Replacement of the Conserved G·U with a G-C Pair at the Cleavage Site of the *Tetrahymena* Ribozyme Decreases Binding, Reactivity, and Fidelity†

Anna Marie Pyle,‡,§ Sean Moran,‖ Scott A. Strobel,‡ Teresa Chapman,‡ Douglas H. Turner,‖ and Thomas R. Cech*,‡

Department of Chemistry and Biochemistry and Howard Hughes Medical Institute, University of Colorado, Boulder, Colorado 80309, and Department of Chemistry, University of Rochester, Rochester, New York 14627

Received June 28, 1994; Revised Manuscript Received August 31, 1994⊗

ABSTRACT: There is a phylogenetically conserved G·U pair at the 5′-splice site of group I introns. When this is mutagenized to a G-C pair, splicing of these introns is greatly reduced. We have used a ribozyme derived from the *Tetrahymena* group I intron to compare the binding and reactivity of oligonucleotides that form either a G·U or a G-C pair at this position. Ribozyme binding of oligonucleotides at 42 °C was measured by native gel electrophoresis and equilibrium dialysis. Binding of GGCCCUCC (C(-1)P), which base-pairs with the ribozyme guide sequence to form a G-C at the cleavage site, was 10-fold weaker than the binding of GGCCCUCU (U(-1)P), which maintains the conserved G·U pair at the cleavage site. This is surprising since a terminal G-C enhances the binding between oligonucleotides by 20-fold relative to a terminal G·U. Thermal denaturation studies indicate that C(-1)P and several analogs with deoxy substitutions bind the guide-sequence oligonucleotide, GGAGGGAAA, as strongly as they bind the ribozyme. In contrast, U(-1)P binds 240-fold more strongly to the ribozyme than to GGAGGGAAA, a difference that is decreased by deoxy substitutions. Thus, while U(-1)P binds the ribozyme through a combination of base-pairing and specific 2-OH and other tertiary interactions, C(-1)P may bind by base-pairing alone. The substrate GGCCCUCCAAAAA (C(-1)S) is cleaved 100-fold more slowly than GGCCCUCUAAAAA (U(-1)S) and also has a higher propensity to be cleaved at the wrong nucleotide position. Taken together, the results suggest that a G-C pair at the ribozyme cleavage site makes docking of the guide-sequence·substrate helix into the catalytic site less favorable than a G·U pair. The resulting consequences of weaker binding, slower reaction, and reduced cleavage fidelity provide a rationale for the phylogenetic conservation of the G·U.

Group I self-splicing introns and ribozymes are rich in phylogenetically conserved nucleotides throughout a central core region that is apparent in models of their secondary and tertiary structures (Cech, 1988; Michel & Westhof, 1990). Nucleotides surrounding this core and near the intronic termini generally are not conserved. An important exception is the conserved G·U pair at the 5′-splice site. Natural mutation of this G·U to a G-C has not yet been reported. For the *Tetrahymena* intron, the conserved G·U pairing is important for the first, but not the second, step of splicing (Barfod & Cech, 1989). Site-directed mutagenesis studies showed that, of all possible pairings in place of the conserved G·U, the A·C pair retains the highest level of activity in reactions by the ribozyme (Doudna et al., 1989). This suggested that the conformation of the base pair is important for the catalysis and perhaps recognition of the reaction site.

There are two steps involved in recognition of the group I intron 5′-splice site. The first step involves base-pairing between an internal guide sequence (IGS)[1] on the intron and complementary sequences on the 5′-exon to form a helix designated P1 (Been & Cech, 1986; Waring et al., 1986; Zaug et al., 1986; Figure 1). In a kinetically separable event, the intact P1 helix docks into the core of the intron (Bevilacqua et al., 1992; Herschlag, 1992), forming a network of tertiary interactions that include contacts to 2′-OH groups on both strands of the incoming helix. This enhances binding by about 4 kcal/mol (Bevilacqua & Turner, 1991; Pyle & Cech, 1991; Strobel & Cech, 1993). These backbone functionalities form specific hydrogen bonds to groups in the core of the ribozyme (Pyle et al., 1992; Michel & Westhof, 1994).

The conserved G·U pair at the cleavage site and the array of 2′-OH groups along the P1 helix represent two distinct

† Supported by NIH Grants GM28039 (T.R.C.) and GM22939 (D.H.T.). A.M.P. was a fellow of the Jane Coffin Childs Fund for Medical Research. S.A.S. is a fellow of the Life Sciences Research Foundation sponsored by the Howard Hughes Medical Institute. D.H.T. is a Guggenheim Fellow and an American Cancer Society Scholar at the University of Colorado. T.R.C. is an Investigator of the Howard Hughes Medical Institute.

* Author to whom correspondence should be addressed.

‡ University of Colorado.

§ Present address: Department of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, New York, NY 10032.

‖ University of Rochester.

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

[1] Abbreviations: K_d, equilibrium dissociation constant; K_d^G equilibrium dissociation constant for the ribozyme−guanosine complex; $\Delta G°$, standard free energy of association; P1, the RNA helix containing the 5′-splice site (and G·U pair) of group I introns; tRNA, transfer RNA; S, oligonucleotide substrate for ribozyme-catalyzed cleavage; P, oligonucleotide representing the 5′-terminal product of ribozyme-catalyzed cleavage; U(-1)P, GGCCCUCU; U(-1)S, GGCCCUCUAAAAA; U(-1)P′, CCCUCU; C(-1)P, GGCCCUCC; C(-1)S, GGCCCUCCAAAAA; C(-1)P′, CCCUCC; d(N), a deoxynucleotide indicated at a particular position, where N = G, C, U, or A; GMP, guanosine 5′-monophosphate; IGS, internal guide sequence; PAGE, polyacrylamide gel electrophoresis. Note that the prime on U(-1)P′ indicates the absence of two G's on the 5′-end, whereas the prime on P′ indicates the absence of a U on the 3′-end of the oligonucleotide.

0006-2960/94/0433-13856$04.50/0

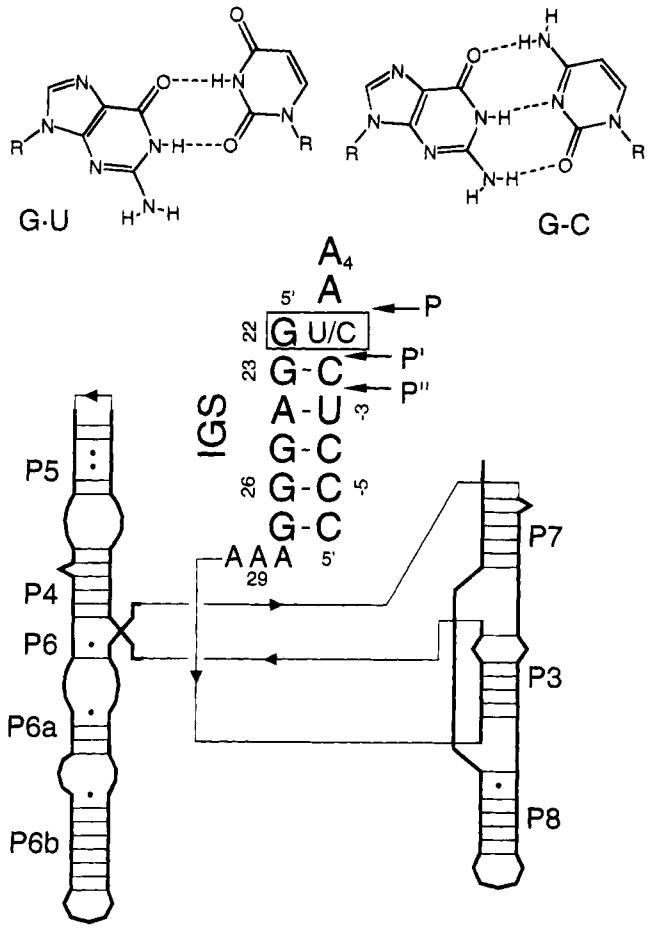

FIGURE 1: Diagram of the P1 helix and the catalytic core of the group I intron [modified from Salvo and Belfort (1992) and Cech et al. (1994)]. The P1 helix is shown in capital letters with both the G·U and G-C pairs at the cleavage site (box). The most likely hydrogen-bonding patterns of the G·U and G-C base pairs are shown at the top of the figure. The numbering of specific bases within the helix are shown on both sides of the P1 helix. The positions of miscleavage that result in P′ and P″ products are also shown. Helices P3−P8, the residues that are proposed to constitute the active site of the ribozyme, are shown as solid lines.

factors important for P1 recognition. Of these two contributions, only the G·U pair provides sequence specificity. It was therefore of interest to determine the consequences for individual steps of the ribozyme reaction of replacing this G·U with a G-C pair.

MATERIALS AND METHODS

Preparation and Purification of RNA. Oligonucleotides were synthesized (Usman et al., 1987) on Applied Biosystems Model 380B or 392 DNA/RNA synthesizers using phosphoramidites from Applied Biosystems. Following workup under standard conditions (Scaringe et al., 1990), the oligonucleotides used in gel electrophoresis and kinetic studies were purified by PAGE. Following base deprotection, oligonucleotides used in thermal denaturation and dialysis experiments were treated with triethylamine hydrogen fluoride and desalted (Bevilacqua et al., 1994). The pyrimidine-rich sequences were purified on Baker Si500F silica TLC plates using *n*-propanol/H_2O/NH_4OH (55:10:35) as the solvent system. The guanosine-rich sequences were heated to 75 °C in the presence of denaturing gel loading buffer for 20 min to break up any aggregation. After purification by PAGE, all oligonucleotides were UV-shadowed and eluted from the gel slices with sterile H_2O or buffer. After ethanol precipitation and desalting with a Waters Sep-Pak column, the purity of all oligonucleotides was checked by 5′-end-labeling and PAGE. For all RNAs that were not radiolabeled, concentrations were calculated using the extinction coefficients of the nearest-neighbor pairs and the nucleotides that make up the oligomer (Borer, 1975; Richards, 1975). For 5′-^{32}P-labeled RNAs, concentrations were determined from the specific activities. The L-21ScaI ribozyme used in this study was transcribed from plasmid pT7L-21 using T7 RNA polymerase (Zaug et al., 1988). For most experiments, the RNA was purified by PAGE, elution, and ethanol precipitation. For dialysis experiments, transcription reactions were phenol/$CHCl_3$ extracted, purified on a Qiagen anion-exchange column, and precipitated with isopropyl alcohol. Ribozyme was resuspended in 10 mM Tris (pH 7.5) and 0.1 mM EDTA and the concentration was determined using $\epsilon_{260} = 3.2 \times 10^6$ M^{-1} cm^{-1} (Zaug et al., 1988).

Gel-Shift Studies of Ribozyme Binding. Electrophoretic analyses of oligonucleotide binding to the ribozyme were performed using methods described previously (Pyle et al., 1990), except that longer incubation times were used, as discussed here (Bevilacqua & Turner, 1991; Pyle et al., 1992). Following ribozyme preincubation at 50 °C for 10 min, ribozyme and product oligonucleotides (0.5 nM) were incubated for 15 min at 42 °C before immediate loading on a nondenaturing polyacrylamide gel preequilibrated to 42 °C. Ribozyme concentrations were varied from significantly below to far above the K_d. The K_d for each trial was determined from the nonlinear least-squares fit to a complete binding curve. The K_d for each oligonucleotide represents the average of two complete trials, and error is calculated from the difference in these values, as described in Table 1. Significantly tighter binding was not observed at longer incubation times or lower concentrations of GGCCCUCC, which binds the tightest of all of the variant oligonucleotides measured in this work. As a general rule, when performing binding studies, it is important to measure K_d using several different ligand concentrations to assure that binding is a simple bimolecular reaction. In addition, it is also important to measure K_d after the samples have been incubated for several different incubation times in order to check for complete equilibration. The time to reach equilibrium will be dependent on the rates for the binding reaction and the concentrations used. For example, for the association of ribozyme and substrate, when either is in large excess, the rate (τ^{-1}) for approach to equilibrium is given by (Bernasconi, 1976; Fersht, 1985)

$$\tau^{-1} = k_{on}([E] + [S]) + k_{off} \quad (1)$$

Here, k_{on} and k_{off} are the apparent association and dissociation rates, and [E] and [S] are the final equilibrium concentrations of ribozyme and substrate, respectively. The association will be about 95% complete after 3τ. For the reactions studied here, k_{on} is expected to be about 10^8 M^{-1} min^{-1} (Herschlag & Cech, 1990; Bevilacqua et al., 1992). For a dissociation constant of 3 nM, this gives $k_{off} = K_d k_{on} \approx 0.3$ min^{-1}. Thus, for [ribozyme] = 3 nM and [oligonucleotide] = 0.5 nM, τ = 2 min.

Equilibrium Dialysis Studies of Ribozyme Binding. Equilibrium dialysis experiments were performed as previously

Table 1: Binding of Oligonucleotides to the L-21 *ScaI* Ribozyme and to an Oligonucleotide That Mimics the IGS

oligonucleotide[a]		ribozyme binding[b]		duplex formation[c]		tertiary energy[d]
name	sequence	K_d (nM)[e]	$\Delta G°_{42°C}$ (kcal/mol)	K_d (nM)	$\Delta G°_{42°C}$ (kcal/mol)	$\Delta\Delta G°_{42°C}$ (kcal/mol)
U(-1)P	GGCCCUCU	0.3	−13.7	70	−10.3	−3.4
U(-1)P′	CCCUCU	0.2[f]	−14.0[f]			−3.7
C(-1)P	GGCCCUCC	3.5	−12.2	3.5	−12.2	0
C(-1)P′	CCCUCC	6.4[g]	−11.8			+0.4
C(-1)Pd(-2)	GGCCCUd(C)C	16	−11.2	7.8	−11.7	+0.5
C(-1)Pd(-3)	GGCCCd(U)CC	30	−10.8	17	−11.2	+0.4
C(-1)Pd(-4)	GGCCd(C)UCC	9.6	−11.5			
C(-1)Pd(-2,-3)	GGCCCd(UC)C	43	−10.6	18	−11.2	+0.6
C(-1)Pd(-4,-5)	GGCd(CC)UCC	18	−11.1			
U(-1,-5)P	GGCUCUCU	22	−11.0	24,000	−6.7	−4.3
U(-1,-5)P′	CUCUCU	6.7[g]	−11.8			−5.1

[a] Gel mobility shift experiments used oligonucleotides with GG at the 5′-end. For all other experiments, the 5′-end started with C. [b] Oligonucleotide binding to the L-21 *ScaI* ribozyme from *Tetrahymena*. Values were measured using the gel mobility shift assay or by equilibrium dialysis. Dialysis samples contained 5 mM pdG to fill the G site but prevent reaction. [c] Oligonucleotide binding to the sequence GGAGGGAAA. Values were calculated from $1/T_M$ vs log(concentration) measurements of duplex stability (Table 2) made more favorable by 0.2 kcal/mol for the GG at the 5′-end. [d] $\Delta\Delta G°_{42°C} = \Delta G°_{42°C}$ (binding to ribozyme) − $\Delta G°_{42°C}$ (binding to GGAGGGAAA). [e] For the gel mobility shift experiments, the mean standard deviation in these values $([\Sigma(\text{mean value} - \text{experimental value})^2/N]^{1/2}$, where N represents two trials) was no more than 50% of the value shown and in most cases was significantly smaller. Each K_d is the average of values obtained from two complete binding curves. [f] From equilibrium dialysis using either straight counts or counts from a gel (no reaction observed). Samples contained 5 mM pdG. [g] From equilibrium dialysis using counts from a gel. Samples contained 5 mM pdG.

described (Bevilacqua & Turner, 1991; Moran et al., 1993), with the following modifications. The dialysis buffer contained 5 mM dGMP to prevent side reactions (Bevilacqua & Turner, 1991). At 42 °C, however, CCCUCC was both cleaved (~30%) and elongated (~5%) during the 3 days needed for dialysis samples to reach equilibrium. Presumably, this is due to a reaction similar to the disproportionation of C_5 (Zaug & Cech, 1986). CUCUCU was also cleaved to an oligonucleotide two nucleotides shorter (~15%). Because of these side reactions, K_d values for CCCUCC and CUCUCU were determined from the radioanalytical quantitation of intact substrate bands after electrophoresis of equilibrated dialysis samples (Moran et al., 1993). Dissociation constants determined in this manner are therefore approximations, since a reaction is proceeding while the equilibrium is being measured. Depletion of starting substrate does not affect the linearization of the expression for K_d, since ribozyme is still in excess over radiolabeled substrates. The fact that some ribozyme−substrate complex is reacting, however, means that there will be less full-length substrate on the ribozyme side of the dialysis membrane than should be present at equilibrium, if the rate of diffusion of full-length substrate from the oligonucleotide chamber is not rapid enough to reestablish equilibrium. We estimate that this could give an apparent K_d value of up to 50% higher than the true value.

Thermal Denaturation Analysis of Duplex Stability. Optical melting curves were measured at 280 nm in a buffer of 50 mM Tris (pH 7.6, 22 °C), 10 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM Na_2EDTA. Duplexes used to measure the stability of the P1 pairing interaction were annealed in the following manner. Concentrated stock solutions of IGS analogs were heated to 90 °C for 10 min to break up any aggregation. They were then combined with pyrimidine-rich oligonucleotides and buffer without magnesium. After vortex mixing, $MgCl_2$ was added to a final concentration of 10 mM. Thermodynamic parameters for duplex formation were determined by two methods (Petersheim & Turner, 1983). First, individual melting curves at a given total strand concentration (C_T) were fit to a two-state model for non-self-complementary duplexes (Longfellow et al., 1990). Thermodynamic parameters from the fits of a range of C_T values (range in concentration ~100-fold) were then averaged and compared to those obtained from a plot of the inverse of the melting temperature ($1/T_M$) vs the logarithm of $C_T/4$ (Borer et al., 1974):

$$1/T_M = (R/\Delta H°) \ln(C_T/4) + \Delta S°/\Delta H° \qquad (2)$$

Values of $\Delta G°_{42°C}$, $\Delta H°$, and $\Delta S°$ obtained by these two methods agree to within 10%, consistent with the two-state model.

Kinetics Measurements of Ribozyme-Catalyzed Cleavage Reactions. Kinetics methods followed those previously described (Herschlag & Cech, 1990; Herschlag et al., 1991), including preincubation of the L-21 *ScaI* ribozyme with 10 mM $MgCl_2$ at 50 °C for 20 min to create a homogeneously folded ribozyme population prior to cooling the ribozyme to 30 °C. Reactions were initiated by the addition of oligonucleotide substrate and GMP. Portions of the reaction were removed at various times and quenched on ice with 2 vol 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).

Measurements of k_{cat} and K_m^G were performed as per McConnell et al. (1993). Single-turnover reactions of saturating ribozyme (100 nM) with C(-1)S substrate were performed under varying GMP concentrations (5−4000 μM) at 30 °C in 10 mM $MgCl_2$, 50 mM HEPES (pH 7.0), and 0.5 nM ^{32}P-5′-end-labeled C(-1)S. The reactions obeyed first-order kinetics over 4 half-lives, with an end point of approximately 10% unreactive substrate.

Cleavage of U(-1)P (GGCCCUCU) and C(-1)P (GGCCCUCC) to produce P′ (GGCCCUC) and P″ (GGCCCU) was assayed under single-turnover conditions with saturating ribozyme and GMP. Rate constants for P′ and P″ production ($k_{cat}^{P'}$ and $k_{cat}^{P''}$) were measured at 50 °C in 50 mM HEPES (pH 7.0), 10 mM $MgCl_2$, and 5.0 mM GMP using a trace amount (≈0.2 nM) of ^{32}P-radiolabeled U(-1)P or C(-1)P and 200 nM ribozyme. The ribozyme and GMP concentrations

Table 2: Thermodynamic Parameters for Duplex Formation[a]

	from $1/T_M$ vs $\ln(C_T/4)$ plots				from average of fits			
sequence	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$-\Delta G°_{42°C}$ (kcal/mol)	T_M[b] (°C)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$-\Delta G°_{42°C}$ (kcal/mol)	T_M[b] (°C)
			Binding to GGAGGG					
CCCUCU	58.15 ± 0.5	157.5 ± 2	8.52 ± 0.02 (7.7)[c]	52.5	57.15 ± 2	154.4 ± 6	8.51 ± 0.06	52.7
CCCUCC	65.54 ± 1	175.5 ± 3	10.22 ± 0.05 (8.8)[c]	60.3	64.66 ± 3	172.8 ± 9	10.22 ± 0.2	60.5
			Binding to GGAGGGAAA					
CCCUCU	69.43 ± 0.8	188.2 ± 3	10.12 ± 0.04 (9.4)[c]	58.7	63.37 ± 0.9	169.8 ± 3	9.85 ± 0.09	58.8
CCCUCC	77.98 ± 0.7	209.4 ± 2	12.00 ± 0.05 (10.5)[c]	65.3	73.21 ± 2	195.2 ± 6	11.71 ± 0.1	65.4
CUCUCU	65.61 ± 1	187.7 ± 4	6.46 ± 0.02 (6.7)[c]	41.2	59.60 ± 1	168.4 ± 5	6.54 ± 0.1	41.5
CCCUd(C)C	74.24 ± 0.9	199.1 ± 3	11.49 ± 0.05	64.1	71.09 ± 1	189.7 ± 3	11.31 ± 0.07	64.2
CCCd(U)CC	73.23 ± 1	197.4 ± 4	11.01 ± 0.06	62.0	69.68 ± 1	186.7 ± 3	10.83 ± 0.05	62.2
CCCd(UC)C	76.71 ± 1	208.7 ± 3	10.96 ± 0.06	60.8	69.74 ± 1	187.6 ± 4	10.61 ± 0.09	61.0

[a] Buffer: 50 mM Tris (pH 7.6, 22 °C), 10 mM NaCl, 10 mM $MgCl_2$, and 0.1 mM Na_2EDTA. Additional significant figures are given to allow accurate calculation of T_M and other parameters. Melts of U(-1)P′ and C(-1)P′ with 90% pure 5′-GGAGGG-3′ in 10 mM sodium cacodylate (pH 7.0), 1 M NaCl, and 0.5 mM Na_2EDTA (no Mg^{2+}) gave the following results: for U(-1)P′, $-\Delta H° = 56.81$ kcal/mol, $-\Delta S° = 153.4$ eu, $-\Delta G°_{42°C} = 8.47$ kcal/mol; for C(-1)P′, $-\Delta H° = 63.08$ kcal/mol, $-\Delta S° = 168.6$ eu, $-\Delta G°_{42°C} = 9.95$ kcal/mol. [b] Calculated for $C_T = 0.1$ mM. [c] Values in parentheses are predicted from the nearest-neighbor model (Turner et al., 1988). Predictions for binding to 5′-GGAGGGAAA-3′ include −1.7 kcal/mol for 3′ unpaired AAA.

were doubled with no detectable increase in the reaction rate, demonstrating that the reactions were saturating with respect to ribozyme and GMP binding. Under these conditions, the observed rates represent k_{cat} for the conversion of E·P·G to products P′ and P″. Reported values are initial reaction rates obtained by monitoring the first 20−30% of the reaction. It was not possible to measure rates over several half-lives because P′ is readily cut to produce P″.

RESULTS

Oligonucleotides That Form a G·U Pair Bind More Tightly to the Ribozyme Than Those That Form a G-C Pair. The initial experiments in this study compared binding to the ribozyme of two oligonucleotides, GGCCCUCU (U(-1)P) and GGCCCUCC (C(-1)P) (Table 1). [The U(-1)P oligonucleotide has the sequence of the last few nucleotides of the cleaved 5′-exon of the self-splicing RNA. The capital P indicates that the oligonucleotide is the 5′-terminal product of the ribozyme-catalyzed cleavage.] At 42 °C, the K_d of U(-1)P is 0.3 nM and the K_d of C(-1)P is 3.5 nM, indicating that C(-1)P binds 10-fold more weakly to the ribozyme. This was surprising because the base-pairing of C(-1)P to the ribozyme IGS should be stronger than that of U(-1)P due to the presence of a strong G-C pair in place of the normal G·U pair at the P1 terminus. With a full complement of base-pairing energy and 4 kcal/mol of tertiary binding energy at 42 °C, the C(-1)P oligonucleotide would be expected to bind with a K_d of 6 pM.

The effect of the U to C substitution was also measured by equilibrium dialysis. These experiments used CCCUCU (U(-1)P′) and CCCUCC (C(-1)P′) to avoid potential aggregation or adsorption due to the GG tails of U(-1)P and C(-1)P. The GG tails are not expected to have a significant effect on binding to L-21 *Sca*I, because at 42 °C a terminal GA mismatch in this neighboring base-pair context is only 0.1 kcal/mol more favorable than a terminal unpaired A (Freier et al., 1986; Petersheim & Turner, 1983). The equilibrium dialysis results are also listed in Table 1, and they show that U(-1)P′ binds 30-fold more tightly than C(-1)P′. Thus, the U to C substitution in the oligonucleotide results in similarly weaker binding to ribozyme when assayed by either equilibrium dialysis or gel shift. There is also reasonable agreement with the absolute values of K_d from gel shift. The presence of 5 mM dGMP in the equilibrium dialysis experiments is expected to enhance the binding of U(-1)P′ and CUCUCU, due to cooperativity. For example, at 15 °C, CUCU labeled with pyrene at the 5′-end binds 4-fold more tightly in the presence of 15 mM dGMP (Bevilacqua et al., 1992, 1993).

For comparison, the G-C pair at the -5 position of U(-1)P was changed to a G·U pair, which is found naturally in the *Tetrahymena* intron but is not conserved among other group I introns. In this case, the oligonucleotide that forms a G·U pair binds more weakly to the ribozyme, as expected from base-pairing rules. Specifically, changing U(-1)P to U(-1,-5)P decreases binding to the ribozyme by 70-fold (22/0.3, Table 1), in contrast to the tighter binding seen when C(-1)P is changed to U(-1)P. Thus, not all G·U pairs enhance binding to the ribozyme.

The difference in base-pairing energy between terminal G·U and G-C pairs was measured by optical melting of U(-1)P′ and C(-1)P′ bound to GGAGGG and to GGAGGGAAA, which are oligonucleotide mimics of the IGS of the ribozyme (see Figure 1). As expected, binding to a guide-sequence oligonucleotide by C(-1)P′ is 1.7−1.9 kcal/mol tighter than binding by U(-1)P′, which forms a weaker G·U pair ($\Delta G°_{42°C}$ values, first 4 rows of Table 2). In fact, C(-1)P′ binds GGAGGGAAA as tightly as it binds ribozyme. [About 1.8 kcal/mol of C(-1)P′ binding free energy can be attributed to the AAA tail of the IGS mimic, which is somewhat more than the 1.1 kcal/mol expected from a single 3′ unpaired A (Turner et al., 1988). In the ribozyme, it is likely that at least the first A of the sequence A28-A29-A30 is available, since it is used to bind an oligopyrimidine for intron cyclization (Been & Cech, 1987) and circle opening (Sugimoto et al., 1988). The first A of the AAA sequence is also modified by dimethyl sulfate (Banerjee et al., 1993).]

The G-C Pair Affects 2′-OH Tertiary Interactions with the P1 Duplex. The equivalence between the binding of C(-1)P to ribozyme and to GGAGGGAAA was most easily

Table 3: Kinetic Characterization of C(-1)S and C(-1)P Cleavage

base pair	$(k_{cat}/K_m)^G$ (10^4 M^{-1} min^{-1})	k_{cat} (min^{-1})	K_m^G (μM)	infidelity $k_{cat}^{P'}$ (10^{-4} min^{-1})	infidelity $k_{cat}^{P''}$ (10^{-4} min^{-1})
GU	15 ± 3	13[a]	90 ± 7 (P1 docked) 360 ± 30 (P1 undocked)	1.1 ± 0.2	10.3 ± 0.5
GC	0.15 ± 0.02	0.56 ± 0.02	310 ± 40	96 ± 11	580 ± 50

[a] Estimated value for the maximum rate of U(-1)S cleavage at pH 7.0 and 30 °C. The value is based upon the maximum rate of cleavage at pH 5.5 and 30 °C and extrapolated to pH 7.0 assuming a linear relationship between pH and the log of the rate of cleavage (Herschlag et al., 1993b; Herschlag & Khosla, 1994). For cleavage of U(-1)S throughout this pH range, k_{cat} is equal to k_c, which is the rate constant for the chemical transition.

interpreted by a model in which the C(-1)P•ribozyme complex was stabilized by base-pairing only, without the 2′-OH tertiary interactions characteristic of the U(-1)P•ribozyme complex. It was also possible, however, that the 2′-OH interactions were still present, but were offset by the loss of interactions specific to the G•U pair; i.e., the correspondence of the ribozyme binding and duplex formation energies might be coincidental. It therefore was important to evaluate directly the 2′-OH tertiary interactions of the C(-1)P•ribozyme complex.

To address this issue, C(-1)P chimeras containing single and double deoxynucleotide substitutions were synthesized. If 2′-OH tertiary interactions were still formed upon C(-1)P binding to ribozyme, then substitution with deoxy at the -3 position would be expected to decrease binding to the ribozyme by 33-fold (Pyle & Cech, 1991; Pyle et al., 1992). Instead, a factor of 9 decrease was observed (30/3.5, Table 1), which is similar to the factor of 5 decrease observed in binding to the oligonucleotide, GGAGGGAAA (17/3.5, Table 1). In addition, if 2′-OH tertiary interactions were retained, oligonucleotide C(-1)Pd(-3) would bind at least an order of magnitude more weakly than C(-1)Pd(-4), because a -4 deoxynucleotide does not significantly disrupt tertiary interactions (Pyle & Cech, 1991; Bevilacqua & Turner, 1991; Herschlag et al., 1993a). Only a 3-fold difference was observed. Binding studies showed that the chimeric oligonucleotides bind the ribozyme more weakly than C(-1)P itself (Table 1), but that they also form intrinsically weaker duplexes with the oligonucleotide mimics of the IGS (Table 2). Deoxynucleotide substitution at position -2 or -3 was 0.4−0.5 kcal/mol more destabilizing for the ribozyme•oligonucleotide complex than for duplex formation, which is suggestive of a small amount of residual tertiary interaction energy in the C(-1)P background. Results in the ribozyme binding column of Table 1 indicate that the location of the deoxynucleotide substitution makes no major difference in the binding of C(-1)P chimeric oligonucleotides, in contrast to previous results on U(-1)P (Pyle & Cech, 1991; Herschlag et al., 1993a) and CUCU (Bevilacqua & Turner, 1991). The small difference in the binding strengths of C(-1)Pd(-4) and C(-1)Pd(-3) indicates that a trace of interaction may exist at the -3 position, but if there is any interaction it is significantly weaker than that in the case of U(-1)P. Thus, the results are consistent with reduced 2′-OH tertiary interactions when the conserved G•U pair is changed to G-C.

The G-C Pair Decreases Reaction Rate and Fidelity. In addition to ground-state analyses of the G-C substitution, kinetic studies were used to evaluate the effects of a G-C pair on reactivity. Under single-turnover conditions, $(k_{cat}/K_m)^G$ represents the second-order rate constant for the reaction of E•S (substrate GGCCCUCCAAAAA fully bound to the ribozyme but the G site unoccupied) with free G. The kinetic constants, k_{cat} and K_m^G, can be determined individually by measuring reaction rates from low to saturating G concentrations (Figure 2). Previous work has demonstrated that, for a G•U pair, k_{cat} is equal to the rate constant for the chemical step, k_c, and that K_m^G is equal to K_d^G (Herschlag & Cech, 1990; McConnell et al., 1993). Of particular interest to this study, K_d^G is sensitive to docking of the substrate helix in the active site. If the helix is in an undocked conformation, K_d^G is 4-fold weaker than if the helix is docked (McConnell et al., 1993), due to energetic coupling between G and substrate binding. If substitution of a G-C pair at the cleavage site results in undocking of the substrate helix, then coupling between G and substrate binding should be lost. Substitution of a G-C pair at the cleavage site reduces $(k_{cat}/K_m)^G$ by 100-fold at pH 7.0 (Table 3). This is the result of an approximately 25-fold reduction in k_{cat} and a 4-fold increase in K_m^G (Table 3). The value of K_m^G suggests that guanosine and substrate binding are energetically uncoupled, as would be expected if the substrate helix were not docked when there is a G-C pair at the cleavage site.

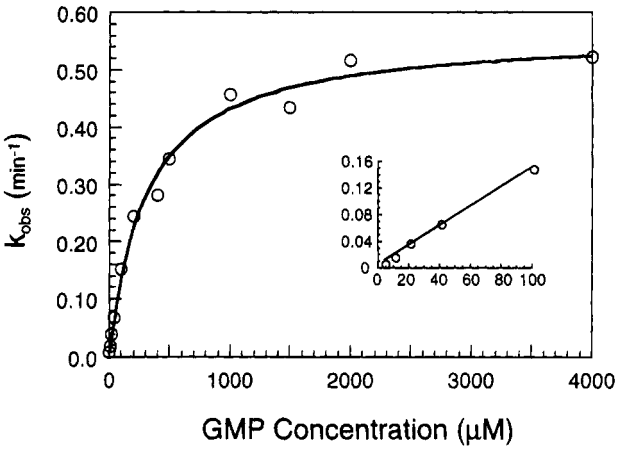

FIGURE 2: Measurement of $(k_{cat}/K_m)^G$, K_m^G, and k_{cat} for C(-1)S cleavage. Each line represents the best nonlinear least-squares fit of the data to the equation $k_{obs} = k_{cat}[G]/\{K_d^G + [G]\}$, solving for k_{cat} and K_d^G (McConnell et al., 1993). The inset is a plot of a subset of the data at low GMP concentrations (5−100 μM). The slope of the line is equal to $(k_{cat}/K_m)^G$.

Another assay that assesses the docking of the substrate helix is cleavage infidelity (Herschlag, 1992; Strobel & Cech, 1994). Although the ribozyme cleaves the substrate GGCCCUCU↓AAAAA (U(-1)S) primarily at U(-1) (arrow), cleavage at C(-2) and U(-3) is also observed (P′ and P″ in Figure 1). Previous studies have demonstrated that disruption of 2′-OH's important for docking results in a 30−80-fold higher rate of miscleavage (Herschlag, 1992; Strobel & Cech, 1994). Infidelity results from translocation of the substrate helix into other binding registers. This places different phosphodiesters in the ribozyme active site and results in

substrate miscleavage. If the U to C mutation had an effect on substrate helix docking, there should be a higher level of miscleavage at both the C(-2) and U(-3) positions. Miscleavage rates were measured for C(-1)P, and as predicted, miscleavage at C(-2) and U(-3) is 87- and 56-fold, respectively, higher than for U(-1)P (Table 3). This provides another line of evidence that the C mutation at least partially disrupts P1 docking into the ribozyme active site.

DISCUSSION

The G·U pair at the 5′-splice site in the P1 helix is conserved among group I introns (Cech, 1988; Michel & Westhof, 1990). In no naturally occurring intron has the substitution of a G-C pair been reported. In this work, we have used a ribozyme derived from the *Tetrahymena* group I intron to investigate the consequences on binding, reactivity, and fidelity of replacing the conserved G·U with G-C. The G·U pair gives the tighter binding of a 5′-exon analog. Additionally, the reaction is faster and has a much greater tendency to take place at the proper site. All of the results can be explained by a model in which the equilibrium for docking of the P1 helix into the catalytic core becomes less favorable upon replacement of G·U with G-C.

The combination of ribozyme and duplex binding experiments gave a seemingly paradoxical result. Oligonucleotide C(-1)P′, which forms a terminal G-C pair, binds 20-fold more strongly than U(-1)P′ to oligonucleotide mimics of the IGS, GGAGGGAAA or GGAGGG. C(-1)P, however, binds to the IGS sequence in the ribozyme 10-fold more weakly than does U(-1)P. Thus, C(-1)P binds 200-fold more weakly than expected. This indicates that the terminal G-C pair is incompatible with tertiary interactions associated with P1 docking, or at least derives little net binding energy from such interactions. From Table 1, the ΔG° associated with tertiary interactions can be calculated as

$$\Delta G^\circ_{3^\circ} = \Delta G^\circ_{\text{ribozyme}} - \Delta G^\circ_{2^\circ}$$

Here $\Delta G^\circ_{\text{ribozyme}}$ and $\Delta G^\circ_{2^\circ}$ are the free energy changes for binding to the ribozyme and the oligonucleotide mimic, respectively. If the three adenosines between the IGS and the P2 helix of the ribozyme are free to stack on the end of the P1 helix (see Figure 1), then the ΔG° measured from melts with GGAGGGAAA can be used as $\Delta G^\circ_{2^\circ}$, yielding (Table 1)

$$\Delta G^\circ_{3^\circ}(\text{GGCCCUCU}) = -3.4 \text{ kcal/mol}$$

$$\Delta G^\circ_{3^\circ}(\text{GGCCCUCC}) = 0.0 \text{ kcal/mol}$$

Thus, oligonucleotide C(-1)P would appear to derive no net ground-state energy from docking. If, however, only the first adenosine (A28) in the ribozyme stacks on the P1 helix, then a better estimate of $\Delta G^\circ_{2^\circ}$ would be the ΔG° measured from melts with GGAGGG, made more favorable by 1.2 kcal/mol for the stabilizing effect of a single terminal GA mismatch (Freier et al., 1986). This would give

$$\Delta G^\circ_{3^\circ}(\text{GGCCCUCU}) = -4.0 \text{ kcal/mol}$$

$$\Delta G^\circ_{3^\circ}(\text{GGCCCUCC}) = -0.8 \text{ kcal/mol}$$

This is still a loss of over 3 kcal/mol in tertiary interaction energy.

The decrease in tertiary interaction energy could result from two possible effects, which are not mutually exclusive. G·U may provide tertiary contacts that favor docking and these are absent with G-C, and/or the shape of G·U may be compatible with the docking site but G-C is not and therefore has unfavorable interactions. Either way, one function of the distinctive G·U pair at the cleavage site is to provide a sequence marker that allows the formation of 2′-OH group interactions that help staple the P1 helix into place. This leads to an increase in the efficiency and the accuracy of cleavage at the 5′-splice site. The loss of 3 kcal/mol in tertiary interactions for the U to C substitution at the -1 position contrasts with the loss of 0.9 kcal/mol for the U to C substitution at position -5 (Table 1), providing an additional indication that the G·U at the cleavage site is unique.

Previous studies demonstrated that disruption of 2′-OH tertiary contacts important for docking the substrate helix into the active site results in a 30−80-fold higher rate of miscleavage (Herschlag, 1992; Strobel & Cech, 1994). The dramatic increase in miscleavage observed for the C(-1)S substrate indicates that a U to C change at the cleavage site dislodges tertiary interactions that stabilize P1 docking. The fact that C(-1)S is still cut significantly more frequently at C(-1) than at C(-2) or U(-3) suggests that other factors also contribute to the establishment of a proper docking register. These include the single-stranded region between P1 and P2 (Young et al., 1991) and the interface between P2 and P2.1, which are suggested to be helically stacked (Downs & Cech, 1990, 1994). Previous studies on the intron form of this ribozyme have demonstrated significant reaction infidelity upon changing the purine−pyrimidine alignment of the base pair, i.e., G·U to U-A, C-G, or U·G (Barfod & Cech, 1989; Downs & Cech, 1994). This study now demonstrates that the more conservative change of G·U to G-C also increases the level of miscleavage.

It is somewhat surprising that the 2′-OH's do not appear to make much of a thermodynamic contribution to C(-1)P binding (beyond their modest contribution to duplex stability), considering that previous experiments clearly demonstrated that 2′-OH's are important for docking in other cleavage registers (Herschlag, 1992; Strobel & Cech 1994). For cleavage at U(-1), the normal 5′-splice site, 2′-OH's at positions -3, 22, and 25 are important for substrate helix docking. For miscleavage at C(-2), which forms a base pair with G23, a commensurate set of 2′-OH's at base positions -4, 23, and 26 serve to dock the helix. If docking the helix to each cleavage register involves a set of 2′-OH groups making tertiary contacts with the active site, then why after mutation of the 5′-cleavage site to a G-C pair do none of the 2′-OH's at positions -1 through -5 appear to contribute tertiary binding energy? This apparent discrepancy is readily explained by the fact that the two assays measure different functional states of the ribozyme. Thermodynamic data measure the equilibrium ground state for the P1 helix. For the G-C mutation, the ground state may be the undocked conformation, where the 2′-OH's do not make a contribution. Cleavage requires helix docking and utilizes the 2′-OH binding energy. The contribution made by the hydroxyls appears in k_{cat} rather than K_d because the docked configuration does not contribute significantly to the ground-state equilibrium. Therefore, for both the case of mutation of the 5′-splice site to a G-C pair and miscleavage at the G-C pair at C(-2), the 2′-OH contribution to docking is a transition-

state rather than a ground-state stabilization.

Several kinetic assays indirectly measure the occupancy of the substrate helix in the active site. In agreement with the thermodynamic data, the kinetic results are consistent with the conclusion that docking of the substrate helix with the G·U pair at the cleavage site is more favorable than with G-C. Substitution of a G-C for a G·U pair results in a 4-fold increase in K_m^G. By analogy with previous work (McConnell et al., 1993), this suggests that guanosine and substrate binding have become energetically uncoupled, consistent with improper docking when there is a G-C pair at the cleavage site. The 25-fold reduction in k_{cat} is also consistent with a lower frequency of active site occupancy due to undocking of the substrate helix, although other factors could contribute.

Taken together, our data provide a rationale for the fact that a G-C pair has not been observed in place of the conserved G·U at the 5′-splice site of group I introns. A G-C pair at the 5′-splice site makes the docking of P1 less favorable, which has a number of functional consequences. These include the following: (i) The maximum rate of the chemical step is 25-fold slower for a G-C than for a G·U pair. This is likely the result of the less favorable P1 docking equilibrium and the misalignment of G·U functional groups that are important for chemistry (S. A. Strobel and T. R. Cech, manuscript in preparation). (ii) Less favorable docking of the substrate duplex results in the uncoupling of substrate and GMP binding. This might affect the ability of the intron to distinguish between the first and second steps of splicing (McConnell et al., 1993; Bevilacqua et al., 1993). (iii) The G-C mutation disrupts the proper docking register of the substrate helix, resulting in 50−90-fold more miscleavage within the substrate helix. The observation that such a variety of critical intron functions are affected by this mutation suggests why the G·U pair is so highly conserved.

It is interesting to speculate on the particular aspects of the G·U pair that are important for docking. Crystallographic and NMR studies have indicated that G·U pairs introduce only slight helical distortions when introduced into a duplex (Quigley & Rich, 1976; Holbrook et al., 1991; White et al., 1992). G·U pairs, however, create a unique site in the minor groove of RNA by exposing the exocyclic amine of G to solvent. This exposed amine is uniquely accessible for interaction with hydrogen-bond-acceptor groups. In a G·U pair, the free exocyclic amine is also a site for specific water coordination, as seen in the crystal structures of $tRNA^{Phe}$ and of a duplex RNA containing several mispairs (Holbrook et al., 1978, 1991). The G·U pair may therefore play an essential role in providing determinants for specific interaction with the minor groove of RNA duplexes. These characteristics contribute to the importance of G·U pairs in $tRNA^{Ala}$ recognition by its cognate synthetase (Musier-Forsyth et al., 1991). In the present study, mutation of the G·U to G-C shifted the N2 exocyclic amine into a hydrogen-bonding register with the carbonyl of C (see Figure 1). This would completely change the environment of the minor groove and possibly disrupt tertiary contacts with the ribozyme catalytic center or ordered contacts with solvent. The conservation of the G·U pair might therefore reflect the need of the ribozyme to utilize the exocyclic amine protruding into the minor groove for ground-state tertiary contact and for chemical transition-state stabilization (S. A. Strobel and T. R. Cech, manuscript in preparation).

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.

Been, M. D., & Cech, T. R. (1986) *Cell 47*, 207−216.

Been, M. D., & Cech, T. R. (1987) *Cell 50*, 951−961.

Bernasconi, C. F. (1976) *Relaxation kinetics*, Academic Press, New York.

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.

Bevilacqua, P. C., Li, Y., & Turner, D. H. (1994) *Biochemistry 33*, 11340−11348.

Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., Vol. I, p 589, CRC Press, Cleveland, OH.

Borer, P. N., Dengler, B., Tinoco, I., Jr., & Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843−853.

Cech, T. R. (1988) *Gene 73*, 259−271.

Cech, T. R., Damberger, S. H., & Gutell, R. R. (1994) *Nat. Struct. Biol. 1*, 273−280.

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.

Fersht, A. (1985) *Enzyme Structure and Mechanism*, W. H. Freeman, New York.

Freier, S. M., Kierzek, R., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Biochemistry 25*, 3209−3213.

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

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

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

Herschlag, D., Piccirilli, J., & Cech, T. R. (1991) *Biochemistry 30*, 4844−4854.

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., Sussman, J. L., Warrant, R. W., & Kim, S.-H. (1978) *J. Mol. Biol. 123*, 631.

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

Longfellow, C. E., Kierzek, R., & Turner, D. H. (1990) *Biochemistry 29*, 278−285.

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) *Nat. Struct. Biol. 1*, 5−7.

Moran, S., Kierzek, R., & Turner, D. H. (1993) *Biochemistry 32*, 5247−5256.

Musier-Forsyth, K., Usman, N., Scaringe, S., Doudna, J. A., Green, R., & Schimmel, P. (1991) *Science 253*, 784−786.

Petersheim, M., & Turner, D. H. (1983) *Biochemistry 22*, 5247−5256.

Pyle, A. M., & Cech, T. R. (1991) *Nature 350*, 628−631.

Pyle, A. M., McSwiggen, J. S., & Cech, T. R. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8187−8191.
Pyle, A. M., Murphy, F. L., & Cech, T. R. (1992) *Nature 358*, 123−128.
Quigley, G. J., & Rich, A. (1976) *Science 194*, 796−806.
Richards, E. G. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., Vol. I, p 597, CRC Press, Cleveland, OH.
Salvo, J. L., & Belfort, M. (1992) *J. Biol. Chem. 267*, 2845−2848.
SantaLucia, J., Kierzek, R., & Turner, D. H. (1991) *J. Am. Chem. Soc. 113*, 4313−4322.
Scaringe, S. A., Francklyn, C., & Usman, N. (1990) *Nucleic Acids Res. 18*, 5433−5441.
Strobel, S. A., & Cech, T. R. (1993) *Biochemistry 32*, 13593−13604.
Strobel, S. A., & Cech, T. R. (1994) *Nat. Struct. Biol. 1*, 13−17.
Sugimoto, N., Kierzek, R., & Turner, D. H. (1988) *Biochemistry 27*, 6384−6392.
Turner, D. H., Sugimoto, N., & Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.
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.
White, S. A., Nilges, M., Huang, A., Brunger, A. T., & Moore, P. B. (1992) *Biochemistry 31*, 1610−1621.
Young, B., Herschlag, D., & Cech, T. R. (1991) *Cell 67*, 1007−1019.
Zaug, A. J., & Cech, T. R. (1986) *Science 231*, 470−475.
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.