic Amine			
G·U	1.4 ± 0.1	(0)	(0)
DAP·C	6.4 ± 2.1	0.9	1.3
DAP·^{Me}iC	25 ± 3	1.7	1.1
Wobble Pairs without Exocyclic Amine			
I·U	86 ± 19	2.5	2.0
A·C	310 ± 80	3.2	3.0
A·^{Me}iC	240 ± 10	3.1	2.8
Watson−Crick Pairs with Exocyclic Amine			
G-C[d]	150 ± 20	2.8	2.6
DAP-U	270 ± 90	3.2	3.8
Watson−Crick Pairs without Exocyclic Amine			
I-C	520 ± 90	3.5	3.4
A-U	430 ± 120	3.4	3.4

[a] The rate constants reported in this table are for the reaction: GGCCCUCY + pG → GGCCCU + pGCY. The substrate, GGCCCUCY, in this miscleavage reaction is actually the product of reaction at the proper cleavage site. Because the scissile phosphate is omitted, the substrate for this reaction cannot be cleaved at the normal splice site. [b] Free energy loss for cleavage infidelity calculated relative to the infidelity rate constant for the G·U pair. Values are calculated as $RT \ln(k_{cat}^{XY}/k_{cat}^{GU})$, where R and T equal 0.00198 kcal·mol^{-1}·K^{-1} and 303.15 K, respectively. Positive numbers indicate greater infidelity than the G·U pair. [c] From Table 1 of Strobel and Cech (1995). [d] From Table 3 of Pyle et al. (1994).

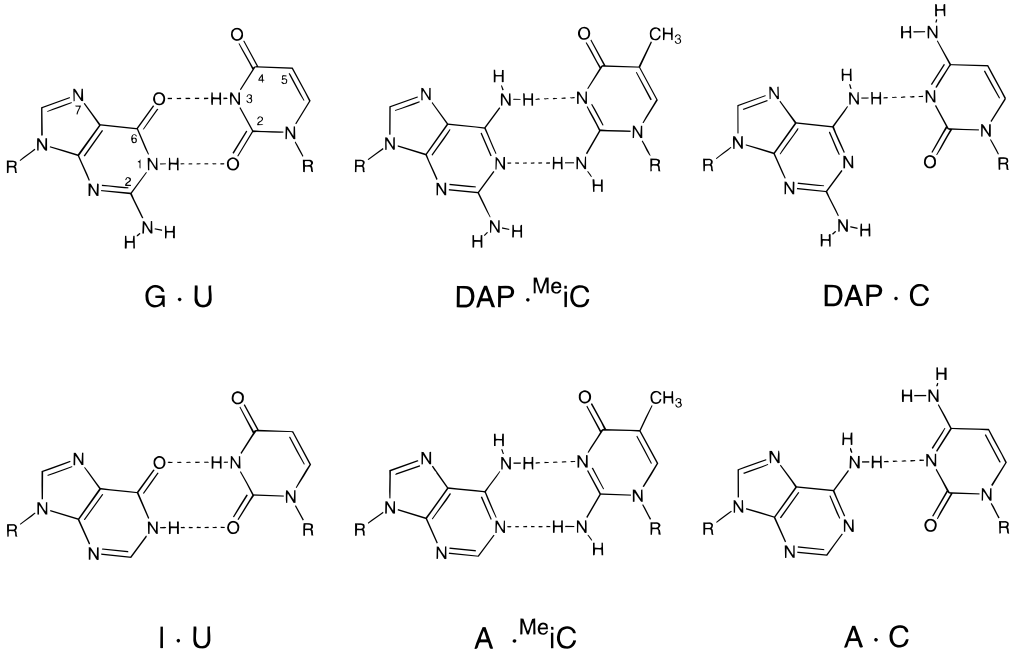

FIGURE 2: Proposed hydrogen bonding of bases in a wobble pairing configuration.

pair, the exocyclic amine is the primary determinant of 5′-splice site selection.

The disparity between the rates for the DAP·^{Me}iC and G·U pairs could be interpreted to mean the O_2 carbonyl of the pyrimidine also makes a positive contribution. However, in the absence of the amine (I·U compared to A· ^{Me}iC) the rates are very similar. This suggests that, instead of the O_2 carbonyl of U making a positive contribution, the N_2 amine of ^{Me}iC is making an indirect negative contribution when the exocyclic amine of DAP is present. The presentation of the exocyclic amine by the DAP·^{Me}iC pair could be distorted because of unfavorable electrostatic interactions between the three adjacent amines of DAP·^{Me}iC (Figure 2).

The Exocyclic Amine Does Not Contribute in a Watson−Crick Context. Previous work has shown that a G-C pair at the cleavage site substantially impairs the ability of the ribozyme to define the exon/intron boundary (Pyle et al., 1994; Knitt et al., 1994). At 30 °C *G22* miscuts C(−1)P more than 100-fold faster than it miscuts U(−1)P (Table 1). A comparably high rate of infidelity is observed for the cleavage of U(−1)P by *DAP22* (Table 1). Although both the G-C and the DAP-U pairs have the critical exocyclic amine, both pairs are in a Watson−Crick configuration, and the amine is locked into a hydrogen bonding register with the O_2 carbonyl of the opposing pyrimidine base (Figure 3).

On the basis of the high rate of miscleavage, it appears the amine is not making a major contribution to splice site selection when presented by a Watson−Crick pair. To test this possibility, the exocyclic amine was deleted in the context of both Watson−Crick pairs through the use of I-C and A-U base pairs (Figure 3). *I22* miscleaves C(−1)P only 3-fold faster than does *G22*. *A22* miscuts U(−1)P less than 2-fold faster than does *DAP22*. This confirms that the exocyclic amine is not contributing significantly to 5′-splice site selection in a Watson−Crick base pair (Figure 5). The mere presence of the exocyclic amine is insufficient. To make its contribution to 5′-splice site selection, it must be presented in a wobble environment.

Correlation of Infidelity Rate Constants and Equilibrium Binding Constants. Infidelity rate constants were also used

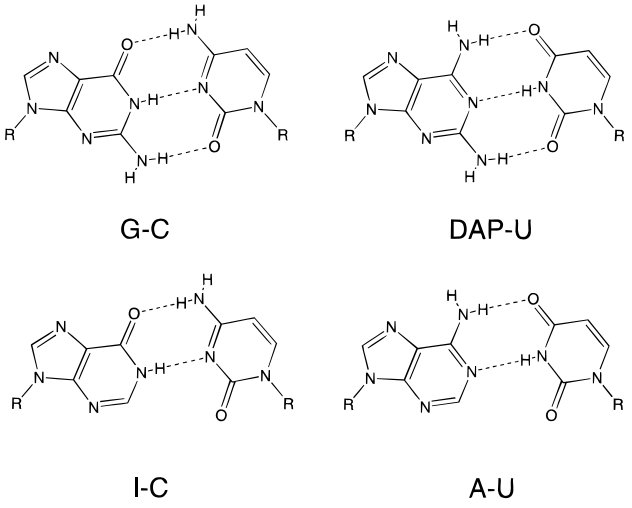

FIGURE 3: Proposed hydrogen bonding of bases in a Watson−Crick pairing configuration.

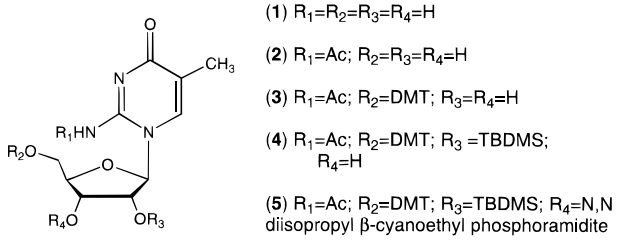

FIGURE 4: Intermediates in the synthesis of 5-methylisocytidine (^{Me}iC) phosphoramidite.

to calculate the difference in free energy of activation ($\Delta\Delta G^{\ddagger}$) relative to the G·U pair (Table 1). These values show good correlation with the $\Delta\Delta G°$ attributed to loss of tertiary binding energy calculated from equilibrium binding constants that were adjusted for differences in duplex stability (Strobel & Cech, 1995). The tertiary $\Delta\Delta G°$ reflects the ground state occupancy (the equilibrium between the docked and undocked states) of the splice site helix in the active site of the ribozyme. Strong agreement between the $\Delta\Delta G$'s for tertiary binding and splice site infidelity argues that the infidelity rates also measure ground state occupancy of the

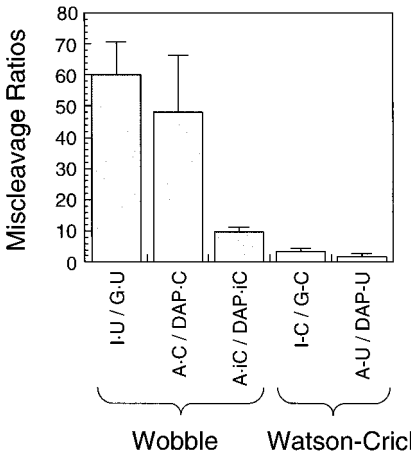

FIGURE 5: Relative contribution of the exocyclic amine to cleavage fidelity for three wobble and two Watson−Crick pairs. Values reflect the relative increase in cleavage infidelity upon deletion of the exocyclic amine for each base pair.

splice site helix and that misdocking, not cleavage at −3U, is the rate-limiting step for miscleavage (Herschlag, 1992).

The Exocyclic Amine also Makes a Modest Contribution to Chemical Transition State Stabilization. In addition to infidelity rates, each of the ribozyme/substrate combinations was kinetically characterized for reactivity at the proper cleavage site. Under single-turnover conditions and saturating ribozyme concentration, $(k_{cat}/K_m)^G$ represents the second-order rate constant for the reaction of E_S (substrate GGCCCUCYAAAAA fully bound to the ribozyme, but the G site unoccupied) with free G. This reaction is analogous to the first step of splicing. The substrate is cleaved at −1Y to produce the product GGCCCUCY. The kinetic constants, k_{cat} and K_m^G can be individually determined by measuring reaction rates from low to saturating G concentrations.

These numbers provide important information about the ground and transition states of the ribozyme. K_m^G is sensitive to the ground state docking of the substrate helix into the active site (McConnell et al., 1993; Bevilacqua et al., 1993). In an undocked conformation, K_m^G is approximately equal to 320 μM. If the helix is docked into the active site, the K_m^G is 4-fold tighter, approximately 80 μM. The difference is due to energetic coupling between G and substrate binding (McConnell et al., 1993; Bevilacqua et al., 1993). With a G·U pair at the cleavage site, K_m^G is equal to the equilibrium dissociation constant for guanosine binding, K_d^G, and k_{cat} is equal to the rate constant for the chemical step, k_c.

Previous work has shown that $(k_{cat}/K_m)^G$ for *I22* cleavage of U(−1)S is 8-fold slower than for *G22* (Strobel & Cech, 1995). Based upon a complete reactivity curve for the I·U pair, GMP binding (K_m^G) is unaffected by the removal of the exocyclic amine (Figure 6 and Table 2). This indicates that for the ground state of the reaction the I·U splice site helix remains docked into the ribozyme active site. This is consistent with the conclusion from the −3U 2′-hydroxyl substitution data previously presented (Strobel & Cech, 1995). Although K_m^G is unaffected, k_{cat} is 6-fold slower (2.3 min^{-1}). Thus, in addition to making a substantial ground state contribution (2 kcal·mol^{-1}), the exocyclic amine also makes a modest contribution (1 kcal·mol^{-1}) to transition state stabilization.

To confirm that the rate-limiting step for k_{cat} remains the chemical transition, we measured the cleavage rate by *I22*

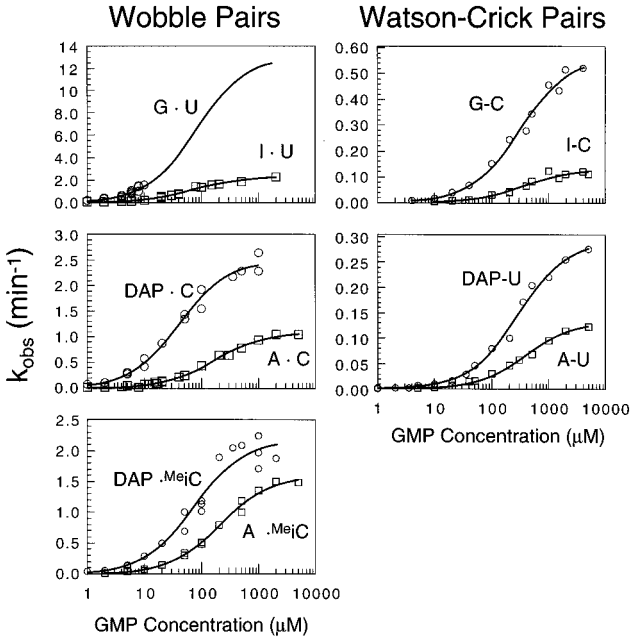

FIGURE 6: Measurement of K_m^G and k_{cat} for each of the nine base pairing combinations for the reaction GGCCCUCYAAAAA + pG → GGCCCUCY + pGAAAAA. Each line represents the best nonlinear least-squares fit of the data to the equation $k_{obs} = k_{cat}[GMP]/(K_m^G + [GMP])$ and solving for k_{cat} and K_m^G (McConnell et al., 1993). Note that the magnitude of the *y*-axis is different for each graph. The plots are grouped as equivalent pairs with and without the exocyclic amine such that the contribution of the amine in the context of each base pair can be directly visualized. The G·U plot is a theoretical line based upon K_m^G, $(k_{cat}/K_m)^G$, and the calculated value for k_{cat} (see Table 2 legend).

of a substrate with an R_P phosphorothioate at the cleavage site (CCCUCU$_{R_P}$A) (Table 2). *I22* showed a phosphorothioate effect of 3.1. This is comparable to the 2.3-fold effect seen for *G22* cleavage of the same substrate (Herschlag et al., 1991). It indicates that for the I·U pair the chemical step remains limiting and that k_{cat} is equal to k_c. Thus, the 6-fold reactivity difference represents a 1 kcal mol^{-1} stabilization of the chemical transition.

The Amine Provides Chemical Transition State Energy Only in the Context of a G·U Pair. Following the same logic as outlined for reaction infidelity, we postulated that a 1 kcal·mol^{-1} difference in transition state energy should be observed between the other wobble pairs with and without the critical exocyclic amine. Kinetic parameters were determined for each of the ribozyme−substrate combinations (Figure 6 and Table 2). For all of the wobble pairs, including A·C, DAP·C, A·^{Me}iC, and DAP·^{Me}iC, k_{cat} is at least 6-fold less than k_{cat} for the G·U pair and approximately the same as I·U. All of the pairs show a thio effect suggesting that chemistry remains the rate-limiting step. (The relationship between k_{cat} and k_c will be discussed in detail below.)

The small differences in k_{cat} between pairs with and without the amine suggest that the transition state energy has not been restored. For example, DAP·^{Me}iC reacts less than 2-fold faster than A·^{Me}iC, and DAP·C is barely 2-fold faster than A·C (Table 2). Even these slight differences in reactivity are unlikely to be due to chemical transition stabilization.

k_{cat} Is Sensitive to the Ground State Occupancy of the Splice Site Helix. At first glance, the reactivity of Watson−Crick pairs with and without the exocyclic amine suggests the amine is providing transition state rescue for these pairs

Table 2: Kinetic Characterization of G·U Wobble Pair Substitutions (pH 7.0, 30 °C)[a]

base pair	$(k_{cat}/K_m)^G$ ($\times 10^4$ M^{-1} min^{-1})	K_m^G(μM)	k_{cat} (min^{-1})	R_p $(k_{cat}/K_m)^G$ ($\times 10^4$ M^{-1} min^{-1})	thio effect
wobble pairs with exocyclic amine					
G·U	15[b]	90[b]	13[c]		2.3[d]
DAP·C	4.3 ± 0.4	41 ± 5	2.5 ± 0.1	2.2 ± 0.2	2.0
DAP·^{Me}iC	2.4 ± 0.2	75 ± 14	2.2 ± 0.1	0.52 ± 0.1	4.6
wobble pairs without exocyclic amine					
I·U	1.9[e]	74 ± 5	2.3 ± 0.1	0.60 ± 0.1	3.1
A·C	0.65 ± 0.08	160 ± 10	1.1 ± 0.1	0.21 ± 0.02	3.1
A·^{Me}iC	0.62 ± 0.06	210 ± 20	1.6 ± 0.1	0.17 ± 0.02	3.6
Watson−Crick pairs with exocyclic amine					
G-C	0.15[b]	310[a]	0.56[b]	0.023 ± 0.003	6.5
DAP-U	0.088 ± 0.01	280 ± 20	0.29 ± 0.01	0.021 ± 0.002	4.2
Watson−Crick pairs without exocyclic amine					
I-C	0.027 ± 0.005	350 ± 70	0.13 ± 0.01	0.0092 ± 0.001	2.9
A-U	0.031 ± 0.004	390 ± 20	0.13 ± 0.01	0.0050 ± 0.001	6.2

[a] $(k_{cat}/K_m)^G$, K_m^G, and k_{cat} are the rate constants for the reaction GGCCCUCYAAAAA + pG → GGCCCUCY + pGAAAAA. R_p $(k_{cat}/K_m)^G$ is the rate constant for the reaction CCCUCY$_{R_p}$A + pG → CCCUCY + pG$_{S_p}$A. The difference between the substrates in the two reactions, the 5′-G's and the length of the 3′-A tail, makes less than a 20% difference in reaction rate for the G·U and I·U pairs. This means the difference in $(k_{cat}/K_m)^G$ compared to R_p $(k_{cat}/K_m)^G$ results from the R_p phosphorothioate. [b] From Table 3 of Pyle et al. (1994). [c] Estimated value for k_{cat} of U(−1)S cleavage. The value is calculated from the second order rate constant $(k_{cat}/K_m)^G$ measured at pH 7.0 and K_m^G measured at pH 5.5. GMP binding is pH independent in this range (McConnell et al., 1993). [d] From Table 2 of Herschlag et al. (1991). [e] From Figure 5 of Strobel and Cech (1995).

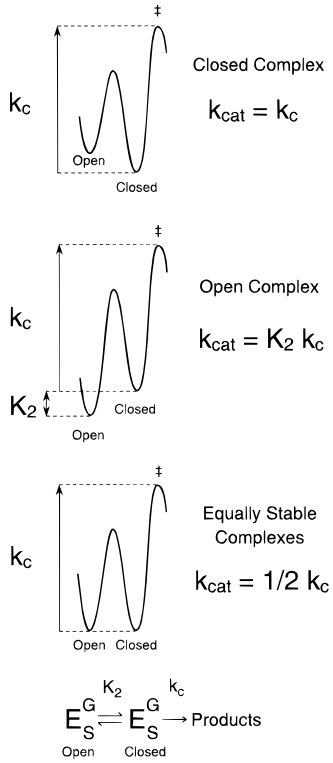

FIGURE 7: Free energy profiles showing the effect of helix docking equilibrium on k_{cat}.

(Figure 6). Although all the Watson−Crick pairs react 20−100-fold slower than G·U, k_{cat} is faster for pairs with the amine than for those without it, and they all show a thio effect indicating that chemistry is limiting (Table 2). While there is only a modest 2-fold difference between DAP-U and A-U, there is more than a 4-fold difference between G-C and I-C. This is approaching the 6-fold total difference observed between G·U and I·U. However, there is reason to suspect that k_{cat} is not equal to k_c for the Watson−Crick pairs and even a subset of the wobble pairs.

This argument derives from two independent pieces of evidence, both related to the ground state occupancy of the splice site helix in the enzyme active site. The first is the contribution made by the 2′-hydroxyl of −3U to substrate binding. The hydroxyl at −3U contributes about 1.5 kcal·mol^{-1} if the helix is in a docked conformation in the ground state (Pyle & Cech, 1991; Bevilacqua & Turner, 1991). It makes no contribution if the helix is undocked in the ground state. On the basis of this argument, all of the Watson−Crick and a subset of the wobble pairs are undocked in the ground state (Strobel & Cech, 1995). These include the G-C, I-C, DAP-U, and A-U Watson−Crick pairs and the A·C and A·^{Me}iC wobble pairs. The substrate helices of the remaining wobble pairs, including G·U, DAP·C, DAP·^{Me}iC, and I·U, are docked into the active site.

This argument is reinforced by measurement of K_m^G for each of these pairs (Table 2). All of the Watson−Crick pairs have a K_m^G consistent with an open complex (Table 2). All the wobble pairs with the critical exocyclic amine, including G·U, DAP·C, and DAP·^{Me}iC, as well as the I·U pair, have a K_m^G consistent with a closed complex (Table 2). A·C and A·^{Me}iC have an intermediate K_m^G that might reflect partial occupancy of the substrate helix in the active site.

The k_{cat} data can be explained using a model that includes a term for the docking equilibrium (Figure 7). For the *Tetrahymena* ribozyme, at least two steps occur after formation of the E_S^G complex (Bevilacqua et al., 1992; Herschlag 1992).

$$\underset{\text{open}}{E_S^G} \overset{K_2}{\leftrightarrow} \underset{\text{closed}}{E_S^G} \xrightarrow{k_c} \text{products} \qquad (1)$$

Following formation of the splice site helix, it is docked into the active site to form a closed complex. This equilibrium association constant is represented by K_2 (Bevilacqua et al., 1992). After docking, the substrate is converted into products at a rate constant equal to k_c. The maximal rate of this reaction, k_{cat}, reflects both the rate of the chemical transition and the docking equilibrium.

$$k_{cat} = k_c \frac{K_2}{K_2 + 1} \qquad (2)$$

Thus, k_{cat} measures a different transition depending upon the magnitude of K_2 (Figure 7).

If docking is favorable, K_2 is large ($\gg$1) and eq 2 reduces to

$$k_{cat} = k_c \quad (3)$$

This is the case for the original G·U pair and the I·U pair that lacks the exocyclic amine (Figure 7). This suggests that the 6-fold difference in k_{cat} is a difference in chemistry attributable to loss of the exocyclic amine.

If docking is unfavorable, K_2 is small ($\ll$1) and eq 2 reduces to

$$k_{cat} = k_c K_2 \quad (4)$$

This is the case for the Watson−Crick pairs (Figure 7). Because K_2 is factored into the magnitude of k_{cat} for these pairs, it is likely that the 4-fold difference between G-C and I-C is not a difference in k_c, but a small difference in the docking equilibrium K_2. This would be consistent with the 4-fold difference in tertiary docking energy between G-C and I-C measured both by gel shift and cleavage infidelity (Strobel & Cech, 1995; Table 1). A similar argument can be used to explain the 2-fold difference between DAP-U and A-U.

For a ribozyme that has equally stable open and closed complexes, K_2 approaches unity (Figure 7). In that case eq 2 reduces to

$$k_{cat} = k_c/2 \quad (5)$$

The intermediate values of K_m^G for A·C and A·^{Me}iC pairs suggest that this might be the situation for these pairs, thus resulting in their 2-fold slower rates compared to DAP·C, DAP·^{Me}iC, or I·U. The data are in agreement with a model in which all the wobble pairs except G·U perform the chemical step at the same rate ($k_c \approx 2$ min^{-1}), but k_{cat} is 2-fold slower for A·C and A·^{Me}iC due to the helix docking equilibrium.

The Contribution to the Chemical Step Made by the Exocyclic Amine Is Independent of the 2′-Hydroxyl of −1U. Having established that the exocyclic amine contributes modestly to the chemical step of the ribozyme reaction, we next sought to define the basis of its contribution. The crystal structure of an RNA duplex containing a G·U wobble pair included a highly ordered water molecule bridging between the exocyclic amine of the G and the 2′-hydroxyl of the U (Holbrook et al., 1991). This lead the authors to speculate that the water might provide transition state stabilization for the ribozyme reaction by increasing the ability of the 2′-hydroxyl to serve as a hydrogen bond donor. This would help stabilize the developing negative charge on the leaving group, the neighboring 3′-oxygen atom. Using a similar logic, Knitt et al. (1994) speculated that the bridging water molecule served an entropic role by freezing out the rotational freedom of the 2′-hydroxyl. In both models, the exocyclic amine makes its contribution through an indirect contact with the 2′-hydroxyl.

The 2′-hydroxyl of −1U contributes about 1000-fold to the chemical transition state (Herschlag et al., 1993). On the basis of the reactivity of the *I22* ribozyme, we have demonstrated that the exocyclic amine contributes approximately 6-fold. If the contribution of the amine is through indirect interaction with the 2′-hydroxyl of −1U, as has been proposed, then deletion of both functional groups should be no worse than deletion of the 2′-hydroxyl alone.

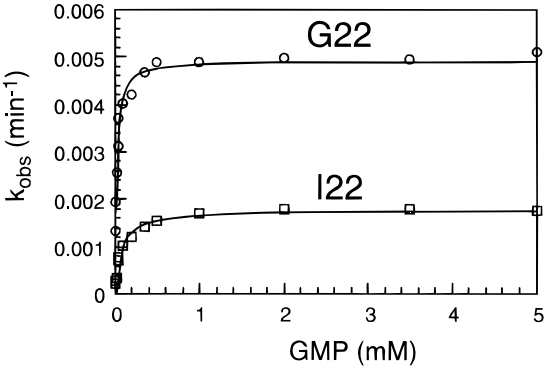

FIGURE 8: Reaction rate constants for cleavage of dU(−1)S [CCCUC(dU)A, where dU represents a 2′-deoxyribose substitution at position −1] by the *G22* and *I22* ribozymes as a function of GMP concentration at 30 °C. k_{cat}, represented by the horizontal lines, equals 0.0051 ± 0.0003 and 0.0016 ± 0.0002 min^{-1} for *G22* and *I22*, respectively.

However, if the contribution of the amine is independent of the 2′-hydroxyl, then the effects should be cumulative.

To distinguish between these possibilities, we measured the complete GMP binding curve for *I22* and *G22* cleavage of dU(−1)S (CCCUCdUA) under saturating enzyme conditions, pH 7.0, 10 mM $MgCl_2$. This substrate has a 2′-deoxy substitution at the cleavage site. Under these conditions, k_{cat} for *G22* is equal to the rate of the chemical transition. Consistent with previously published rates, the 2′-deoxyribose substitution reduced k_c by almost 2500-fold compared with the ribose cleavage rate (0.0051 ± 0.0003 min^{-1} compared with 13 min^{-1}) (Figure 8). However, the *I22* ribozyme reacted an additional 3-fold slower than *G22* (0.0016 ± 0.0002 min^{-1}) (Figure 8). Thus, the exocyclic amine makes a modest contribution to the chemical transition state that is independent of the contribution made by the 2′-hydroxyl. On the basis of a 6-fold difference in k_c between *I22* and *G22* for cleavage of a ribose, and a 3-fold difference for cleavage of a 2′-deoxyribose, the exocyclic amine is contributing at most 0.4 kcal·mol^{-1} (2-fold) via the 2′-hydroxyl (Figure 9). On the basis of models of a G·U wobble pair in an A-form RNA duplex, it is impossible to make a path from the exocyclic amine to the 3′-oxygen except through the 2′-hydroxyl. Thus, the remaining energy (0.7 kcal·mol^{-1}) must not derive from direct or even indirect contact with the phosphate (Figure 9).

DISCUSSION

The G·U wobble pair at the 5′-splice site is nearly universally conserved among known group I introns (Cech, 1988; Michel & Westhof, 1990). In a previous study, we used unnatural base pair substitutions to demonstrate thermodynamically that the exocyclic amine of the G is an important tertiary contact for docking the splice site helix into the active site of the ribozyme (Strobel & Cech, 1995). In the present study, we have kinetically characterized the same set of substitutions and again found that the key to the functional importance of the G·U pair resides in its unique presentation of the exocyclic amine of G. For all parameters tested, including binding of the 5′-exon, accuracy of 5′-splice site selection, and chemical transition state stabilization, the exocyclic amine is an important molecular determinant.

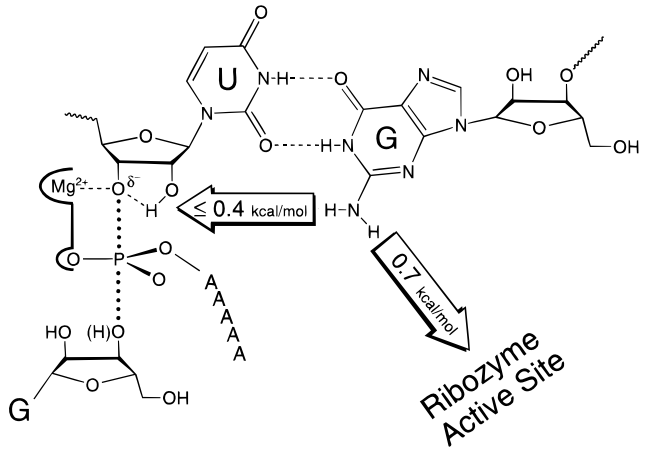

FIGURE 9: Transition state model for the effect of the exocyclic amine of G on the chemical step [modified from Cech et al. (1992)]. The exocyclic amine makes most of its modest contribution to the transition state (0.7 kcal·mol^{-1}) in a manner independent of the 2′-hydroxyl at the cleavage site. The simplest model for its contribution involves a more stable tertiary bond between the amine and its contact in the ribozyme active site during the chemical transition state. The exocyclic amine contributes at most 0.4 kcal·mol^{-1} via the 2′-hydroxyl, possibly through a bridging water molecule (Knitt et al., 1994; Holbrook et al., 1992).

Chemical Transition State Stabilization Provided by the G·U Pair. The *Tetrahymena* ribozyme performs the transesterification reaction with a rate enhancement of approximately 10^{11} over the uncatalyzed rate. Not all of the energy needed to explain this rate enhancement has been accounted for, though some critical elements have been identified (Cech et al., 1992). In the transition state the 3′-oxygen is coordinated to a metal that provides 1000-fold stabilization (Piccirilli et al., 1993). There is also an interaction with the S_P oxygen of the scissile phosphate, perhaps by the same metal ion, though its contribution has not been quantitated. Additional energy is provided by ground state destabilization of the substrate (Narlikar et al., 1994). Previous work has implied that the wobble conformation or specific functional groups of the G·U base pair make a modest (less than 10-fold) contribution (Green et al., 1991; Pyle et al., 1994; Knitt et al., 1994; Strobel & Cech, 1995). On the basis of this study, the extent and molecular basis for this contribution can be quantitated.

Reaction at the normal cleavage site is faster with a G·U pair than with any other base combination. It is 6-fold faster than with an I·U pair, a pair that is identical to the G·U in functional group presentation except for the simple subtraction of the exocyclic amine. Because the chemical step was shown to be limiting for the G·U and I·U pairs, the difference indicates that the exocyclic amine stabilizes the chemical transition state by a modest 1 kcal·mol^{-1}. Characterization of the reactivity difference between G·U and I·U with a 2′-deoxy substitution at the cleavage site showed that the majority (0.7 kcal·mol^{-1}) of this energy is in addition to that provided by the 2′-hydroxyl at −1U, thus arguing against the bridging water molecule model (Holbrook et al., 1991; Knitt et al., 1994). The simplest model to explain the data is that the tertiary interaction between the exocyclic amine and a hydrogen bonding partner in the active site is modestly improved during the chemical transition (Figure 9). This could derive from small changes in the electrostatic environment surrounding the hydrogen bond or an improved geometric alignment of the bonding partners. In either scenario the contribution is small, accounting for only 3-fold of the 10^{11}-fold rate enhancement.

k_{cat} for DAP·^{Me}iC and DAP·C Is Slower than that for G·U. The failure of DAP·^{Me}iC and DAP·C to react as fast as G·U could be interpreted to mean another functional group in the wobble pair makes a modest contribution to the transition state. While we cannot completely rule out this possibility, the more likely explanation is that while the amine is sufficiently aligned by these pairs to make a strong ground state contact, it is not properly positioned to make a transition state contribution. Generally, there are more stringent geometric requirements for transition state stabilization than for stabilization in the ground state. For reasons that are not obvious in the absence of structural information, the geometric requirements for transition state stabilization do not appear to be completely met by the DAP·^{Me}iC and DAP·C pairs.

Quantitative Analysis of 5′-Splice Site Selection. The *Tetrahymena* ribozyme is a million-fold more likely to cleave the proper 5′-exon/intron boundary than the next most frequently used cryptic site. The *G22* ribozyme cleaves the correct site (−1U) at a rate of 13 min^{-1} while the most frequently miscut site (the A-U pair at position −3U) is cut at a rate of 1.4×10^{-5} min^{-1}. This difference corresponds to 8.2 kcal·mol^{-1} of ground or transition state energy that must be defined to quantitatively explain 5′-splice site selection. On the basis of the rate constants measured in this study, the ribozyme achieves much of this specificity through the contributions of the G·U exocyclic amine.

(1) Ground State Binding to the Exocyclic Amine. The G·U wobble pair presents the exocyclic amine as a handle to dock the splice site helix into the proper alignment, presumbably through hydrogen bonding. This engages the −1U scissile phosphate with the catalytic apparatus of the active site. On the basis of the miscleavage rates of G·U compared to I·U, 2.5 kcal·mol^{-1} of alignment energy is contributed through a tertiary interaction with the amine. Equilibrium binding measurements estimated this contribution to be 2.0 kcal·mol^{-1}. The exocyclic amine can make most of this contribution in the context of any of the three wobble pairs tested (G·U, DAP·C, and DAP·^{Me}iC), but it cannot make a contribution when hydrogen bonded in a Watson−Crick configuration.

(2) Preferential Fit for a Wobble Conformation. In addition to the exocyclic amine, the shape of the wobble pair also appears to contribute a modest amount of alignment energy (about 2−5-fold or <1 kcal·mol^{-1}; for example, in Table 1 compare the infidelity seen with I·U to that of the Watson−Crick pairs A-U and I-C). This energy does not seem to derive from a specific functional group, but rather from a preferential fit for a wobble pair in the active site of the ribozyme as originally suggested by Doudna et al. (1989).

(3) Transition State Stabilization by the G·U Pair. The G·U wobble pair provides greater transition state stabilization than other base pairs, including the G-C and A-U pairs at neighboring positions within the splice site helix. When G-C and A-U are substituted for the G·U pair at the primary cleavage site, reactivity decreased 25- and 100-fold, repectively (Table 2). This indicates that the G·U pair contributes an upper limit of 2.8 kcal·mol^{-1} to splice site specificity by providing greater transition state stabilization compared to the A-U pair at the −3 position. About 1 kcal/mol derives from transition state stabilization by the exocyclic amine. Of the remaining 1.8 kcal·mol^{-1}, it is not yet clear how much

arises from a difference in the chemical step versus the ground state contributions already enumerated above. Further assignment of this energy requires measurement of k_c for a Watson−Crick pair in a docked conformation, a value difficult to measure because the Watson−Crick pairs disfavor docking.

(4) Other Structural Elements Constrain the Mobility of the Splice Site Helix. While the G·U pair is a major determinant of 5′-splice site selection, other structural elements within the ribozyme have also been shown to contribute. The splice site helix is conformationally constrained by the single-stranded region linking it to the P2 helix (Herschlag, 1992) and by the helical stacking interface between P2 and P2.1 (Downs & Cech, 1990, 1994). Disruption of either of these structural elements significantly increases the infidelity of the ribozyme. This is proposed to result from the helix having greater translational freedom relative to the active site, thus allowing it to sample alternative docking registers more frequently.

On the basis of the rate constants measured in the study, we can quantitatively estimate the positive contribution to 5′-splice site selection made by these structural elements. By putting an A-U at the preferred site of cleavage, the ribozyme can choose between two A-U pairs, one at the correct cleavage site and one at the most frequently used cryptic site, −3U. Although these two potential sites of reaction appear chemically equivalent, the ribozyme still preferentially cuts the correct site 30-fold faster than the cryptic site (0.13 min^{-1} for the correct site, 0.0043 min^{-1} for the cryptic site) indicating that for this position the other structural elements provide the helix with 2.0 kcal·mol^{-1} of conformational constraint. For cleavage at other positions within the splice site helix, the contribution from other structural elements is substantially larger. For example a G-C pair is about 5-fold more reactive than an A-U pair and is more stably docked into the active site, yet cleavage at −2C (a G-C pair) is at least 10-fold slower than at −3U (an A-U pair). This implies that other structural elements provide a greater degree of conformational constraint against docking in the −2C register.

While sequence nonspecific contacts such as the 2′-hydroxyls at positions −3, 22, and 25 provide greater docking energy, their energetic contribution is categorically different from that provided by the exocyclic amine. Because commensurate sets of hydroxyl groups can be used in alternative docking registers, the energy generated from the 2′-hydroxyls does not provide alignment specificity (Herschlag, 1992; Strobel & Cech, 1994).

In summary, 8.2 kcal·mol^{-1} of energy is needed to explain the ribozyme's specificity for the correct 5′-exon/intron boundary at −1U over the cryptic site at −3U. This has been accounted for in the following way: (i) 2.5 kcal·mol^{-1} of alignment energy from the G·U exocyclic amine, (ii) About 1 kcal·mol^{-1} of alignment energy from the general shape of the G·U wobble pair, (iii) between 1.0 and 2.8 kcal·mol^{-1} of enhanced transition state stabilization from the G·U pair, and (iv) 2.0 kcal·mol^{-1} of conformational constraint derived from other structural elements within the ribozyme. This totals between 6.5 and 8 kcal·mol^{-1}, suggesting that these features make independent energetic contributions and that the majority of the energy involved in 5′-splice site selection has been quantitatively assigned. The total energy and relative contributions of each element will be different when defining the specificity relative to other positions. For example, discrimination against the −2C position relies more heavily on conformational constraint provided by the structural elements linking the splice site helix to the ribozyme.

Organic Synthesis and RNA Enzymology. The G·U wobble pair was one of the first conserved elements identified in group I introns (Michel et al., 1982; Davies et al., 1982). Although mutagenesis of the G·U to other base pairs confirmed that the G·U wobble was important, the molecular basis for its contribution could not be defined until each of the functional groups could be independently modified within the context of a wobble configuration. Through synthetic organic chemistry, six different wobble pairs were generated that systematically altered the majority of the functional groups of the G·U pair, thus providing another demonstration of the utility of synthetic chemistry for answering questions of macromolecular enzymology.

Systematic chemical alteration of the functional groups in the G·U pair demonstrated that the N_2 exocyclic amine makes a significant contribution to binding and 5′-splice site selection. The lone contact to the G is somewhat surprising, given that the U is actually the nucleotide at the 5′-exon/intron boundary. It appears, however, that the U is not directly contacted by the active site of the enzyme. Instead the U positions the G into the proper conformation for the exocyclic amine to make its contribution.

ACKNOWLEDGMENT

We thank Phil Bevilacqua and Tim McConnell for helpful discussion, and Anne Gooding for oligonucleotide synthesis.

REFERENCES

Barfod, E. T., & Cech, T. R. (1989) *Mol. Cell. Biol. 9*, 3657−3666.
Been, M. D., & Cech, T. R. (1986) *Cell 47*, 207−216.
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.
Bevilacqua, P. C., Johnson, K. A., & Turner, D. H. (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90, 8357−8361.
Cech, T. R. (1987) *Science 236*, 1532−1539.
Cech, T. R. (1988) *Gene 73*, 259−271.
Cech, T. R.. Herschlag, D., Piccirilli, J. A., & Pyle, A. M. (1992) *J. Biol. Chem. 267*, 17479−17481.
Cech, T. R., Damberger, S. H., & Gutell, R. R. (1994) *Nature Struct. Biol. 1*, 273−280.
Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A., & Scazzacchio, C. (1982) *Nature 300*, 719−724.
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. (1990) *Biochemistry 29*, 5605−5613.
Downs, W. D., & Cech, T. R. (1994) *Genes Dev. 8*, 1198−1211.
Green, R., Szostak, J. W., Benner, S. A., Rich, A., & Usman, N. (1991) *Nucleic Acids Res. 19*, 4161−4166.
Herschlag, D. (1992) *Biochemistry 31*, 1386−1399.
Herschlag, D., & Cech, T. R. (1990) *Biochemistry 29*, 10159−10171.
Herschlag, D., Piccirilli, J., & Cech, T. R. (1991) *Biochemistry 30*, 4844−4854.
Herschlag, D., Eckstein, F., & Cech, T. R. (1993) *Biochemistry 32*, 8312−8321.
Holbrook, S. R., Cheong, C., Tinoco, I., & Kim, S.-H. (1991) *Nature 353*, 579−581.
Kimura, J., Yagi, K., Suzuki, H., & Mitsunobu, O. (1980) *Bull. Chem. Soc. Jpn. 53*, 3670−3677.
Knitt, D. S., Narlikar, G. J., & Herschlag, D. (1994) *Biochemistry 33*, 13856−13879.

McConnell, T. S., Cech, T. R., & Herschlag, D. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8362−8366.
Michel, F., & Westhof, E. (1990) *J. Mol. Biol. 216*, 585−610.
Michel, F., & Westhof, E. (1994) *Nature Struct. Biol. 1*, 5−7.
Michel, F., Jacquier, A., & Dujon, B. (1982) *Biochimie 64*, 867−881.
Moore, M. J., & Sharp, P. A. (1992) *Science 256*, 992−997.
Narlikar, G. J., Gopalakrishnen, V., McConnell, T. S., Usman, N., & Herschlag, D. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 3668−3672.
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*, 628−631.
Pyle, A. M., Murphy, F. L., & Cech, T. R. (1992) *Nature 358*, 123−128.
Pyle, A. M., Moran, S., Strobel, S. A., Chapman, T., Turner, D. H., & Cech, T. R. (1994) *Biochemistry 33*, 13856−13863.
Reddy, M. P., Hanna, N. B., & Forooqui, F. (1994) *Tetrahedron Lett. 35*, 4311−4314.
Salvo, J. L., & Belfort, M. (1992) *J. Biol. Chem. 267*, 2845−2848.
Scaringe, S. A., Francklyn, C., & Usman, N. (1990) *Nucleic Acids Res. 18*, 5433−5441.
Slim, G., & Gait, M. J. G. (1991) *Nucleic Acids Res. 19*, 1183−1188.
Strobel, S. A., & Cech, T. R. (1993) *Biochemistry 32*, 13593−13604.
Strobel, S. A., & Cech, T. R. (1994) *Nature Struct. Biol. 1*, 13−17.
Strobel, S. A., & Cech, T. R. (1995) *Science 267*, 675−679.
Strobel, S. A., Cech, T. R., Usman, N., & Beigelman, L. (1994) *Biochemistry 33*, 13824−13835.
Switzer, C. Y., Moroney, S. E., & Benner, S. A. (1993) *Biochemistry 32*, 10489−10496.
Tor, Y., & Dervan, P. B. (1993) *J. Am. Chem. Soc. 115*, 4461−4467.
Usman, N., Ogilvie, K. K., Jiang, M.-Y., & Cedergren, R. J. (1987) *J. Am. Chem. Soc. 109*, 7845−7854.
Waring, R. B., Towner, P., Minter, S. J., & Davies, R. W. (1986) *Nature 321*, 133−139.
Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Tracz, D., Workman, C., Sweedler, D., Gonzalez, C., Scaringe, S., & Usman, N. (1995) *Nucleic Acids Res. 23*, 2677−2684.
Zaug, A. J., Been, M. D., & Cech, T. R. (1986) *Nature 324*, 429−433.
Zaug, A. J., Grosshans, C. A., & Cech, T. R. (1988) *Biochemistry 27*, 8924−8931.

BI952244F