Binding of Guanosine and 3′ Splice Site Analogues to a Group I Ribozyme: Interactions with Functional Groups of Guanosine and with Additional Nucleotides†

Sean Moran,‡ Ryszard Kierzek,§ and Douglas H. Turner*,‡

Department of Chemistry, University of Rochester, Rochester, New York 14627-0216, and Institute of Bioorganic Chemistry, Polish Academy of Sciences, 60-704 Poznan, Noskowskiego 12/14, Poland

Received September 17, 1992; Revised Manuscript Received February 2, 1993

ABSTRACT: Dissociation constants, K_d, were measured by equilibrium dialysis at 5 °C for a series of substrates binding to the L-21 *Sca*I ribozyme derived from the *Tetrahymena thermophila* self-splicing large subunit (LSU) ribosomal RNA intron. These substrates are analogues for the 3′ exon splice site, the cyclization site, and the exogenous G that initiates group I splicing. UCG has a K_d of 17 μM. Lengthening the substrate to GUCG and GGUCG enhances binding but by less than expected from potential base pairing. Functional groups on the 3′-terminal G of GUCG were replaced with H to test their effect on binding. GUC(2′dG) binds slightly tighter than the all-ribose molecule but shows no reactivity as a substrate. GUC(3′dG) binds weaker than GUCG. Inosine and 2-aminopurine ribonucleoside at the 3′ position weaken binding by 16- and 26-fold, respectively, but both tetramers are reactive. Thus hydrogen bonds to Watson–Crick pairing positions of the 3′G of GUCG contribute 1–2 kcal/mol to the free energy change for binding. Similar results are found in comparisons of UCG with UC(2′dG), UC(3′dG), and UCI. The nonreactive substrate GUCdGA includes a phosphodiester bond 3′ to the guanosine that is the site of chemistry for the all-ribose substrate GUCGA; GUCdGA binds 50 times more weakly than GUCdG. A similar result is obtained for GUCdGU. Competition experiments show that guanosine and guanosine 5′-monophosphate bind with dissociation constants of about 0.9 mM. The monomers 2′dG and 3′dG have K_d's of 0.5 and ≥3 mM, respectively. This suggests that sugar pucker and/or interactions with hydroxyl groups affect binding. Implications for ribozyme catalysis, splicing, cyclization, and design of antisense oligomers are discussed.

Except for Watson–Crick base pairing, little is known about the interactions important for recognition of RNA and for RNA folding. Knowledge of such interactions is important for predicting RNA structure (Turner et al., 1988) and for rational design of drugs targeted to RNA. For example, it has been suggested that a self-splicing group I intron in *Pneumocystis carinii* is a possible target for a therapeutic (Lin et al., 1992). Substrate binding by ribozymes provides an opportunity to study the interactions important for selective binding. For example, the self-splicing group I LSU rRNA intron of *Tetrahymena thermophila* initially chooses 2 out of ~6400 possible sites for cleavage and ligation (Bass & Cech, 1984; see Figure 1A). Selection of the 5′ cleavage site depends largely on base pairing to an internal guide sequence (IGS) (Davies et al., 1985; Been & Cech, 1986) but also on recognition of 2′-OH groups (Sugimoto et al., 1989; Pyle & Cech, 1991; Bevilacqua & Turner, 1991). This 5′ cleavage site is attacked by an exogenous guanosine, and several novel interactions have been proposed for the binding of this guanosine cofactor (Bass & Cech, 1984, 1986; Michel et al., 1989a; Yarus et al., 1991b). Selection of the 3′ splice site and subsequent positioning of a terminal G residue for a cyclization reaction are also partially dependent on recognition of G (Inoue et al., 1986; Price & Cech, 1988; Been & Perrotta, 1991). Exon ligation and cyclization are further aided by formation of base pairs with the two nucleotides immediately preceding the terminal G to form a pairing designated P9.0 (Burke, 1989; Michel et al., 1989a, 1990; Burke et al., 1990; see Figure 1). In this paper, the free energy contributions of some interactions contributing to the recognition of guanosine and 3′ splice and cyclization site analogues, in the presence of $d(CT)_3$ as an analogue of the 5′ exon, are studied by equilibrium dialysis (Uhlenbeck, 1972; Bevilacqua & Turner, 1991).

MATERIALS AND METHODS

Preparation of Ribozyme. Plasmid pT7L-21 (Zaug et al., 1988) was linearized by cutting with restriction endonuclease *Sca*I. T7 RNA polymerase (Davanloo et al., 1984) was used to transcribe cut plasmid following a modified procedure (12 mM Mg^{2+}, 10 mM DTT, 4 mM spermidine) of Grosshans and Cech (1989). The transcription mixture was extracted with phenol/$CHCl_3$, purified with a QIAGEN Tip-500 anion-exchange column, 2-propanol precipitated, and stored in 10 mM Tris (pH 7.5) and 0.1 mM EDTA at −20 °C. The purity of transcripts was checked by electrophoresis on 4% polyacrylamide/8 M urea gels with ethidium bromide visualization. The concentration of ribozyme was determined by measuring the absorbance at 260 nm and using $\epsilon_{260} = 3.2 \times 10^6\ M^{-1}\ cm^{-1}$ (Zaug et al., 1988).

Synthesis and Characterization of Substrates. Nucleoside phosphoramidites protected at the 2′-OH position with tetrahydropyranyl (thp) groups (Markiewicz et al., 1984) were used to synthesize oligomers either by hand (Kierzek et al., 1986) or on an Applied Biosystems 392 DNA/RNA synthesizer. When the oligomer GUCdGA was prepared with thp-blocked phosphoramidites, acid deblocking caused depurination at the dG residue. To prevent this, GUCdGA and related oligomers were synthesized with silyl protecting groups

† This work was supported by the Office of Naval Research and by National Institutes of Health Grant GM22939. D.H.T. is a Guggenheim Fellow.

* Author to whom correspondence should be addressed.
‡ University of Rochester.
§ Polish Academy of Sciences.

0006-2960/93/0432-5247$04.00/0

at the 2′-OH position (Usman et al., 1987). After complete deblocking, oligomers were purified on silica TLC plates (Baker Si500F) developed in 1-propanol/H_2O/NH_4OH (55:10:35). Concentrations were determined optically using ϵ_{260} values calculated from extinction coefficients for nearest neighbors (Borer, 1975; Richards, 1975). For oligomers containing inosine (I) and 2-aminopurine ribonucleoside (2AP), ϵ_{260} values were estimated by adding the extinction coefficients for the monomers I and 2AP to that calculated for the portion of the oligomer containing C, G, and U. Oligomers were radiolabeled with T4 polynucleotide kinase (New England Biolabs) and [γ-^{32}P]ATP (New England Nuclear), purified on 20% polyacrylamide/8 M urea gels, visualized by autoradiography, eluted with 0.5 M $NH_4C_2H_3O_2$, 0.1 mM EDTA, and 0.1% SDS, and dialyzed against 0.2 mM EDTA. For GUCdG, the presence of a deoxynucleotide at the 3′ end was confirmed by treating p*GUCG and p*GUCdG with $NaIO_4$ (Wells & Cantor, 1977; Agrawal et al., 1986). Gel electrophoresis of the reaction products showed altered mobility of the all-ribose tetramer (due to ring opening of the 3′-terminal ribose), while the mobility of the oligomer ending in deoxyribose was unchanged. The oligomers p*GUCG, p*GUCGA, and p*GUCdGA were partially hydrolyzed by heating to 90 °C in aqueous ammonia (pH ~12) for 1–6 min; the absence of a tetramer fragment from the p*GUCdGA oligomer proved the deoxyribose character of the fourth nucleotide from the 5′ terminus.

The labeled monomer p*G was prepared by kinasing Gp with [γ-^{32}P]ATP and removing the 3′-phosphate with P1 nuclease (Bass & Weintraub, 1989). Guanosine nucleosides and nucleotides and NpG's (N = A, C, and U) were obtained from Sigma. Monomer ddG was prepared by dephosphorylating ddGTP with alkaline phosphatase (calf intestinal, New England Biolabs), purified by TLC (see above), and desalted with a Sep-pak C-18 cartridge (Waters). To check for aggregation, imino proton NMR spectra of various G monomers were taken at 5 °C in 10 mM sodium phosphate buffer (pH 7.0), 5 mM $MgCl_2$, 135 mM NaCl, and 10% D_2O using a Varian VXR-500S spectrometer with a 1:$\bar{3}$:3:$\bar{1}$ pulse sequence to suppress water (Hore, 1983). Downfield-shifted peaks indicative of hydrogen-bonded imino protons (such as would occur upon formation of G quartets) were observed for 0.5 M pG but were absent from the spectra of 50 mM pG and 0.5 mM pG, G, dG (=2′dG), 3′dG, and ddG. NMR results were unchanged after a 3-day incubation at 5 °C.

NMR spectra of the nonexchangeable protons in 0.5 mM G, dG, and 3′dG at 10 °C were measured in D_2O containing dialysis buffer [without d(CT)$_3$] using a Bruker AMX 400-MHz spectrometer.

Equilibrium Dialysis. Experiments were performed at 5 and 15 °C with a Hoefer Scientific Instruments equilibrium microdialyzer, Model 101. Dialysis buffer was 50 mM HEPES (25 mM sodium salt, 25 mM free acid, pH 7.7), 5 mM $MgCl_2$, 135 mM NaCl, and 100 μM d(CT)$_3$. The all-deoxyribose analogue to the 5′ exon, d(CT)$_3$, was added to suppress side reactions (van der Horst & Inoue, 1993). Ribozyme was renatured by heating at 90 °C for 3 min in 10 mM Tris (pH 7.5) and 0.1 mM Na_2EDTA and centrifuging at room temperature for 10 s. After being cooled at room temperature for 15 min and chilled on ice for an additional 15 min, aliquots of ribozyme were combined with 10 μL of 5× dialysis buffer and the proper volume of sterile H_2O to give a 50-μL sample of a desired ribozyme concentration. These samples were then incubated for ~12 h at the desired temperature before dialysis was started. Samples of 3′ splice site analogues were prepared by mixing 10 μL of 5× dialysis buffer with 40 μL of a stock solution of the appropriate radiolabeled oligomer. For competition experiments, the appropriate amount of cold nucleoside or oligomer was added to both ribozyme and radiolabeled oligomer before incubation. Following incubation, samples (~50 μL) were injected into chambers separated by a dialysis membrane (Hoefer Scientific Instruments EMD 103) with a molecular weight (MW) cutoff of 12 000–14 000. This membrane allows passage of monomers and small oligomers (MW <2000) but prevents passage of ribozyme molecules (MW ~124 000). At various times, aliquots were taken from each side of the dialysis membrane, and R values [counts per minute (cpm) on ribozyme side divided by cpm on oligomer side] were measured either by direct scintillation counting of samples in 3 mL of Ecoscint A (National Diagnostic) or by electrophoresis of samples on a 20% polyacrylamide/8 M urea gel, visualization by autoradiography, and cutting and counting bands. In early experiments, R values were checked each day following the initiation of dialysis. Equilibration was judged to be complete after 3 days since there was no further increase in R. Equilibrium dialysis under conditions where [L-21 *Sca*I] $\gg$ [oligomer] allows determination of dissociation constants by fitting the R values and ribozyme concentrations (constrained to pass through the origin) to

$$R - 1 = [\text{L-21 } \textit{Sca}\text{I}]/K_d \quad (1)$$

Reactions. Radiolabeled oligomers at ≤20 nM were incubated with 1 μM renatured L-21 *Sca*I ribozyme and 10 μM CUCUA$_3$ in dialysis buffer [without d(CT)$_3$] at either 5 or 15 °C. Aliquots were removed after specified time periods, and the reaction was quenched by addition of an equal volume of gel loading buffer (20 M formamide, 0.1% xylene cyanol, 0.1% bromphenol blue, 90 mM Tris–borate, 2 mM EDTA). Extent of reaction was measured by cutting out reactant and product bands after gel electrophoresis of samples (see above) and determining the radioactivity of each band by scintillation counting.

RESULTS

A Deoxy Analogue of the 5′ Substrate Is Needed To Prevent Side Reactions. Initial dialysis experiments with only 3′ cyclization site analogues revealed a number of unexpected reactions between radiolabeled oligomers and ribozyme. For instance, dialysis of p*GUCG for 3 days at 5 °C yields p*G, some radiolabeled oligomers longer than p*GUCG, and radiolabeled ribozyme. The monomer can be explained by hydrolysis of p*GUCG in a reaction similar to hydrolysis at the 3′ splice junction or circle opening of the intron (Inoue et al., 1986; Zaug et al., 1984). Alternatively, both the monomer and other reaction products could be generated by G-exchange reactions (Kay et al., 1988). For example, p*GUCG$_{OH}$ + p*GpUCG → p*GUCGpUCG + p*G$_{OH}$. Similar reactions between p*GUCG$_{OH}$ and the G-rich IGS would yield radiolabeled ribozyme with heterogeneous 5′ termini (Kay & Inoue, 1987). No attempt was made to characterize these reactions.

The oligomer d(CT)$_3$ was used as an inhibitor to prevent the side reactions. The dissociation constant of p*d(CT)$_3$ was measured as 8.4 μM at 5 °C in separate equilibrium dialysis experiments in the absence of other substrates. Accordingly, at 100 μM d(CT)$_3$ and 0.5–20 μM ribozyme, approximately 90% of the ribozyme has d(CT)$_3$ bound at 5 °C. This concentration of inhibitor is sufficient to reduce the side reactions to negligible amounts and therefore was included in all subsequent equilibrium dialysis experiments. It has been reported that 5′ exon analogues and rG bind indepen-

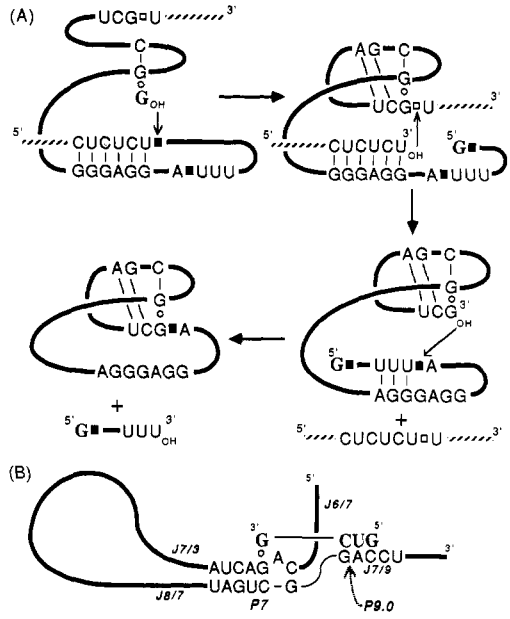

FIGURE 1: (A) Diagram outlining self-splicing and cyclization of the intervening sequence (IVS) of the LSU rRNA of *T. thermophila* (Cech, 1990). Bold lines represent the IVS sequence and striped lines the exon sequences. The solid squares represent phosphates subject to attack by guanosine in the first step of splicing and cyclization. The open square represents the phosphate where exon ligation takes place. The open circle is the base triple between guanosine and the G_{264}-C_{311} base pair of conserved secondary structure element P7. Hydrogen bonds between sequence elements involved in P1 and P9.0 are indicated by thin lines. (B) Local secondary structure around the G-binding site of the IVS. Bold letters indicate the 3′ splice and cyclization site analogue. Thin lines are phosphodiester linkages that are elongated to allow a two-dimensional representation of this area. Bold lines are continuations of the ribozyme sequence. Primary and secondary structure elements are labeled following the convention for group I introns (Burke et al., 1987).

dently (Pyle et al., 1990; Herschlag & Cech, 1990a). If this is also true for $d(CT)_3$ and the G-containing substrates discussed below, then the equilibrium dialysis experiments report simply the binding of the G-containing substrates. Preliminary stopped-flow experiments, however, suggest the binding of pyrene-labeled all-ribose CUCU and dG to be cooperative (Bevilacqua et al., 1992). If this applies to $d(CT)_3$ and the G-containing substrates, then equilibrium dialysis will slightly overestimate the dissociation constant in the presence of $d(CT)_3$ since not all ribozyme will have $d(CT)_3$ bound. For GUCG at 5 °C, equilibrium dialysis in the absence and presence of $d(CT)_3$ gave dissociation constants of 2.5 ± 0.5 and 1.8 ± 0.3 μM, respectively, consistent with small or negligible cooperativity.

UCG and d(TCG) Bind Ribozyme Tightly. The cyclization step that follows splicing involves recognition of UCG at the 3′ end of the intron (Figure 1A). Studies of an intermolecular guanosine addition reaction showed that the trimer UCG also attacks the cyclization site when the 3′-terminal guanosine of the intron has been removed (Tanner & Cech, 1987). Equilibrium dialysis with p*UCG and L-21 *Sca*I gives a K_d of 17 μM at 5 °C (Figure 2 and Table I). The all-deoxy analogue d(TCG) binds over 3 times more weakly with a K_d of 57 μM at 5 °C.

Addition of Extra Nucleotides at the 5′ End of UCG Increases Binding to Ribozyme. The native sequence at the 3′ end of the *T. thermophila* intron is $C_{411}U_{412}C_{413}G_{414}$. Gel electrophoresis of samples after equilibrium dialysis of p*CUCG and L-21 *Sca*I ribozyme revealed the presence of a ladder of radiolabeled oligomers longer than p*CUCG on

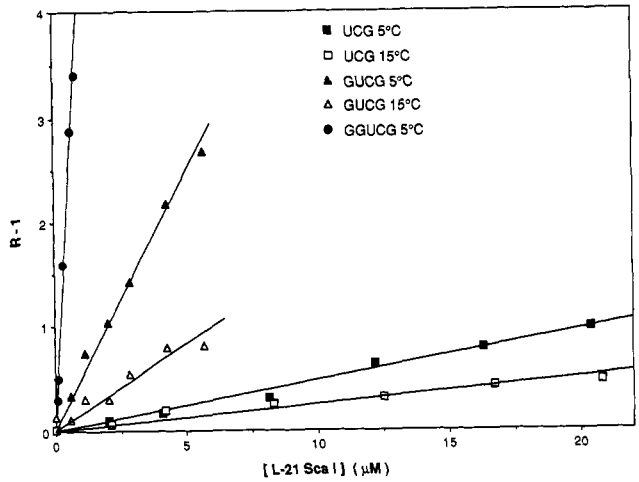

FIGURE 2: Plots of $R-1$ vs [L-21 *Sca*I] for the series of 3′ substrate analogues NUCG (N = no additional nucleotide, G, GG). Equilibrium dialysis conditions are 5 or 15 °C in 50 mM HEPES/25 mM Na^+ (pH 7.7), 5 mM $MgCl_2$, 135 mM NaCl, and 0.1 mM $d(CT)_3$. For $R-1=1$, [L-21 *Sca*I] = K_d.

the ribozyme side of the dialysis chamber. At 10 μM ribozyme, approximately 30% of the initial p*CUCG is consumed in 3 days. Cutting and counting bands of unreacted p*CUCG allowed calculation of R values. The resulting K_d for CUCG is 7 μM at 5 °C. This is only an estimate of K_d since the system is not at equilibrium due to the reactivity of p*CUCG. Calculations based on nearest neighbor parameters (Turner et al., 1988) suggest this K_d of 7 μM could arise from binding to the G site, to $G_{23}A_{24}G_{25}G_{26}$, and/or to $G_{406}A_{407}G_{408}U_{409}$.

Equilibrium dialysis with the fluorescent oligomer (2AP)-UCG gave a K_d of 5.6 μM at 5 °C. This factor of 3 tighter binding than UCG is close to the factor of 1.7 expected from adding a 5′ unpaired purine adjacent to a UA pair (Turner et al., 1988).

As shown in Figure 1B, C_{315} and C_{316} are directly 3′ to P9.0. Presumably, a substrate containing additional G residues could form GC pairs with C_{315} and C_{316}. Therefore, binding of GUCG and GGUCG was measured (Figure 2). GUCG and GGUCG bind roughly 9- and 60-fold more tightly than UCG (Table I).

It is possible that the ribozyme has more than one binding site for these oligomers. One way to check this is to obtain a Scatchard plot, which measures the moles of oligomer bound per mole of ribozyme (ν) as a function of free oligomer concentration, [oligo] (Scatchard, 1949; Cantor & Schimmel, 1980). For a set of independent binding sites, each with an identical K_d:

$$\nu/[\text{oligo}] = n/K_d - \nu/K_d \qquad (2)$$

Here n is the number of binding sites on the ribozyme. As the free oligomer concentration increases to infinity, all the sites on the ribozyme become saturated ($\nu = n$). If there is only one binding site on the ribozyme for the oligomer, then the Scatchard plot will have an X-intercept of 1. Figure 3 shows a Scatchard plot for GUCG binding to L-21 *Sca*I ribozyme. The X-intercept is 1.02 ± 0.06, suggesting only one binding site for GUCG. The K_d obtained from the slope of this plot is 2.6 μM, in good agreement with the K_d of 1.8 μM obtained from $R-1$ plots.

Atomic Mutation of the 3′G Residue of GUCG Reveals Effects of Functional Groups on Binding. Equilibrium dialysis of oligomers with 2′- or 3′-deoxyguanosine, inosine, or 2-aminopurine ribonucleoside replacing guanosine at the 3′-terminal position of GUCG shows the importance of various functional groups of the 3′G for interacting with the G-binding

Table I: Dissociation Constants[a] for Guanosine Monomers and 3′ Splice Site Analogues

(A) Effect of Adding Groups 5′ to Guanosine

monomer or 3′ splice site analogue	K_d (μM)	$-\Delta G°_5$ (kcal/mol)	exptl $\Delta\Delta G°_5$[b] (kcal/mol)	predicted $\Delta\Delta G°_5$[c] (kcal/mol)
G[d]	910 ± 330	3.87 ± 0.20		
pG[d]	820 ± 270	3.93 ± 0.18	−0.06 ± 0.27	
	[1500 ± 550][e]	[3.72 ± 0.21][e]		
CG[d]	230 ± 60	4.63 ± 0.15	−0.76 ± 0.25	
UCG	17 ± 4	6.07 ± 0.13	−1.44 ± 0.20	−3.4
	22 ± 2[d]	5.93 ± 0.05[d]		
	[36 ± 5][e]	[5.86 ± 0.08][e]		
CUCG	(7.0 ± 1.4)[f]	(6.56 ± 0.11)[f]	(−0.49 ± 0.17)[f]	0.6
(2AP)UCG	5.6 ± 0.3	6.68 ± 0.04	−0.61 ± 0.14	−0.3
GUCG	1.8 ± 0.5	7.31 ± 0.16	−1.24 ● 0.21	−2.9
	2.6 ± 0.4[g]	7.11 ± 0.09[g]		
	[6.6 ± 0.6][e]	[6.83 ± 0.06][e]		
GGUCG	0.28 ± 0.09	8.34 ± 0.18	−1.03 ± 0.24	−3.9

(B) Effect of Removing Functional Groups

monomer or 3′ splice site analogue	K_d (μM)	$-\Delta G°_5$ (kcal/mol)	$\Delta\Delta G°_5$[b] (kcal/mol)
G[d]	910 ± 330	3.87 ± 0.20	
dG[d]	470 ± 150	4.24 ± 0.18	0.37 ± 0.27
3′dG[d]	≥3000	≤3.2	≤−0.67
2′,3′ddG[d]	530 ± 120	4.17 ± 0.13	0.30 ± 0.24
pG[d]	820 ± 270	3.93 ± 0.18	
pdG[d]	710 ± 200	4.01 ± 0.16	0.08 ± 0.24
UCG	17 ± 4	6.07 ± 0.13	
UCI[d]	860 ± 350	3.90 ± 0.23	−2.17 ± 0.26
UCdG	5.9 ± 1.6	6.65 ± 0.15	0.58 ± 0.20
UC(3′dG)[d]	180 ± 20	4.77 ± 0.06	−1.30 ± 0.15
d(TCG)	57 ± 9	5.40 ± 0.09	−0.67 ± 0.16
GUCG	1.8 ± 0.5	7.31 ± 0.16	
GUCI	29 ± 6	5.77 ± 0.12	−1.54 ± 0.19
GUC(2AP)	47 ± 6	5.51 ± 0.07	−1.80 ± 0.17
GUCdG	0.96 ± 0.17	7.66 ± 0.10	0.35 ± 0.19
GUC(3′dG)	4.1 ± 0.6	6.86 ± 0.08	−0.45 ± 0.18
GUC[d]	(360 ± 190)[f]	(4.38 ± 0.29)[f]	(−2.93 ± 0.33)[f]

(C) Effect of Adding Groups 3′ to Guanosine

monomer or 3′ splice site analogue	K_d (μM)	$-\Delta G°_5$ (kcal/mol)	$\Delta\Delta G°_5$[b] (kcal/mol)
G[d]	910 ± 330	3.87 ± 0.20	
Gp[d]	1900 ± 750	3.46 ± 0.22	−0.41 ± 0.30
UCdG	5.9 ± 1.6	6.65 ± 0.15	
UCdGA	≥200	≤4.7	≤−2.0
UCdGU	≥200	≤4.7	≤−2.0
GUCdG	0.96 ± 0.17	7.66 ± 0.10	
GUCdGA	48 ± 7	5.50 ± 0.08	−2.16 ± 0.13
GUCdGU	32 ± 7	5.72 ± 0.12	−1.94 ± 0.16
d(TCG)	57 ± 9	5.40 ± 0.09	
d(TCGT)	≥300	≤4.5	≤−0.9

[a] Conditions for equilibrium dialysis experiments: 50 mM HEPES (25 mM free acid, 25 mM Na^+ salt, pH 7.7), 5 mM $MgCl_2$, 135 mM NaCl, and 100 μM $d(CT)_3$ at 5 °C unless otherwise noted. [b] $\Delta\Delta G°_5 = \Delta G°_5(N+1) - \Delta G°_5(N)$. [c] Free energy increment predicted at 5 °C for adding a base pair or 5′ dangling end to the A-form RNA duplex (Turner et al., 1988). [d] Competition experiment with unlabeled monomer or oligomer competing with p*GUCG. [e] Equilibrium dialysis experiment at 15 °C. [f] See text for discussion of uncertainties in determining K_d of CUCG and GUC. [g] Determined from a fit of the Scatchard plot (Figure 3). [h] $\Delta\Delta G°_5$ is relative to the underlined substrate at the head of each subgroup. Example: $\Delta\Delta G°_5(UCI) = \Delta G°_5(UCG) - \Delta G°_5(UCI)$.

site (Figure 4 and Table I). Replacing the 2′-OH with H decreases K_d by almost 50% (binding of GUCdG is tighter). A similar result is obtained with UCdG; it binds almost 3-fold more tightly than UCG. Replacing the NH_2 of the 3′G of GUCG with H weakens binding by a factor of 16. Replacing the O6 of G with H coupled with deletion of the imino proton weakens binding by a factor of 26. If the 3′G residue is removed completely, $R-1$ is ~0.1 at ribozyme concentrations

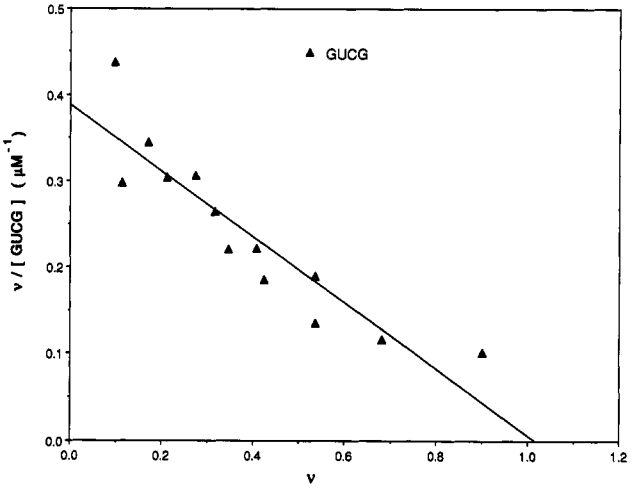

FIGURE 3: Scatchard plot of p*GUCG binding to the L-21 *Sca*I ribozyme at 5 °C. The buffer is the same as in Figure 2. ν is the moles of oligomer bound per mole of ribozyme, or fraction of saturation; [GUCG] is the concentration of free oligomer at equilibrium. An X-intercept of 1 indicates a single binding site.

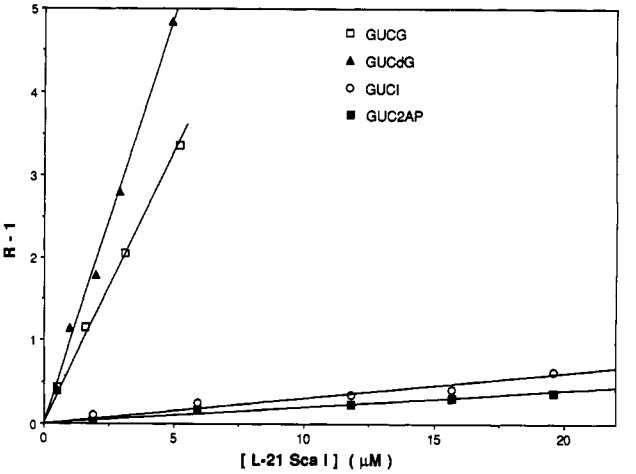

FIGURE 4: $R-1$ vs [L-21 *Sca*I] plots for the series of 3′ substrate analogues GUCX (X = G, dG, 2AP, I). Equilibrium dialysis conditions are 5 °C with same buffer as in Figure 2.

as high as 20 μM, giving a K_d of ~200 μM, more than a factor of 100 weaker than GUCG. As described below, K_d for GUC was also measured by competition with GUCG.

Addition of a Nucleotide at the 3′ End of GUCdG or d(TCG) Decreases Binding. To determine the effect of lengthening the substrate in the 3′ direction, K_d's were measured for GUCdGA and GUCdGU. Deoxyribose analogues were used because ribozyme-catalyzed hydrolysis (Inoue et al., 1986; Zaug et al., 1984) cleaves off the 3′-terminal nucleotide when the preceding guanine is attached to a ribose sugar. GUCdGA and GUCdGU bind more weakly than GUCdG by at least 30-fold. Attempts were also made to measure the binding of UCdGA, and UCdGU, and d(TCGT). All are too weak to measure directly. UCdGA and UCdGU bind at least 30-fold more weakly than UCdG; d(TCGT) binds at least 5-fold more weakly than d(TCG).

Competition Experiments Measure Dissociation Constants for Weakly Binding Substrates. When p*G is used as the radiolabeled substrate in equilibrium dialysis with a ribozyme concentration of ~20 μM, the measured $R-1$ values are close to zero. Control experiments that check equilibration of labeled oligomers across the dialysis membrane in the absence of ribozyme typically give $R-1 \leq 0.05$. From this, it is concluded that pG does not bind measurably to ribozyme at [L-21 *Sca*I] ≤20 μM. A lower limit can be placed on the

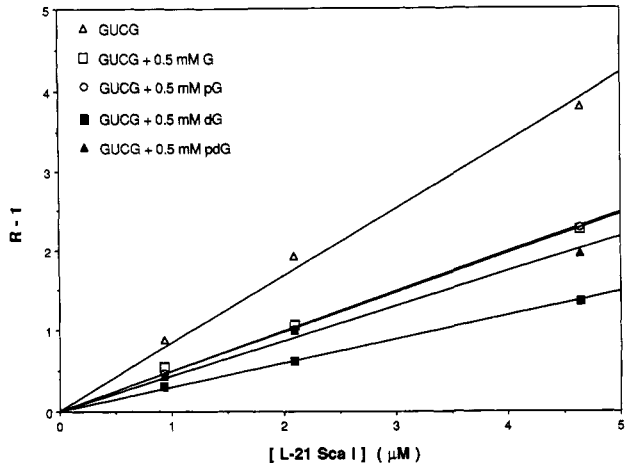

FIGURE 5: $R - 1$ vs [L-21 *Sca*I] plots for competition experiments at 5 °C in the buffer listed in the caption to Figure 2. Binding of G monomers to ribozyme reduces the effective concentration of ribozyme available to bind p*GUCG. From eq 5, the higher the apparent dissociation constant for p*GUCG in the presence of a fixed concentration of a given G monomer (i.e., the lower the slope of the $R - 1$ plot), the tighter the binding of the monomer to the ribozyme.

dissociation constant of pG: $K_d(pG) \geq [\text{L-21 } Sca\text{I}]/(R - 1) = 20\ \mu M/0.05 = 0.4$ mM. Since aggregation of L-21 *Sca*I becomes a problem at concentrations ≥20 μM, the binding constant of monomeric G and other weakly binding substrates is best measured by competition. Including a high concentration, [C], of unlabeled, weakly binding substrate in samples used for equilibrium dialysis with p*GUCG causes some of the ribozyme to bind to unlabeled substrate, lowering the amount of free ribozyme available to bind p*GUCG. The apparent dissociation constant, K_{app}, for p*GUCG is increased from that measured in the absence of competitor, K_d:

$$K_{app} = K_d(1 + [C]/K_c) \quad (3)$$

where K_c is the dissociation constant of the competing substrate (Uhlenbeck, 1972). Since only binding that competes with p*GUCG is observed, K_c is the dissociation constant for the competitor binding to the same site as p*GUCG.

For guanosine monomers, concentrations of 100 μM or greater are required to increase K_{app} for GUCG (Figure 5). The dissociation constants for various G nucleosides and nucleotides are listed in Table I. Guanosine, guanosine 5′-monophosphate, and 2′-deoxyguanosine 5′-monophosphate all bind with K_d's just under 1 mM; 2′-deoxyguanosine and dideoxyguanosine bind slightly tighter, and guanosine 3′-monophosphate and 3′-deoxyguanosine bind weaker. In control experiments, A, C, and U nucleosides and their 5′-monophosphates have no competitive effect on the binding of p*GUCG (data not shown).

The dimers AG, CG, and UG were also tested for competitive binding against p*GUCG. (GG was not assayed because of solubility problems.) CG binds about 4-fold tighter than pG, while AG and UG bind about the same as pG (data not shown).

Competition was also used to measure the effect of separately replacing the exocyclic amino group and 3′-OH of G in UCG with a hydrogen. UCI and UC(3′dG) bind more weakly than UCG by factors of 50 and 10, respectively (Table I). As a check of the competition method, K_c for UCG was measured in competition with p*GUCG. The value of 22 μM obtained is in good agreement with the directly measured K_d of 17 μM. This is consistent with UCG and GUCG having the same binding site.

When the trimer GUC is used in a competition experiment with p*GUCG, 20–60% of p*GUCG is elongated to products that appear to be five and six nucleotides long. Although these products were not characterized, they may be due to a reaction similar to the first step of splicing (p^*GUCG_{OH} + GUpC → $p^*GUCGpC + GU_{OH}$) and/or to a G-exchange reaction (p^*GUCG_{OH} + GpUC → $p^*GUCGpUC + G_{OH}$). Separation of the radiolabeled components in the dialysis cells by gel electrophoresis allows calculation of R values from the portion of p*GUCG remaining intact and yields an estimated dissociation constant of 360 μM for GUC (Table I). This value is consistent with the lower limit to K_d determined from direct equilibrium dialysis.

Substrates Bind Weaker at Higher Temperature. When the equilibrium dialysis temperature is raised to 15 °C, dissociation constants for pG, UCG, and GUCG increase (Table I) by factors of 2, 2, and 4, respectively, relative to 5 °C. This indicates that the $\Delta H°$ values for the binding of these molecules are negative. The $\Delta H°$ values obtained from the van't Hoff equation are −10.0 ± 7.9 kcal/mol for pG, −11.9 ± 4.7 kcal/mol for UCG, and −20.6 ± 4.2 kcal/mol for GUCG. The $\Delta H°$ values would be more reliable if a larger temperature interval could be used. Unfortunately, the side reactions observed for GUCG (and, to a lesser degree, UCG) proceed to a greater extent at 15 °C. Thus, similar equilibrium dialysis experiments are not feasible at higher temperatures.

The Preferred Sugar Pucker of Unbound 3′dG Is C3′-Endo, While That of G and dG Is C2′-Endo. Coupling constants from NMR spectra can be used to determine sugar puckers in solution (Westhof et al., 1977). The $J_{1'2'}$ and $J_{3'4'}$ coupling constants measured for G, dG, and 3′dG at 10 °C are listed in Table II. The percentage of each nucleotide in the C3′- or C2′-endo conformation can be calculated from these coupling constants and is listed in Table II. For rG and 3′dG, the $J_{1'2'}$ splitting gives the most reliable value, while for 2′dG, $J_{3'4'}$ is most reliable (Westhof et al., 1977). The free energy changes for interconversion of sugar puckers are also listed in Table II. The results are similar to those derived previously for A, dA, and 3′dA and are expected to be essentially independent of temperature (Westhof et al., 1977).

Reactivity Assays at 15 °C Indicate All Oligomers except GUCdG and GUCdGA Are Substrates for Ribozyme Reactions. Transesterification reactions that mimic the first step of self-splicing and the cyclization step test the ability of the 3′ splice site analogues to serve as substrates for ribozyme-catalyzed endonuclease reactions (Zaug et al., 1986):

$$p^*NG + CUCUA_3 \rightarrow p^*NGA_3 + CUCU \quad (4)$$

N = nothing, UC, GUC, GGUC

$$p^*GUCX + CUCUA_3 \rightarrow p^*GUCXA_3 + CUCU \quad (5)$$

X = G, dG, I, 2AP

Figure 6 shows an autoradiograph of typical reactions. For p*G at 15 °C, only 7% reaction is observed after 30 min; even at 2 h, the extent of reaction is just over 20%. In comparison, after only 30 min the extent of reactions for UCG, GUCG, and GGUCG is over 50%, with GUCG and GGUCG reacting slightly faster than UCG; at 2 h the reactions are essentially complete (extent of reaction ~90%). When the 3′-terminal G of GUCG is modified, reactivity decreases precipitously. GUCI and GUC(2AP) have reacted only 25% and 5%, respectively, after 12-h incubations. There is no detectable reaction for the GUCdG substrate after 12 h. This inactivity

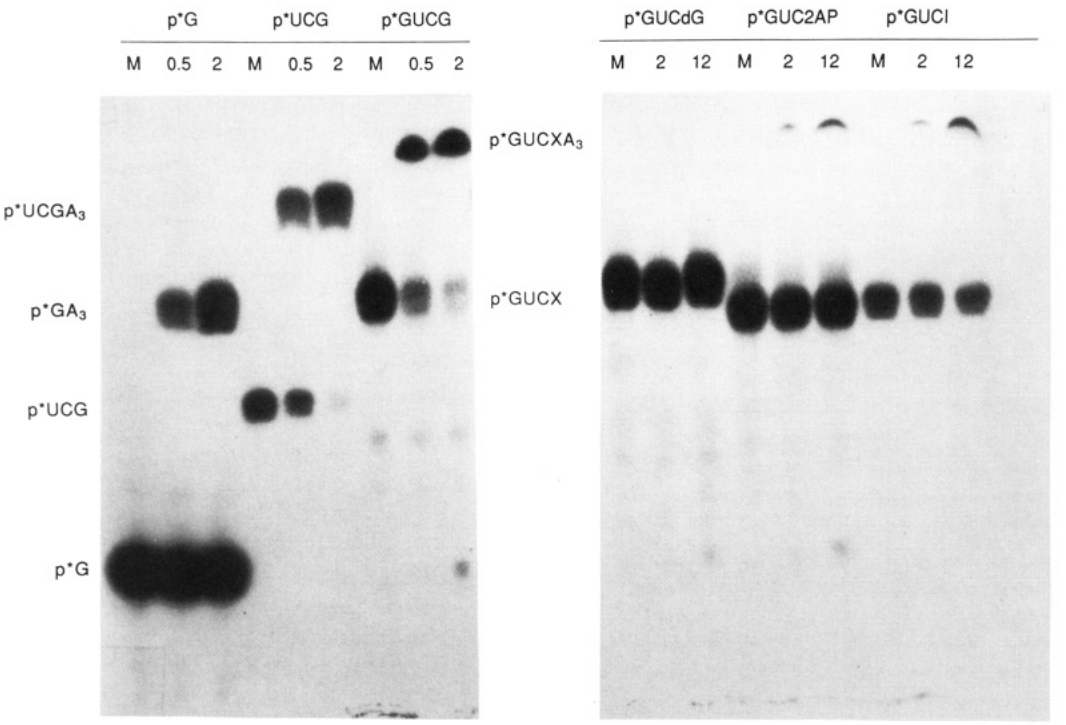

FIGURE 6: Autoradiogram of p*NUCX + $CUCUA_3$ (10 μM) reactions catalyzed by L-21 *Sca*I ribozyme (1 μM) at 15 °C. The buffer is the same as that used for equilibrium dialysis experiments except $d(CT)_3$ omitted. Radiolabeled 3′ substrates and time points (in hours) are listed above each set of reactions (M = marker). Substrate and product bands are labeled on the side.

Table II: *J* Coupling Constants for G Nucleosides[a]

nucleoside	$J_{1',2'}$ (Hz)	% C3′-endo sugar pucker[b]	ΔG °[c] (kcal/mol)	$J_{3',4'}$ (Hz)	% C3′-endo sugar pucker[b]	ΔG °[c] (kcal/mol)
rG	6.0	36	–0.3	3.5	37	–0.3
2′dG	6.7	29	–0.5	3.2	34	–0.4
3′dG	1.7	82	0.9	8.1	87	1.1
2′,3′ddG[d]	3.4	64	0.3	8.4	89	1.2

[a] From NMR spectra of 0.5 mM guanosine nucleosides in equilibrium dialysis buffer without $d(CT)_3$ at 10 °C. [b] Calculated using the procedure of Westhof et al. (1977). [c] $\Delta G° = -0.001\,987T \ln(\%\ C2'\text{-endo}/\%\ C3'\text{-endo})$. [d] Results at 21 °C from Jagannadh et al. (1991).

of deoxyguanosine also becomes apparent when the 3′ splice site analogues, GUCGA and GUCdGA, are tested for cleavage via hydrolysis or reaction with CUCU:

$$\text{p*GUCGA} + H_2O \rightarrow \text{p*GUCG} + \text{pA} \quad (6)$$

$$\text{p*GUCGA} + \text{CUCU} \rightarrow \text{p*GUCG} + \text{CUCUA} \quad (7)$$

For reactions 6 and 7 at 15 °C, more than 75% of p*GUCGA is converted to p*GUCG after only 2 h, while p*GUCdGA shows no reaction after 24 h (data not shown). Decreasing the temperature to 5 °C slows down the reactions for p*UCG and p*GUCG with $CUCUA_3$ (these were the only 3′ splice site analogues tested for reaction 4 at 5 °C; data not shown). After 30 min at 5 °C, 20% of UCG and 45% of GUCG are converted to product. After 2 h, the reactions are still not complete (60% reaction for p*UCG, 80% for p*GUCG), but it is apparent that p*GUCG is reacting roughly twice as fast as p*UCG.

DISCUSSION

Understanding the contributions of various interactions to folding, substrate binding, and catalysis by RNA is important for prediction and modeling of RNA structure (Turner et al., 1988) and for design of drugs and ribozymes. The equilibrium dialysis results presented above provide insight into the interactions used by a group I ribozyme to bind G substrate and the 3′ splice and cyclization site. The preliminary activity results suggest some interactions that are important in the transition state. In general, it appears that hydrogen-bonding groups can make contributions of 1–2 kcal/mol to binding. The contributions of base pairs, however, are less than predicted from model systems. Binding results for GUCdGA, GUCdGU, UCdGA, UCdGU, and d(TCGT) suggest that unfavorable interactions with the nucleotide following the splice and cyclization site may also be important for catalysis.

G Binds Relatively Weakly. The results presented in Table I provide the first equilibrium measurements of the dissociation constant for G binding to a group I ribozyme. At 5 °C, the K_d's for pG and G are about 0.9 mM. [Note that *Tetrahymena* live between 1 and 40 °C (Bick, 1972).] Evidently, a 5′-phosphate does not affect binding, just as it does not affect the reactivity of guanosine monomers in splicing (Bass & Cech, 1984). At 50 °C, K_d for G has been estimated from kinetics experiments as 0.3–1.1 mM (Herschlag & Cech, 1990b; Herschlag et al., 1991). If the change in K_d for pG in going from 5 to 15 °C is used to extrapolate to 50 °C using the van't Hoff equation, then the predicted K_d at 50 °C is about 10 mM. The difference between the kinetic estimate and the thermodynamic extrapolation could be due to the large uncertainty in $\Delta H°$ for pG (see Results), to the different buffer conditions for the experiments, or to the assumptions required for the kinetic estimate. Given current knowledge, the agreement between kinetic and thermodynamic results is reasonable.

At 5 °C, the K_d's of dG and pdG are 0.5 and 0.7 mM, respectively (Table I). Apparently, dG and pdG bind slightly tighter than their ribose counterparts, although the difference is within experimental error. These values are similar to the inhibition constant, K_I, of 1.1 mM measured for dG in the splicing reaction for LSU rRNA precursor at 30 °C (Bass & Cech, 1986). This is consistent with the dG binding site in

the L-21 *Sca*I ribozyme being the same as in the rRNA precursor.

While the K_d for dG binding to L-21 *Sca*I is similar to the K_I for dG in splicing, the extrapolated K_d's for G and pG are about 100–1000 fold larger than the "K_M's" of 20 and 5 μM, respectively, reported for the single turnover reactions of splicing at 30 °C (Bass & Cech, 1984) and of an intermolecular G-addition reaction at 42 °C (Tanner & Cech, 1987). (The quotation marks around K_M indicate that splicing and intermolecular G addition are single rather than multiple turnover reactions.) Herschlag and Cech (1990b) have pointed out that this could be due to "a G-independent conformational change that is rate-limiting at saturating but not subsaturating G". Another mechanism that is consistent with the data involves two effectively irreversible steps with the second irreversible step (k_3) rate-limiting and neither irreversible step attainable under the conditions used for equilibrium dialysis:

$$S{\cdot}E + G \overset{K_d}{\rightleftarrows} S{\cdot}E{\cdot}G \overset{k_2}{\rightarrow} (SEG)' \overset{k_3}{\rightarrow} E + P \qquad (8)$$

Here S is the substrate that mimics the 5′ splice site. For this mechanism with the first step in rapid equilibrium, $K_M = K_d k_3/(k_2 + k_3)$ (Fersht, 1985). Thus, if $k_3 \ll k_2$, then $K_M \ll K_d$. An example of such a mechanism is one in which (SEG)′ is a covalent intermediate. Another possibility is a mechanism in which both k_2 and k_3 are rates for covalent rearrangements of substrates. Note that the third step in the mechanism can represent a combination of steps.

Adding Potential Base Pairs 5′ of G Strengthens Binding but Not by as Much as Expected for Simple Base Pairing. Phylogenetic comparisons and site-directed mutagenesis indicate that U_{412} and C_{413}, which precede the terminal G_{414} of the LSU intron, pair with G_{313} and A_{314} to form P9.0. This aids in 3′ splice site selection and in positioning the 3′-terminal G of the intron for cyclization (see Figure 1) (Burke, 1989; Michel et al., 1989a, 1990; Burke et al., 1990). Enzymatic digestion and chemical modification studies of the circular form of the *T. thermophila* intron (Jaeger et al., 1990; Inoue & Cech, 1985) and of the L-21 *Sca*I ribozyme (Banerjee et al., 1993) indicate that G_{313} and A_{314} are unpaired. For the next two nucleotides, C_{315} and C_{316}, available data have been interpreted as indicating pairing to G_{405} and G_{406} (Michel & Westhof, 1990), no pairing (Young et al., 1991), or C_{315} unpaired and C_{316} paired to G_{331} (Burke et al., 1987). The effect of lengthening the G substrate to allow additional pairing was studied with the series CG, UCG, GUCG, and GGUCG (Figure 1B). The results in Table I indicate that these oligomers bind more tightly than G by factors of 4, 50, 500, and 3000, respectively. These correspond to free energy increments, $\Delta\Delta G°_5$ of −0.8, −1.4, −1.2, and −1.0, for the addition of C, U, G, and G, respectively. The $\Delta\Delta G°_5$'s expected for base pair formation by the U, G, and G in short oligomers are −3.4, −2.9, and −3.9 kcal/mol, respectively (Freier et al., 1986). Evidently, pairing in this binding site is not easily related to simple duplex formation. It is specific, however. AG and UG bind more weakly than CG; CUCG and (2AP)UCG bind more weakly than GUCG (Table I). The results suggest that constraints on the ribozyme nucleotides prevent their full base-pairing potential from being realized. Alternatively or in addition, formation of additional base pairs may alter the G-binding site, thus weakening binding interactions.

UCG and d(TCG) Bind Similarly. The K_d at 5 °C for the duplex $\overset{5'UCG3'}{3'CAGC5'}$ is predicted to be 1.5 mM (Freier et al., 1986; Turner et al., 1988), almost 100-fold weaker than the K_d for UCG binding to ribozyme. Similar enhanced binding has been observed for analogues of the 5′ splice site (Sullivan & Cech, 1985; Sugimoto et al., 1988, 1989; Bevilacqua & Turner, 1991). The magnitude of the extra favorable free energy for binding has been attributed partially (Bevilacqua & Turner, 1991) or completely (Pyle & Cech, 1991) to interactions with 2′-OH groups. Replacing UCG with d(TCG) weakens binding by only a factor of 3, or 0.2 kcal/mol of 2′-OH (Table I). Decreases of this magnitude and larger are observed in comparisons of duplex formation by RNA·RNA and RNA·DNA oligomers (Martin & Tinoco, 1980; Hall & McLaughlin, 1991; Bevilacqua & Turner, 1991). Thus it seems unlikely that interactions with 2′-OH groups are responsible for the tight binding of UCG. The relatively small free energy increments measured for addition of the U and C (see Table I) are also consistent with this interpretation. The enhanced binding appears to be due to the 3′G.

Hydrogen-Bonding Groups on G Can Each Contribute 1–2 kcal/mol to Binding. GUCG binds to L-21 *Sca*I 200-fold tighter than GUC, corresponding to a $\Delta\Delta G°_5$ of −2.9 kcal/mol for the 3′-terminal G (Table I). The origins of this binding energy were explored by deleting functional groups on the 3′G (Table I). The results are summarized in Figure 7, which also shows two proposed models for the G-binding site (Michel et al., 1989a; Yarus et al., 1991a,b). GUCdG and UCdG have somewhat tighter binding than GUCG and UCG, respectively, indicating the 2′-OH of G weakens ground-state binding. GUCI and UCI bind 16 and 50 times more weakly than GUCG and UCG, respectively, indicating that interactions with the amino group of the 3′G contribute at least 1.5 and 2.2 kcal/mol to the $\Delta G°$ for binding at 5 °C. The contribution of an amino group may be even larger than the measured $\Delta\Delta G°_5$ because its removal strengthens any hydrogen bond to the imino proton by lowering its pK_a (Santa Lucia et al., 1991). The difference between the $\Delta\Delta G°_5$'s for GUCG and UCG may indicate that the additional GC pair formed with GUCG affects the binding site for the 3′G. The fact that the $\Delta\Delta G°_5$ of −2.9 kcal/mol between GUC and GUCG is less favorable than the $\Delta G°_5$ of −3.9 kcal/mol for binding monomer pG also suggests a perturbation of the G-binding site. GUC(2AP) binds 26 times more weakly than GUCG, indicating that either the imino proton or the carbonyl group or a combination of both contribute 1.8 kcal/mol to binding.

The free energy increments measured for substitution of functional groups are consistent with models proposed for the G-binding site (Figure 7) and with suggested free energy increments for hydrogen bonds to bases in RNA (Freier et al., 1986; Turner et al., 1987; SantaLucia et al., 1991, 1992). The base triple model of Michel et al. (1989a) has single hydrogen bonds to the G amino and imino protons. For this model, the substitution results would indicate that each hydrogen bond contributes about 2 kcal/mol to binding. This is consistent with previous suggestions that an optimum hydrogen bond in a Watson–Crick pair can contribute 2 kcal/mol to binding (Turner et al., 1987). The axial III model of Yarus et al. (1991b) has two hydrogen bonds to the G amino group, one to the imino proton, and the equivalent of one to the carbonyl group. For this model, the substitution results would indicate that each hydrogen bond contributes about 1 kcal/mol to binding. This is within the range observed for hydrogen bonds in RNA (Freier et al., 1986; Turner et al., 1987; SantaLucia et al., 1992). The contributions of hydrogen bonds to folding and binding free energies are controversial (Fersht, 1987; Abeles et al., 1992; Lesser et al., 1990; Shirley et al., 1992; Williams et al., 1991; Williams, 1992; SantaLucia et al., 1992).

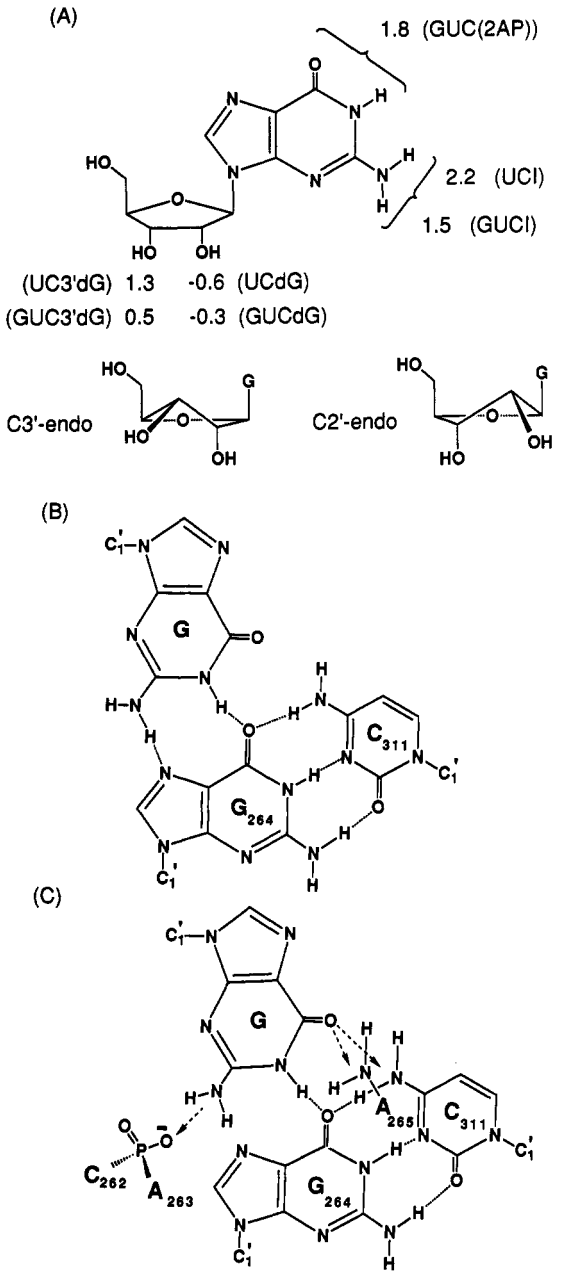

FIGURE 7: (A) Structure of guanosine with free energy increments for replacement of various functional groups with H (see Table I). A positive value indicates weaker binding upon replacement. (B) Diagram of the base triple G-binding site proposed by Michel et al. (1989a). All three bases lie in a plane. (C) Axial model of the G-binding site proposed by Yarus et al. (1991a,b). The substrate guanine base is rotated out of the plane of ribozyme base pair G_{264}-C_{311} such that the carbonyl group lies above the plane of the page and the exocyclic amine lies below. The exocyclic amine of the substrate guanine base interacts with the charged phosphodiester group between A_{263} and C_{262}, while the carbonyl fits into a pocket of positive charge formed by the amino hydrogens of C_{311} and A_{265}.

The above results suggest that hydrogen bonds to bases can contribute 1–2 kcal/mol to substrate binding to a ribozyme.

The preliminary kinetic studies indicate that interactions with the functional groups deleted in the binding studies are important in the transition state corresponding to the second step of splicing and to cyclization. GUCdG and GUCdGA are not reactive even though they bind. This suggests that an interaction with the 2′-OH of the 3′G is essential in the transition state (Bass & Cech, 1986). At 15 °C, GUCI and GUC(2AP) react more than 40 and 150 times slower than GUCG under conditions where the rate is determined by k_{cat}/K_M. These reductions in rate are equal to or greater than the reductions in binding at 5 °C, suggesting that contacts with the functional groups are maintained or even enhanced in the transition state.

The effects on steady-state kinetics of replacing monomer G with inosine and 2-aminopurine have been studied for splicing at 30 °C (Bass & Cech, 1984) and for an intermolecular G-addition reaction at 42 °C (Tanner & Cech, 1987). In both cases, the effects on "K_M" are larger than the effects on K_d reported here. For example, the "K_M"s for I and 2AP in splicing are 130- and 100-fold larger than the "K_M" for G (Bass & Cech, 1984). The differences are even larger for the G-addition reaction (Tanner & Cech, 1987). As discussed above, there is also a large difference between the "K_M" and K_d for G, which suggests a mechanism with an intermediate not attainable in equilibrium dialysis (for example, a covalent intermediate). Thus, interpretation of the different effects of substitution on "K_M" and K_d will require determination of the detailed mechanism.

Binding of the 3′ Substrate Has a Small Dependence on Temperature. The binding of pG, UCG, and GUCG was measured at 5 and 15 °C (Table I). All bind more weakly at higher temperature, indicating $\Delta H°$ for binding is negative.

There is a Trend to Weaker Binding as the Size of the 3′ Group on G Increases. GUCdG binds more tightly than GUCdGA or GUCdGU; d(TCG) binds more tightly than d(TCGT). Moreover, it appears that pG and G bind slightly more tightly than Gp. Evidently, binding gets weaker as the size of the 3′ group increases. This may have functional significance. First, the phosphodiester bond between the guanosine in the binding site and the nucleotide 3′ to this guanosine is the site of chemistry for exon ligation. Binding to this group in the ground state can be either unfavorable or neutral (Fersht, 1985; Menger, 1992). Unfavorable interactions in the ground state, however, can increase the catalytic rate if these interactions are relieved or replaced by favorable interactions in the transition state (Jencks, 1975; Fersht, 1985; Menger, 1992). The three-dimensional model of Michel and Westhof (1990) has a sharp turn in the sugar–phosphate backbone following the G at the catalytic site. The binding results are consistent with such a model since this strained backbone conformation could weaken binding. The results are also consistent with the susceptibility of UCGA and of the 3′ splice and cyclization site to hydrolysis [see Results and Inoue et al. (1986) and Zaug et al. (1984, 1985)]. The strained bond following G should react faster because the free energy of the ground state is brought closer to the free energy of the transition state.

A second functional role for unfavorable interactions with the 3′ groups following G may involve directing the order of reactions in splicing. The first step in splicing requires binding of an exogenous G (Cech et al., 1981). Presumably, this intermolecular binding must compete with the intramolecular binding of UCGU to the G-binding site. The binding experiments indicate that if both reactions are intermolecular, then UCG binds about 50 times more tightly than G. Unfortunately, the hydrolysis reaction precludes a direct comparison of G and UCGU by equilibrium dialysis. The results for GUCdGU, GUCdG, UCdGU, and UCdG, however, suggest that adding a 3′U to UCG will weaken binding by at least 30-fold. This would make the intermolecular binding of G and UCGU similar. Thus, unfavorable contacts with the nucleotide following G_{414} in the rRNA precursor may allow exogenous G to compete for the G-binding site.

The last step of the splicing cascade involves attack by the 3′-OH of G_{414} at the cyclization site $U_{15}A_{16}$ (Zaug et al., 1983). At this stage, it is advantageous for UCG to be held

in the active site long enough for reaction to occur but that it leave the active site after reaction. Differences in K_d's for RNA associations often reflect differences in off rates because on rates tend to be relatively independent of length (Pörschke & Eigen, 1971; Craig et al., 1971; Pörschke et al., 1973; Turner et al., 1990). If this is true for GUCdG and GUCdGA, and if similar effects are observed for the all-ribose sequences UCG and UCGA, then the off rate for UCGA will be about 50 times faster than for UCG. Thus the site of cyclization would be released much faster after cyclization than before, thereby suppressing reverse cyclization.

The 2'-OH of G Disfavors Binding. The 3'-OH of G Favors Binding. The 2'-OH of G Is Essential for the Transition State. There is a trend for substrates to bind slightly more tightly when the 2'-OH of G is replaced by H. Thus GUCdG, UCdG, dG, and pdG bind more tightly than GUCG, UCG, G, and pG, respectively (Table I). The preference is small, corresponding to an average $\Delta\Delta G°_5$ of 0.3 kcal/mol. In contrast, replacing the 3'-OH of G with H to give GUC(3'dG), UC(3'dG), and 3'dG clearly reduces binding. Consistent with expectations (Saenger, 1984; Olson, 1982; Westhof et al., 1977), NMR experiments indicate that G and dG in solution favor C2'-endo sugar puckers, while 3'dG favors C3'-endo (Table II). Thus the results suggest the G-binding site may favor a C2'-endo conformation. Dideoxy-G, however, favors a C3'-endo conformation in solution (Jagannadh et al., 1991) but binds as tightly as dG. This suggests that the weak binding of 3'dG may be partially due to an unfavorable contact with the 2'-OH group in the C3'-endo conformation. Since the sugar is the site of chemistry, it would not be surprising for the ribozyme to control sugar pucker. Alternatively or in addition, the 3'-OH group may be involved in a favorable contact in ground-state binding. If this involved a hydrogen bond to the H, then this would increase the nucleophilicity of the oxygen. It would be surprising if the contact involved the 3'-oxygen since this is the atom transferred to give the final 3'–5' linkage (Cech et al., 1981). Thus binding to this group in the ground state would likely slow the reaction for an intramolecular substrate (Fersht, 1985; Menger, 1992).

While the difference in binding between ribose and deoxyribose is subtle, the difference in reactivity is dramatic. No dG analogue has been found to react either under the conditions reported here or under a wide variety of additional variations in temperature and salt conditions (Moran and Turner, unpublished experiments; Bass & Cech, 1986). We estimate that replacing the 2'-OH of G with H decreases the rate for reactions 4 and 5 by more than a factor of 1000. This contrasts with results for 5' exon mimics in a reverse cyclization reaction of the *Tetrahymena* intron, which is similar to the reverse of reaction 4. For that case, rCdU and dCdT react only about 5-fold slower than rCrU (Sugimoto et al., 1989). Similarly, CUCdU reacts less than 10-fold slower than CUCU with UCGA (see reaction 4) (Bevilacqua & Turner, 1991). The results suggest that the 2'-OH of G is required for the first step of splicing, either as a group that is transferred or as a catalytic group. An absolute requirement for a 2'-OH group is also seen in the first step of group II and nuclear intron splicing (Michel et al., 1989b; Padgett et al., 1986).

Implications for Design of Reagents for RNA Engineering. Ribozymes like L-21 *Sca*I have been used as endonucleases for cleaving RNA molecules at defined sequences (Zaug et al., 1986). The results presented here indicate L-21 *Sca*I and related ribozymes should also serve to ligate two sequences using reactions such as 5 and 7. For example, reaction 5 with UCG and GUCG was essentially complete after a 2-h incubation. The observation that GUCG binds about 500 times more tightly than G suggests that base pairing of nucleotides preceding G will give adequate specificity. Reaction 7 is essentially the reverse of reaction 5, so that ligation can be used to add a sequence 3' to either UCG or CUCU. The specificity of both binding sites can be altered by site-directed mutagenesis (Been & Cech, 1986; Murphy & Cech, 1989; Michel et al., 1989a; Burke et al., 1990). Even the G-binding site can be altered to accept other nucleotides (Michel et al., 1989a; Been & Perrotta, 1991). Presumably, extending the substrate sequence in the 5' direction will not alter these properties, so that ribozymes can be engineered to cut and paste large RNA sequences.

Implications for Design of Therapeutics. Effective drugs often have dissociation constants in the micromolar range. It has previously been shown that CUCU binds to forms of the *T. thermophila* intron with dissociation constants in the nanomolar to micromolar range, depending on conditions (Sugimoto et al., 1988, 1989; Bevilacqua & Turner, 1991). Here GUCG binds to L-21 *Sca*I with a K_d in the micromolar range. In both cases, the binding is much tighter than expected for simple Watson–Crick pairing, thus implicating additional interactions. The results show that small oligonucleotides can bind tight enough to serve as therapeutics. Since this tight binding is likely dependent on the shape of a particular binding pocket, it is also likely to be highly specific. The observation that GUCdG is completely unreactive shows that small modifications can turn a substrate into an inhibitor (Bass & Cech, 1986). This suggests that small oligonucleotides can be designed to act as drugs. For example, group I introns with secondary structures similar to that of the *T. thermophila* LSU intron have been discovered in *P. carinii* (Lin et al., 1992; Liu et al., 1992), and it has been suggested that inhibitors targeted to the G-binding site might serve as therapeutics (Lin et al., 1992). The results presented here provide valuable insights for rational design of such inhibitors. A similar approach may be useful for attacking other pathogens in which transesterification reactions may be important. Sleeping sickness and elephantiasis are possible examples (Cech, 1991; Blum et al., 1991; Simpson & Shaw, 1989; Simpson, 1990; Benne, 1990). This approach has many advantages over current antisense strategies. In particular, much shorter oligomers need to be synthesized, and potential problems with specificity (Herschlag, 1991; Roberts & Crothers, 1991) are avoided by depending on tertiary structure in the target for recognition (Bevilacqua & Turner, 1991).

ACKNOWLEDGMENT

We thank Brian Cleary, Liyan He, Charlie Bailey, and Matt Fountain for assistance with the NMR and Dr. Phil Bevilacqua for helpful discussions.

REFERENCES

Abeles, R. H., Frey, P. A., & Jencks, W. P. (1992) *Biochemistry*, Jones & Bartlett, Boston, MA.
Agrawal, S., Christodoulou, C., & Gait, M. J. (1986) *Nucleic Acids Res. 14*, 6227–6245.
Banerjee, A. R., Jaeger, J. A., & Turner, D. H. (1993) *Biochemistry 32*, 153–163.
Bass, B. L., & Cech, T. R. (1984) *Nature 308*, 820–826.
Bass, B. L., & Cech, T. R. (1986) *Biochemistry 25*, 4473–4477.
Bass, B. L., & Weintraub, H. (1989) *Cell 55*, 1089–1098.
Been, M. D., & Cech, T. R. (1986) *Cell 50*, 207–216.
Been, M. D., & Perrotta, A. T. (1991) *Science 252*, 434–437.
Benne, R. (1990) *Trends Genet. 6*, 177–181.
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.
Bick, H. (1972) *Ciliated Protozoa*, p 74, World Health Organization, Geveva.

Blum, B., Sturm, N. R., Simpson, A. M., & Simpson, L. (1991) *Cell 65*, 543–550.
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.
Burke, J. M. (1989) *FEBS Lett. 250*, 129–133.
Burke, J. M., Belfort, M., Cech, T. R., Davies, R. W., Schweyen, R., Shub, D., Szostak, J. W., & Tabak, H. (1987) *Nucleic Acids Res. 15*, 7217–7221.
Burke, J. M., Esherick, J. S., Burfeind, W. R., & King, J. L. (1990) *Nature 344*, 80–82.
Cantor, C. R., & Schimmel, P. R. (1980) in *Biophysical Chemistry*, pp 852–859, W. H. Freeman & Co., San Francisco, CA.
Cech, T. R. (1990) *Annu. Rev. Biochem. 59*, 543–568.
Cech, T. R. (1991) *Cell 64*, 667–669.
Cech, T. R., Zaug, A. H., & Grabowski, P. J. (1981) *Cell 27*, 487–496.
Craig, M. E., Crothers, D. M., & Doty, P. (1971) *J. Mol. Biol. 62*, 383–401.
Davanloo, P., Rosenberg, A. H., Dunn, J. J., & Studier, F. W. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 2035–2039.
Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A., & Scazzacchio, C. (1985) *Nature 300*, 719–724.
Fersht, A. R. (1985) in *Enzyme Structure and Mechanism*, 2nd ed., Chapter 12, W. H. Freeman & Co., New York.
Fersht, A. R. (1987) *Trends Biochem. Sci. 12*, 301–304.
Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373–9377.
Grosshans, C. A., & Cech, T. R. (1989) *Biochemistry 28*, 6888–6894.
Hall, K. B., & McLaughlin, L. W. (1991) *Biochemistry 30*, 10606–10613.
Herschlag, D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 6921–6925.
Herschlag, D., & Cech, T. R. (1990a) *Biochemistry 29*, 10159–10171.
Herschlag, D., & Cech, T. R. (1990b) *Biochemistry 29*, 10172–10180.
Herschlag, D., Piccirilli, J. A., & Cech, T. R. (1991) *Biochemistry 30*, 4844–4854.
Hore, P. J. (1983) *J. Magn. Reson. 55*, 283–300.
Inoue, T., & Cech, T. R. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 648–652.
Inoue, T., Sullivan, F. X., & Cech, T. R. (1986) *J. Mol. Biol. 189*, 143–165.
Jaeger, J. A., Zuker, M., & Turner, D. H. (1990) *Biochemistry 29*, 10147–10158.
Jagannadh, B., Reddy, D. V., & Kunwar, A. C. (1991) *Biochem. Biophys. Res. Commun. 179*, 386–391.
Jencks, W. P. (1975) *Adv. Enzymol. 43*, 219–410.
Kay, P. S., & Inoue, T. (1987) *Nature 327*, 343–346.
Kay, P. S., Menzel, P., & Inoue, T. (1988) *EMBO J. 7*, 3531–3537.
Kierzek, R., Caruthers, M. H., Longfellow, C. E., Swinton, D., Turner, D. H., & Freier, S. M. (1986) *Biochemistry 25*, 7840–7846.
Lesser, D. R., Kurpiewski, M. R., & Jen-Jacobson, L. (1990) *Science 250*, 776–786.
Lin, J., Niu, M. T., Yoganathan, T., & Buck, G. A. (1992) *Gene 119*, 163–173.
Liu, Y., Rocourt, M., Pan, S., Liu, C., & Leibowitz, M. J. (1992) *Nucleic Acids Res. 20*, 3763–3772.
Markiewicz, W. T., Biala, E., & Kierzek, R. (1984) *Bull. Pol. Acad. Sci., Chem. 32*, 433–451.
Martin, F. H., & Tinoco, I., Jr. (1980) *Nucleic Acids Res. 8*, 2295–2299.
Menger, F. M. (1992) *Biochemistry 31*, 5368–5373.
Michel, F., & Westhof, E. (1990) *J. Mol. Biol. 216*, 585–610.
Michel, F., Hanna, M., Green, R., Bartel, D. P., & Szostak, J. W. (1989a) *Nature 342*, 391–395.
Michel, F., Umesono, K., & Oseki, J. (1989b) *Gene 82*, 5–30.
Michel, F., Netter, P., Xu, M.-Q., & Shub, D. A. (1990) *Genes Dev. 4*, 777–788.
Murphy, F. L., & Cech, T. R. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9218–9222.
Olson, W. K. (1982) *J. Am. Chem. Soc. 104*, 278–286.
Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S., & Sharp, P. A. (1986) *Annu. Rev. Biochem. 55*, 1119–1150.
Pörschke, D., & Eigen, M. (1971) *J. Mol. Biol. 62*, 361–381.
Pörschke, D., Uhlenbeck, O. C., & Martin, F. H. (1973) *Biopolymers 12*, 1313–1335.
Price, J. V., & Cech, T. R. (1988) *Genes Dev. 2*, 1439–1447.
Pyle, A. M., & Cech, T. R. (1991) *Nature 350*, 628–631.
Pyle, A. M., McSwiggen, J. A., & Cech, T. R. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 8187–8191.
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.
Roberts, R. W., & Crothers, D. M. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9397–9401.
Saenger, W. (1984) *Principles of Nucleic Acid Structure*, p 61, Springer-Verlag, New York.
SantaLucia, J., Jr., Kierzek, R., & Turner, D. H. (1991) *J. Am. Chem. Soc. 113*, 4313–4322.
SantaLucia, J., Jr., Kierzek, R., & Turner, D. H. (1992) *Science 256*, 217–219.
Scatchard, G. (1949) *Ann N.Y. Acad. Sci. 51*, 660–672.
Shirley, B. A., Stanssens, P., Hahn, U., & Pace, C. N. (1992) *Biochemistry 31*, 725–732.
Simpson, L. (1990) *Science 250*, 512–513.
Simpson, L., & Shaw, J. (1989) *Cell 57*, 355–366.
Sugimoto, N., Kierzek, R., & Turner, D. H. (1988) *Biochemistry 27*, 6384–6392.
Sugimoto, N., Tomka, M., Kierzek, R., Bevilacqua, P. C., & Turner, D. H. (1989) *Nucleic Acids Res. 17*, 355–371.
Sullivan, F. X., & Cech, T. R. (1985) *Cell 42*, 639–648.
Tanner, N. K., & Cech, T. R. (1987) *Biochemistry 26*, 3330–3340.
Turner, D. H., Sugimoto, N., Kierzek, R., & Dreiker, S. D. (1987) *J. Am. Chem. Soc. 109*, 3783–3785.
Turner, D. H., Sugimoto, N., & Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167–192.
Turner, D. H., Sugimoto, N., & Freier, S. M. (1990) in *Nucleic Acids* (Saenger, W., Ed.) Vol. 1C, pp 201–227, Springer-Verlag, Berlin.
Uhlenbeck, O. C. (1972) *J. Mol. Biol. 65*, 25–41.
Usman, J., Ogilvie, K. K., Jiang, M. Y., & Cedergren, R. J. (1987) *J. Am. Chem. Soc. 109*, 7845–7854.
van der Horst, G., & Inoue, T. (1993) *J. Mol. Biol. 229*, 685–694.
Wells, B. D., & Cantor, C. R. (1977) *Nucleic Acids Res. 4*, 1667–1680.
Westhof, E., Plach, H., Cuno, I., & Lüdemann, H.-D. (1977) *Nucleic Acids Res. 4*, 939–953.
Williams, D. H. (1992) *Aldrichimica Acta 25*, 9.
Williams, D. H., Cox, J. P. L., Doig, A. J., Gardner, M., Gerhard, U., Kaye, P. T., Lal, A. R., Nicholls, I. A., Salter, C. J., & Mitchell, R. C. (1991) *J. Am. Chem. Soc. 113*, 7020–7030.
Yarus, M., Illangesekare, M., & Christian, E. (1991a) *Nucleic Acids Res. 19*, 1297–1304.
Yarus, M., Illangesekare, M., & Christian, E. (1991b) *J. Mol. Biol. 222*, 995–1012.
Young, B., Herschlag, D., & Cech, T. R. (1991) *Cell 76*, 1007–1019.
Zaug, A. J., Grabowski, P. J., & Cech, T. R. (1983) *Nature 301*, 578–583.
Zaug, A. J., Kent, J. R., & Cech, T. R. (1984) *Science 224*, 574–578.
Zaug, A. J., Kent, J. R., & Cech, T. R. (1985) *Biochemistry 24*, 6211–6218.
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.