NMR Structures of (rGC**U**GA**G**GCU)$_2$ and (rGC**G**GA**U**GCU)$_2$: Probing the Structural Features That Shape the Thermodynamic Stability of GA Pairs[†,‡]

Blanton S. Tolbert,[§] Scott D. Kennedy,[§] Susan J. Schroeder,[∥] Thomas R. Krugh,[⊥] and Douglas H. Turner*[,⊥,#]

Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, New York 14642, Department of Chemistry and Biochemistry, University of Oklahoma, Norman, Oklahoma 73019-3051, Department of Chemistry, University of Rochester, Rochester, New York 14627-0216, and Center for Pediatric Biomedical Research and Department of Pediatrics, University of Rochester School of Medicine and Dentistry, Rochester, New York 14642

Received July 4, 2006; Revised Manuscript Received October 16, 2006

ABSTRACT: The NMR structures of $^{5'GCUGAGGCU3'}_{3'UCGGAGUCG5'}$ and $^{5'GCGGAUGCU3'}_{3'UCGUAGGCG5'}$ are reported. The internal loop, $^{5'UGAG3'}_{3'GAGU5'}$, is about 2 kcal/mol more stable than $^{5'GGAU3'}_{3'UAGG5'}$ at 37 °C. The duplexes assemble into similar global folds characterized by the formation of tandem sheared GA pairs. The different stabilities of the loops are accompained by differences in the local structure of the closing GU pairs. In the $^{5'UGAG3'}_{3'GAGU5'}$ internal loop, the GU pairs form canonical wobble configurations with two hydrogen bonds, whereas in $^{5'GGAU3'}_{3'UAGG5'}$, the GU pairs form a single hydrogen bond involving the amino group, GH22, and the carbonyl group, UO4. This pairing is similar to the GU closing pair of the 690 hairpin loop found in *E. coli* 16S rRNA. The $^{5'UGAG3'}_{3'GAGU5'}$ and $^{5'GGAU3'}_{3'UAGG5'}$ structures reveal how the subtle interplay between stacking and hydrogen bonding determines sequence dependent conformation and thermodynamic stability. Thus, this work provides structural and thermodynamic benchmarks for theoreticians in the ongoing effort to understand the sequence dependence of RNA physicochemical properties.

In the post-genomic era, it is evident that RNA actively choreographs many biological events ranging from catalysis to gene regulation (*1*−*4*). This emerging awareness of RNA function compels development of a fundamental understanding of the sequence dependence of RNA stability, structure, and dynamics. The impetus derives from the difficulty of empirically determining important physicochemical properties for every interesting RNA molecule revealed by genome sequencing projects. By systematically studying model systems, rules can be developed for predicting these properties from a primary sequence (*5*−*7*). This will greatly assist the rapid extraction of important functional RNAs from genome-wide searches (*8*, *9*).

RNA typically exists as a single stranded molecule capable of folding into secondary and tertiary structural elements. One important element of RNA secondary structure is the internal loop, which consists of stretches of non-Watson−Crick pairs flanked by canonical helical segments. The sequence dependence of internal loop stability and structure has been extensively studied experimentally (*10*−*38*), and some computational approaches have also been tested (*39*−*41*). Despite this effort, there are many unknowns regarding the nature of the physicochemical forces shaping internal loop stability, structure, and dynamics (*19*, *42*, *43*). In order to gain further insight into the nature of these forces, this article presents the solution 3D structures of tandem GA pairs closed by GU pairs.

GA pairs are highly represented in known RNA secondary structures, and their occurrence has been linked to both functional and structural roles (*44*, *45*). The sequence dependence of both the stability and local structure of tandem GA pairs has been investigated (*11*, *15*, *19*, *23*, *25*−*27*, *37*, *46*). The thermodynamic stabilities of symmetric tandem GA pairs vary in free energy over 4 kcal/mol depending on the identity of the adjacent/closing base pair (Table 1) (note negative $\Delta G°$ values are stabilizing) (*15*). The stability increment follows the trend: $\frac{5'G}{3'C} > \frac{5'C}{3'G} > \frac{5'U}{3'G} \geq \frac{5'A}{3'U} \geq \frac{5'U}{3'A} > \frac{5'G}{3'U}$.

The most stable tandem GA pair internal loop is $^{5'GGAC3'}_{3'CAGG5'}$, which has a favorable loop free energy increment of −2.6 kcal/mol (*15*, *26*). The NMR structure of this loop (*23*) has head-to-head GA imino pairs (cis Watson−Crick/Watson−Crick) (*47*), with Watson−Crick GC closing pairs. The most destabilizing tandem GA internal loop is $^{5'GGAU3'}_{3'UAGG5'}$, which has an unfavorable loop free energy increment of about 2 kcal/mol at 37 °C (*15*, *26*). Thus, the two C to U substitutions in going from $^{5'GGAC3'}_{3'CAGG5'}$ to $^{5'GGAU3'}_{3'UAGG5'}$ produce an approximately 1000-fold change in the equilibrium constant for folding at

[†] This work was supported by NIH Grant GM22939 (to D.H.T.). We also thank NIGMS for providing minority supplement grant GM22939 to B.S.T.

[‡] The coordinates have been deposited as Protein Data Bank entries 2IRN and 2IRO.

* Corresponding author. Phone: 585-275-3207. Fax: 585-276-0205. E-mail: turner@chem.rochester.edu.

[§] Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry.

[∥] University of Oklahoma.

[⊥] Department of Chemistry, University of Rochester.

[#] Center for Pediatric Biomedical Research and Department of Pediatrics, University of Rochester School of Medicine and Dentistry.

10.1021/bi061350m CCC: $37.00 © 2007 American Chemical Society
Published on Web 01/20/2007

Table 1: Thermodynamics of Loop Formation and Type of Sugar Pucker for Internal Loops

loop	$\Delta G^\circ_{37\,\mathrm{loop}}$[a] (kcal/mol)	sugar pucker of loop internal G[b]	GA type
5′G<u>GA</u>C3′ / 3′C<u>AG</u>G5′	−2.6	C3′-endo (N)	imino
5′C<u>GA</u>G3′ / 3′G<u>AG</u>C5′	−0.7	C3′-endo (N)	sheared
5′U<u>GA</u>G3′ / 3′G<u>AG</u>U5′	0.1	C3′/C2′-endo (N/S)	sheared
5′A<u>GA</u>U3′ / 3′U<u>AG</u>A5′	0.3	not determined	sheared
5′U<u>GA</u>A3′ / 3′A<u>AG</u>U5′	0.7	C2′-endo (S)	sheared
5′G<u>GA</u>U3′ / 3′U<u>AG</u>G5′	1.8	C2′-endo (S)	sheared

[a] The reported loop free energy increments for the GU/UG loops are updated (*15*) with the most recent nearest neighbor parameters (*5*, *7*). References for the original measurements are given in ref *15*. The thermodynamic stability increments were calculated from the following expression (*80*): $\Delta G^\circ_{37,\mathrm{loop}} = \Delta G^\circ_{37,(\mathrm{duplex\ w/loop})} - \Delta G^\circ_{37,(\mathrm{duplex\ w/out\ loop})} + \Delta G^\circ_{37,(\mathrm{interrupted\ basepair})}$. As an illustration, the loop free energy increment of $(\underline{GGAU})_2$ is calculated as follows: ΔG°_{37} $(\underline{GGAU})_2 = \Delta G^\circ_{37}$ $(\mathrm{rGCG}\underline{GAUGC})_2 - \Delta G^\circ_{37}$ $(\mathrm{rGCGUGC})_2 + \Delta G^\circ_{37}$ $(\underline{GU})_2$, where ΔG°_{37} $(\mathrm{rGCG}\underline{GAUGC})_2$ is the measured value of the duplex containing the internal loop (*26*), ΔG°_{37} $(\mathrm{rGCGUGC})_2$ is the measured value of the duplex without the loop (*26*), and ΔG°_{37} $(\mathrm{GU})_2$ is the free energy increment of the nearest neighbor interaction interrupted by the loop (*5*). [b] Sugar puckers are taken from refs 27 (CGAG), 23 (GGAC), 37 (UGAA), and from this work. They are based on coupling constants.

37 °C. This contrasts with a destabilization of about 1 kcal/mol at 37 °C in going from 5′C<u>GA</u>G3′ / 3′G<u>AG</u>C5′ to 5′U<u>GA</u>G3′ / 3′G<u>AG</u>U5′, which corresponds to less than a 10-fold change in the equilibrium constant for folding. The 5′G<u>GA</u>U3′ / 3′U<u>AG</u>G5′ internal loop is less stable than the 5′UGAG3′ / 3′GAGU5′ internal loop by 1.7 kcal/mol at 37 °C (*15*, *26*). According to a hydrogen-bonding model, 5′G<u>GA</u>U3′ / 3′U<u>AG</u>G5′ and 5′U<u>GA</u>G3′ / 3′G<u>AG</u>U5′ should be isoenergetic, yet they differ by at least a factor of 10 in K_{eq} at 37 °C.

What are the underlying forces responsible for the disparate stabilities? Part of the $\Delta\Delta G^\circ_{\mathrm{loop}}$ in going from 5′G<u>GA</u>C3′ / 3′C<u>AG</u>G5′ to 5′G<u>GA</u>U3′ / 3′U<u>AG</u>G5′ can be parsed into a change in the number of hydrogen bonds in the closing base pairs. The average difference between closing an internal loop with two GC or two AU pairs is 1.3 kcal/mol (*19*) and is expected to be similar for the GC to GU closure if only the number of hydrogen bonds is important. The difference in stabilities of 5′C<u>GA</u>G3′ / 3′G<u>AG</u>C5′ and 5′U<u>GA</u>G3′ / 3′G<u>AG</u>U5′ is consistent with this expectation. This leaves unexplained, however, ∼3 kcal/mol of sequence dependence of the free energy between 5′G<u>GA</u>C3′ / 3′C<u>AG</u>G5′ and 5′G<u>GA</u>U3′ / 3′U<u>AG</u>G5′.

To provide insight into correlations between the energetics and structure for tandem GA mismatches with GU closing pairs, the local structures of $(\mathrm{rGC}\mathbf{U}\underline{GA}\mathbf{G}\mathrm{GCU})_2$ and $(\mathrm{rGC}\mathbf{G}\underline{GA}\mathbf{U}\mathrm{GCU})_2$ have been determined by solution NMR and restrained molecular dynamic simulations (the residues in the antiparallel strand will be denoted with an * throughout the text). The two loops assemble into similar three-dimensional structures with sheared GA pairs (trans Hoogsteen/Sugar edge (*47*)) but with subtle differences in hydrogen bonding and stacking. Comparisons to the thermodynamics and structures of symmetrical tandem GA internal loops with Watson−Crick closing pairs provide new benchmarks for theoreticians investigating the intricate details of the interactions molding tandem GA internal loop stability (*39*).

MATERIALS AND METHODS

RNA Synthesis and Purification. rGCU<u>GA</u>GGCU and rGCG<u>GA</u>UGCU were prepared by solid-phase synthesis on an ABI 392 DNA/RNA synthesizer following the phosphoramidite method (*48*, *49*). CPG support and phosphoramidites were purchased from Proligo and Azco, respectively. Following synthesis, the resulting support bound oligonucleotides were incubated at 55 °C overnight in a 3:1 ammonium hydroxide/ethanol solution to remove the base-protecting groups and to cleave the oligonucleotides from the support. The soluble oligonucleotides were separated from the support with a Quik-Sep filtration column (BIO-RAD). Subsequently, the silyl-protecting groups were removed by incubating at 55 °C in a 9:1 TEA-3HF (triethylamine trihydrofluoride)/DMF solution for 2−4 h at 55 °C. The deprotected oligonucleotides were desalted on a Sep-Pak C18 column (Waters) and purified on a large preparative Whatman TLC plate (20 cm × 20 cm) using a 55:35:10 1-propanol/ammonia/water mobile phase. The main product bands were visualized by UV shadowing, scraped from the plate, and extracted with RNase free water. The purity of the oligonucleotides was checked by analytical TLC, and the masses were verified by ESI MS with a Hewlett-Packard 1100 LC/MS chemstation.

UV Melting Experiments. Purified oligonucleotides were lyophilized and reconstituted in 1 M NaCl, 20 mM cacodylate (pH 7), and 0.5 mM Na_2EDTA. All experiments were performed on a Beckman Coulter DU-640 spectrophotometer equipped with a Peltier temperature controller cooled with flowing water. Curves of absorbance at 260 nm versus temperature were recorded using a heating rate of 1 °C/min with a 1 min averaging time. The data were fit to a two-state model, assuming linear sloping baselines and temperature independent ΔH° and ΔS° using the MeltWin program (*50*). Additionally, the melting data were analyzed at different concentrations with the following relationship to extract the van 't Hoff thermodynamic parameters (*51*) (T_M, melting temperature in kelvins and C_T, total strand concentration).

$$T_M^{-1} = (2.303R/\Delta H^\circ)\log(C_T) + (\Delta S^\circ/\Delta H^\circ) \quad (1)$$

NMR Sample Preparation. Samples were prepared for NMR by reconstituting the purified oligonucleotides in 500 μL of RNase free water. The samples were then dialyzed exhaustively against 1 L of autoclaved double distilled water. Following dialysis, the samples were lyophilized and reconstituted in 300 μL of NMR buffer containing 80 mM NaCl, 0.5 mM Na_2EDTA, and 10 mM sodium phosphate at pH 5.5 (90% H_2O/10% D_2O). The pH was measured in the NMR tube using an Orion micro-combination pH/sodium electrode (Thermo Electron Corporation) and was 5.30 and 5.18 for $(\mathrm{rGC}\mathbf{U}\underline{GA}\mathbf{G}\mathrm{GCU})_2$ and $(\mathrm{rGC}\mathbf{G}\underline{GA}\mathbf{U}\mathrm{GCU})_2$, respectively. For nonexchangeable proton spectra, the oligonucleotides were lyophilized and reconstituted in 99.96% D_2O three times. The final NMR samples were prepared in 99.996% D_2O (Cambridge Isotopes) at pH 5.85. The single strand concentrations for $(\mathrm{rGC}\mathbf{U}\underline{GA}\mathbf{G}\mathrm{GCU})_2$ and $(\mathrm{rGC}\mathbf{G}\underline{GA}\mathbf{U}\mathrm{GCU})_2$ were ∼1.5 and 2.0 mM, respectively.

NMR Spectroscopy. All of the NMR experiments were performed on a Varian Inova 600 MHz spectrometer. The

Table 2: Thermodynamic Parameters of Duplex Formation[a]

	$1/T_m$ vs $\log(C_T)$				average of curve fits			
sequence	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$-\Delta G°_{37}$ (kcal/mol)	T^b_M (°C)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$-\Delta G°_{37}$ (kcal/mol)	T^b_M (°C)
5′GCUGAGGCU3′ 3′UCGGAGUCG 5′	63.5 ± 3 (65.1 ± 2)	180.0 ± 9.6 (188.4 ± 5)	7.63 ± 0.07 (6.70 ± 0.03)	46.9 (42.0)	69.9 ± 13 (58.4 ± 6)	200.1 ± 40.8 (166.71 ± 20)	7.88 ± 0.36 (6.71 ± 0.06)	47.1 (42.6)
5′GCGGAUGCU3′ 3′UCGUAGGCG 5′	58.3 ± 5 (51.5 ± 1)	168.3 ± 16.2 (151.1 ± 4)	6.12 ± 0.08 (4.64 ± 0.03)	39.4 (30.9)	63.4 ± 9 (50.4 ± 9)	184.6 ± 29.6 (147.18 ± 31	6.19 ± 0.27 (4.79 ± 0.27)	39.5 (31.6)

[a] All measurements were made in 1 M NaCl, 10 mM sodium cacodylate (pH 7.0), and 0.5 mM Na_2EDTA. The values in parentheses were measured without the 3′ dangling U and are taken from ref 26. [b] Calculated at 10^{-4} M oligonucleotide strand concentration.

1D imino proton spectra were recorded in 90% H_2O using an S-shaped excitation pulse with excitation maximum near 12 ppm. SNOESY spectra were recorded at 100, 120, and 200 ms mixing times at various temperatures. NOESY spectra were acquired in D_2O at 100, 200, and 400 ms mixing times at 10 °C for (rGC**U**G**A**G**G**CU)$_2$ and 5 °C for (rGC**G**G**A**U**G**CU)$_2$. Scalar coupled protons were identified in TOCSY spectra recorded at 12 and 40 ms mixing time and in DQFCOSY spectra. Data processing was performed with NMRPipe (*52*).

Restraint Generation and Structure Calculation. Peak integrations were performed in Sparky using the sum-over-box method (*53*). Data from the 100 ms mixing time NOESY spectra were used to derive distance restraints using the two-spin approximation with $1/r^6$ scaling. Data from the 400 ms mixing time NOESY spectra were only used to confirm peaks that were weak at 100 ms. The reference volumes were taken from the H1′−H2′ cross-peaks of the stem bases with the reference distance for this pair of protons set to 2.75 Å. To account for relaxation effects, spin diffusion, or baseline distortions, distance bounds were set to ±30% for the 100 ms mixing time data. Distances derived from overlapped cross-peaks were loosened because of the inaccuracies in their volume measurements. There were several critical loop-defining distance restraints for both (rGC**U**G**A**G**G**CU)$_2$ and (rGC**G**G**A**U**G**CU)$_2$ determined from the SNOESY data, which are discussed below. All of these restraints were placed into bins of 1.8−3.5 and 2.5−5.5 Å based on the average volume of pyrimidine H5−H6 cross-peaks (2.45 Å) at 5 °C for both (rGC**U**G**A**G**G**CU)$_2$ and (rGC**G**G**A**U**G**CU)$_2$. Hydrogen bond restraints were applied to the four GC pairs in the stems of each duplex. The values were set to the average hydrogen bond distance ± the standard deviation measured for a 1.20 Å crystal structure of an all A-form RNA duplex determined by Steitz and co-workers (*54*) (see captions to Tables S4 and S5 (Supporting Information) for details). The GU pairs in (rGC**U**G**A**G**G**CU)$_2$ were restrained to be in a canonical wobble configuration, but the GU pairs of (rGC**G**G**A**U**G**CU)$_2$ were left unrestrained or restrained to various conformations because of the lack of G and U imino NOE cross-peaks (see below).

Backbone torsion angle restraints were semiquantitatively determined from ^{1}H-^{31}P HETCOR spectra. The Watson−Crick region of each oligonucleotide displayed a typical A-form spread in the phosphorus chemical shifts (∼1 ppm) (*55*). As a result, this region was restrained to have A-form-like backbone torsion angles. The loop torsion angles for both duplexes were left unrestrained. The glycosidic torsion angle, χ, was restrained to the anti-conformation for all nucleotides in the duplex. This is consistent with the NMR data because there are no unusually large intraresidue H1′−H8 cross-peaks indicative of a syn conformation. Sugar pucker restraints, δ, were derived from the TOCSY and DQFCOSY spectra. All stem bases were restrained to the C3′ endo conformation based on the absence of H1′−H2′ cross-peaks and the existence of strong H2′−H3′ cross-peaks for the duplexes. In (rGC**U**G**A**G**G**CU)$_2$, the δ torsion angles for the G4, A5, and U9 were restrained to cover both the *C2′* and C3′ endo conformations. In(rGC**G**G**A**U**G**CU)$_2$, the δ torsion angle for G4 was restrained to the C2′ endo conformation because of the presence of a strong H1′−H2′ (>10 Hz) cross-peak in the DQFCOSY spectrum. A5 and U9 were restrained to cover both sugar pucker conformations as in (rGC**U**G**A**G**G**CU)$_2$.

Restrained molecular dynamics and simulated annealing calculations were performed in CNS version 1.1 (*56*). The initial input structures were built as A-form duplexes in the Biopolymer module of Insight II. The rna-dna-all-atom parameter and topology files of CNS were used to assign atomic charges, masses, atom linkages, and force constants (*56*). The dielectric constant was set to unity. The structure calculations proceeded as follows. (1) High temperature dynamics: the system was raised to 1000 K and allowed to develop over 1 ps under the influence of bond, angle, improper, NOE (5 kcal/mol Å^2), and dihedral angle (100 kcal/mol rad^2) restraints. (2) Simulated annealing in Cartesian space: the temperature was cooled from 1000 to 0 K in 100 K steps while holding the NOE and dihedral angle force constants at 150 kcal/mol Å^2 and 200 kcal/mol rad^2, respectively, and increasing the van der Waals scale factor linearly from 1 to 4. The electrostatics were off except between hydrogens and other atoms, which were set to 10%. At each temperature, the system was allowed to develop over 20 ps using 5000 MD steps. (3) Energy minimization: Powell energy minimization was applied with the van der Waals scale factor at 1 and electrostatic terms at 100%. Planarity restraints (25 kcal/mol Å^2) were applied to the GC stem regions during the course of the simulation. Duplexes (rGC**U**G**A**G**G**CU)$_2$ and (rGC**G**G**A**U**G**CU)$_2$ required 2000 and 1000 steps, respectively, to fully optimize van der Waals contacts. All structures agreed with the experimental NOE and dihedral angle restraints within 0.2 Å and 5°, respectively. All models were visualized and analyzed in Pymol (*57*) and helical parameters calculated with X3DNA (*58*).

RESULTS

Thermodynamics. Table 2 shows the thermodynamic parameters measured for the duplex formation of (rGC**U**G**A**G**G**CU)$_2$ and (rGC**G**G**A**U**G**CU)$_2$, along with previous values for the duplexes without the 3′ dangling uracils (*26*). Addition

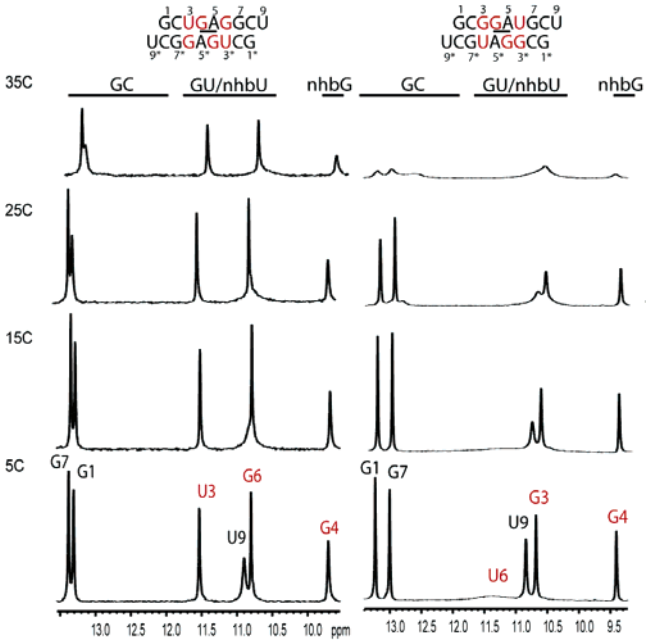

FIGURE 1: NMR spectra for the imino proton region as a function of temperature for (left) (rGC**U**GA**G**GCU)$_2$ and (right) (rGC**G**GA**U**GCU)$_2$. The spectra were recorded in 80 mM NaCl, 0.5 mM Na_2EDTA, and 10 mM sodium phosphate at pH 5.30 and 5.18, respectively. The bars above demarcate the regions expected for base pair types: GC, Watson−Crick GC; GU, wobble GU; nhbU, non-hydrogen bonded U; and nhbG, non-hydrogen bonded G. The loop G and U bases are colored red.

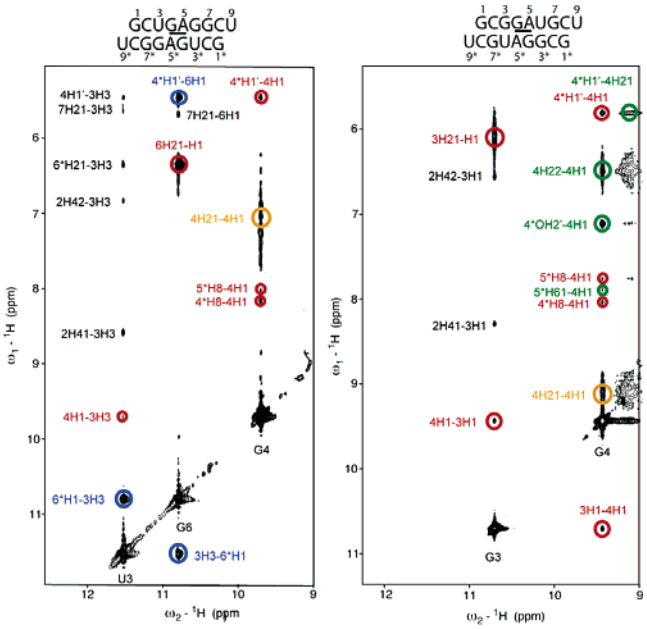

FIGURE 2: Comparison of the imino region of the 120 ms mixing time SNOESY spectra at 5 °C for the internal loops of (left) (rGC**U**GA**G**GCU)$_2$ and (right) (rGC**G**GA**U**GCU)$_2$. The labels and color scheme are as follows: red, similar NOE cross-peaks in both (rGC**U**GA**G**GCU)$_2$ and (rGC**G**GA**U**GCU)$_2$; blue, NOE cross-peaks unique to (rGC**U**GA**G**GCU)$_2$; green, NOE cross-peaks unique to (rGC**G**GA**U**GCU)$_2$; and orange, similar NOE cross-peaks in both (rGC**U**GA**G**GCU)$_2$ and (rGC**G**GA**U**GCU)$_2$ with ~2 ppm chemical shift difference in the ω_1 dimension. Notice that the GU pair in (rGC**U**GA**G**GCU)$_2$ is characterized by a strong NOE between U3/3*H3 and G6*/6H1, whereas in (rGC**G**GA**U**GCU)$_2$, the U6/6*H3 peak is missing (see text for more details).

of the 3′ dangling uracils increases the duplex stability by 0.9 and 1.5 kcal/mol at 37 °C for (rGC**U**GA**G**GCU)$_2$ and (rGC**G**GA**U**GCU)$_2$, respectively. This is less than the 2.2 kcal/mol increase in stability anticipated for two $\frac{5'CU}{3'G}$ interactions per duplex (*59*). Evidently, the internal loops destabilize the 3′ favorable stacking interaction imparted by the dangling U relative to a fully Watson−Crick helix. This is a non-nearest neighbor effect. A similar destabilization for 3′ dangling ends has been observed for bulge loops where the stability increments become less favorable as a function of loop size (*60*).

Exchangeable Proton Assignments. The 1D imino proton spectra for both duplexes are shown in Figure 1. In the 12.5−13.5 ppm region, each duplex gives two imino proton peaks, which correspond to GC base pairs. In (rGC**U**GA**G**GCU)$_2$ at 5 °C, four imino proton peaks are observed upfield of the GC pairs. These peaks belong to the remaining G and U bases. Even at pH 5.2 and 5 °C, only three of the expected four peaks are distinctly observed in an analogous region of the spectrum for (rGC**G**GA**U**GCU)$_2$. A previous 1D NMR analysis at 20 °C of (rGC**G**GA**U**GC)$_2$ without a 3′ dangling U exhibited a broad peak around 13 ppm (*26*). On the basis of the downfield shift of this peak, it was suggested that (rGC**G**GA**U**GC)$_2$ forms a tandem head-to-head Watson−Crick GA pair similar to that found in (rGC**G**GA**C**GC)$_2$ (*23*). The work shown here contradicts that assessment because a similar shifted peak is not observed until the temperature is elevated above 25 °C, which is within 15 °C of the T_m (Figure 1). This peak is probably a result of the conformational exchange of one of the stem guanosines. Also, the spectral assignments of (rGC**G**GA**U**GCU)$_2$ reported here establish that G4H1 resonates at ~9.5 ppm (see below), which is typical for non-hydrogen-bonded imino protons of guanosines.

The imino and stem amino protons were assigned from the SNOESY spectra following established procedures (Figures S1 and S2, Supporting Information) (*61*). Some base-specific assignments of the G and U imino protons were confirmed by natural abundance ^{1}H-^{15}N HSQC spectra (Figure S3, Supporting Information). G1 and G7 imino proton peaks were confirmed by cross-strand NOEs to amino and H5 (by spin diffusion through aminos) protons of C8* and C2*, respectively, indicative of the formation of homoduplexes (Figures S1 and S2, Supporting Information).

In (rGC**U**GA**G**GCU)$_2$, the U3 and G6 imino resonances were assigned by the presence of a strong cross-peak between the two protons (Figure 2), which is expected for a GU wobble pair because the separation distance is less than 3 Å. G6H1 was confirmed to resonate upfield of U3H3 on the basis of the imino resonance connectivity pathway, a strong NOE to its H21 (Figure 2), and from the natural abundance ^{1}H-^{15}N HSQC spectrum (Figure S3, Supporting Information). The U9 imino resonance was confirmed by a very weak cross-peak to G1H1 and to C8H41. The chemical shift of U9 at ~10.9 ppm is consistent with the chemical shift of the 3′ dangling uracil in 5′GGUGGAGGCU/ 3′PCCGAAGCCG (*10*). By the process of elimination, the imino resonance at 9.7 ppm was assigned to G4.

A similar strategy provided the assignments for (rGC**G**GA**U**GCU)$_2$. An important difference between (rGC**G**GA**U**GCU)$_2$ and (rGC**U**GA**G**GCU)$_2$ is the lack of a G3/3*H1 to U6*/6H3 cross-peak because the U6H3 imino

resonance is broad (Figure 1). The lack of a G3/3*H1−U6*/6H3 NOE in (rGC**G**GA**U**GCU)$_2$ (Figure 2) suggests that the closing GU pairs are not in a wobble conformation.

Important NOE restraints for defining the loop topology in both duplexes were obtained from SNOESY spectra (Figure 2). In (rGC**U**GA**G**GCU)$_2$, there is a cross-strand NOE between the G4 imino proton and G4*H8. This NOE is useful in orienting the G4 base from one strand relative to the G4 base in the other strand. Also, G4H1 has a NOE to G4*H1′ and to G4*H2′. Both cross-peaks are cross-strand NOEs because the intranucleotide separation is >6 Å, which is too far to give rise to a cross-peak of the observed size. Thus, the G4 base from one strand is close to the G4 from the other strand.

In (rGC**G**GA**U**GCU)$_2$, G4H1 shows cross-strand NOE patterns similar to those of G4H1 in (rGC**U**GA**G**GCU)$_2$ (Figure 2), that is, the G4H1-G4*H8 and G4H1-G4*H1′ cross-peaks. Additional NOEs were observed in (rGC**G**GA**U**GCU)$_2$ relative to (rGC**U**GA**G**GCU)$_2$ (Figure 2). The differences between the two spectra include the following: (1) G4H1 of (rGC**G**GA**U**GCU)$_2$ has a cross-strand NOE to its own OH2′. (The OH2′ assignment was confirmed by a H_2O TOCSY spectrum (data not shown).) (2) Both A5H61 and A5H62 have a strong cross-strand NOE to G4*H1′. (3) Both G4H21 and G4H22 have a strong cross-strand NOE to G4*H1′. The aforementioned NOE patterns suggest that different interactions are involved in forming the local structure of (rGC**G**GA**U**GCU)$_2$ relative to (rGC**U**GA**G**GCU)$_2$.

Tandem GA pairs as well as some GU pairs have been implicated as metal binding pockets, with cobalt hexamine serving as a spectroscopic probe for binding (*30*, *62*, *63*). The addition of 3 mM cobalt hexamine to (rGC**U**GA**G**GCU)$_2$ did not provide any further NOE information regarding the exchangeable protons, and the nonexchangeable NOE patterns with and without cobalt hexamine were identical. Similarly, the addition of 1 mM cobalt hexamine did not sharpen the line width of U6H3 in (rGC**G**GA**U**GCU)$_2$ or provide any additional information about the other exchangeable protons. Only slight spectral differences were observed in the aromatic region of the spectrum, while leaving the NOESY walk pattern unchanged (data not shown). Evidently, cobalt hexamine does not perturb the structures.

Nonexchangeable Proton Assignments. NOESY spectra were recorded at 10 and 5 °C for (rGC**U**GA**G**GCU)$_2$ and (rGC**G**GA**U**GCU)$_2$, respectively. At higher temperatures, both duplexes show a number of conformational exchange peaks. The 400 ms mixing time NOESY spectra in the (H8/H2/H6)−(H5/H1′) region are shown in Figure 3. Complete sequential NOEs are observed for both duplexes, indicating structured loops. Except for H5′, H5″, and U9H4′ of (rGC**G**GA**U**GCU)$_2$, complete aromatic and sugar proton assignments (Tables S2 and S3, Supporting Information) were made for both duplexes following established procedures (*61*). Identification of the pyrimidine H5−H6 cross-peaks from TOCSY spectra assisted with initiating the assignments (Figure S4, Supporting Information). In (rGC**U**GA**G**GCU)$_2$, the G6H1′ is shifted upfield to around 4.3 ppm. This peak was confirmed by the observation of a cross-peak in the natural abundance ^{1}H-^{13}C HSQC spectrum (Figure S5, Supporting Information). An upfield shift of the G6H1′ of (rGC**U**GA**G**GCU)$_2$ is consistent with this proton being positioned above or below an aromatic ring as observed with GNRA tetraloops and sheared purine−purine pairs (*27*, *64*, *65*). The equivalent H1′ in (rGC**G**GA**U**GCU)$_2$, U6H1′, has a chemical shift of 5.77 ppm (Figures 3 and S6 (Supporting Information)), which is within the standard chemical shift region for H1′/H5 protons. The H5 proton of U6 in (rGC**G**GA**U**GCU)$_2$ is shifted upfield to 4.4 ppm, however (Figures S4 and S6, Supporting Information). The differences in the chemical shifts of H1′ at position 6 of (rGC**U**GA**G**GCU)$_2$ and (rGC**G**GA**U**GCU)$_2$ suggest a difference in the local structure at the base 5 to base 6 step in the duplexes. Also, the A5H2 protons in the two duplexes show medium strength NOEs to both A5H1′ protons and to G6H1′ of (rGC**U**GA**G**GCU)$_2$ or U6H1′ of (rGC**G**GA**G**UCU)$_2$ (Figure 3). The intensity of the A5H2-H1′ NOE for both duplexes suggests that it is an interstrand NOE, implying that the A5 bases stack on each other across the strand and on G6 of (rGC**U**GA**G**GCU)$_2$ or U6 of (rGC**G**GA**U**GCU)$_2$ as has been observed in (GGCAAGCCU)$_2$ and (GCCGAGGC)$_2$ (*14*, *27*).

The assignments of H2′ protons follow directly from the strong H1′−H2′ cross-peaks in NOESY spectra. The assignments were confirmed by strong intrastrand (n)H8/H6−H2′ (n-1) NOEs in the aromatic to sugar region of the spectra.

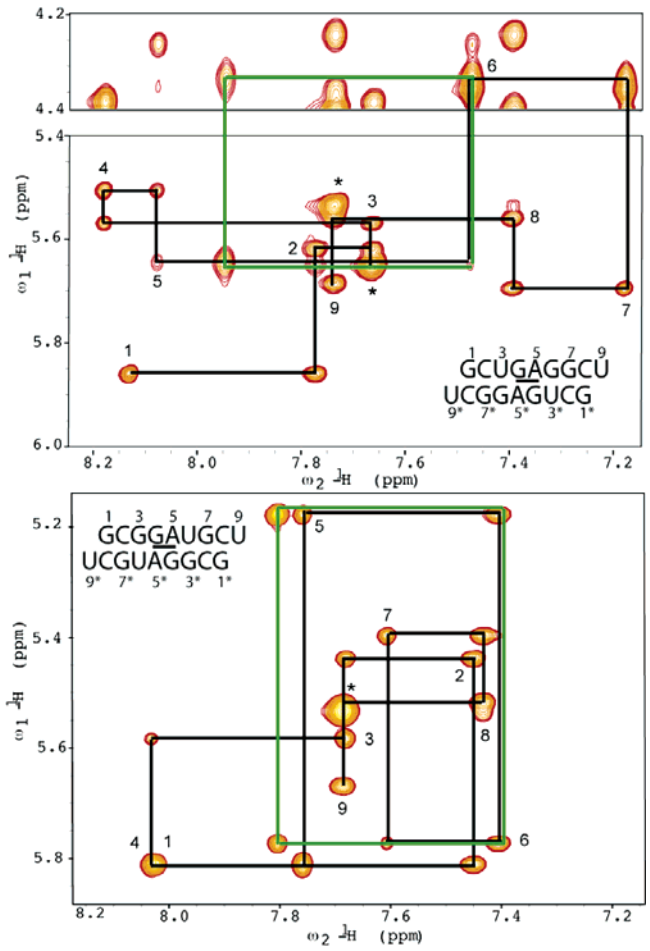

FIGURE 3: The 400 ms mixing time 2D NOESY spectra of the H8/H2/H6 to H1′/H5 region of (top) (rGC**U**GA**G**GCU)$_2$ at 10 °C (pH 5.85) and (bottom) (rGC**G**GA**U**GCU)$_2$ at 5 °C (pH 5.85). The black lines trace the connectivity from H8/H6/H2 to H1′ for both duplexes. The green rectangles indicate identical NOE patterns in both (rGC**U**GA**G**GCU)$_2$ and (rGC**G**GA**U**GCU)$_2$ involving A5H2 to A5*H1′, G6H1′ for (rGC**U**GA**G**GCU)$_2$ and A5H2 to A5*H1′, U6H1′ in (rGC**G**GA**U**GCU)$_2$. Asterisks denote H5−H6 cross-peaks.

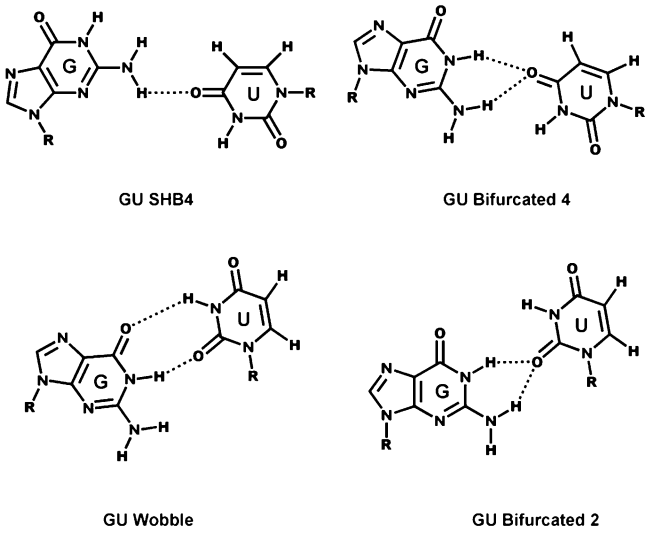

FIGURE 4: Modeled GU pair conformations for (rGC**G**GA**U**GCU)$_2$ (see text for more details).

H3′ assignments were made from cross-peaks to H2′s in DQFCOSY spectra and from ^{1}H-^{31}P HETCOR spectra (Figures S7 and S8, Supporting Information), which also facilitated assignments of phosphorus resonances. The H3′ assignments were further verified by the presence of medium strength intrastrand NOEs from H8/H6−H3′ protons in NOESY spectra. Some H4′ assignments were made from the scalar coupling to H3′ protons observed in DQFCOSY spectra. All aromatic and most ribose proton assignments were further verified by natural abundance ^{1}H-^{13}C HSQC spectra.

Structural Modeling. The structures for (rGC**U**GA**G**GCU)$_2$ and (rGC**G**GA**U**GCU)$_2$ were modeled by restrained molecular dynamics and simulated annealing as described in the Materials and Methods section. The closing GU pairs of (rGC**U**GA**G**GCU)$_2$ form canonical wobble configurations, which is consistent with the strong cross-peak between U3/U3* and G6*/G6 imino resonances (Figure 2). The conformation of the closing GU pairs in (rGC**G**GA**U**GCU)$_2$ is less obvious because of the lack of a NOE between the imino protons of the GU pairs (Figure 2). The NOE is missing because the U6H3 resonance is broad (Figure 1) and lost during the mixing time, presumably because of the rapid exchange with the solvent. The broad line width of U6H3 is unlikely a result of conformational dynamics because the linewidths of its aromatic and ribose protons are comparable to the linewidths of the other aromatic and ribose protons in the duplex. In order to identify possible GU pair structures consistent with the data, the modeling was done several times with different imposed restraints.

Modeling calculations for (rGC**G**GA**U**GCU)$_2$ were performed by restraining the GU pairs in four conformations (Figure 4): (a) single hydrogen bond 4 (SHB4); GH22 to UO4, (b) bifurcated 4; GH1/H21 to UO4, (c) GU wobble, and (d) bifurcated 2; GH1/H21 to UO2 (note that the numbering for the SHB and bifurcated structures refer to the number of the acceptor oxygen atom of U). Each set of structures was analyzed for their agreement with the NMR restraints and chemical shifts. Each conformation yielded structures consistent with the NOE derived distance bounds within 0.2 Å, but only the SHB4 and bifurcated 4 conformations satisfied the dihedral angle selection criteria. Structures derived from the GU wobble and bifurcated 2 conformations consistently displayed a G3 χ dihedral angle violation greater than 5°, which is likely a result of the tight contraints used for the χ dihedral angles ($-150 \pm 20°$). Supporting Information, Table S1 (a−d) shows the average simulated energies for all converged structures along with the NOE energy components. As judged by the simulated energies, the SHB4 (-453.1 ± 14.3 kcal/mol) and bifurcated 4 (-459.1 ± 9.1 kcal/mol) models satisfy the force field parameters better than the wobble (-359.6 ± 9.0 kcal/mol) and bifurcated 2 (-356.6 ± 13.0 kcal/mol) models. This is in part a reflection of the more unfavorable NOE energies associated with the wobble (10.96 ± 1.19) and bifurcated 2 (12.34 ± 2.05) models compared to the SHB4 (1.84 ± 0.36) and bifurcated 4 (1.75 ± 0.25) models (Table S1, Supporting Information).

Table 3: Comparison of Observed and Predicted Chemical Shift Values for U6H5 and H1′ of (rGC**G**GA**U**GCU)$_2$ Modeled with Different GU Hydrogen Bond Conformations[a]

	U6H5^{Obs}	U6H5^{Pred}	U6H1′^{Obs}	U6H1′^{Pred}
GU SHB4	4.41	4.57 ± 0.4	5.77	5.43 ± 0.30
GU bifurcated 4	4.41	5.45 ± 0.10	5.77	5.19 ± 0.15
GU wobble	4.41	5.59 ± 0.07	5.77	3.55 ± 0.46
GU bifurcated 2	4.41	5.59 ± 0.10	5.77	3.58 ± 0.43

[a] The GU SHB4 model shows the best agreement between the observed and predicted shifts (see text for more details). The predicted values were calculated with the program Nuchemics (*66*, *67*). As a comparison, the GU SHB4 structure was also calculated in the Discover 98 software package using the AMBER 95 forcefield. The predicted values for U6H5 and H1′ are 4.26 ± 0.16 and 5.87 ± 0.04, respectively.

The SHB4 pairing (Figure 4) provided the best agreement between the observed chemical shifts for the loop aromatic and H1′ protons and those predicted by the program Nuchemics (*66*, *67*) (Tables 3 and S1 (Supporting Information)). In particular, the unusual U6H5 chemical shift of 4.41 ppm is predicted to be 4.57 ± 0.4 ppm for SHB4 and 5.45 to 5.59 ± 0.1 ppm for the other GU pairings. SHB4 also shows the best agreement between the observed and predicted U6H1′ shifts (Table 3). The unusual U6H5 shift is a result of ring current effects from the neighboring A5 nucleoside. A similar ring current effect explains the upfield shift to 4.32 ppm of G6H1′ of (rGC**U**GA**G**GCU)$_2$, which is consistent with the predicted chemical shift of 4.13 ± 0.33 ppm. Figure S9 in the Supporting Information shows the position of U6H5 and H1′ for the SHB4 and wobble models along with the position of G6H1′ of (rGC**U**GA**G**GCU)$_2$ relative to the respective A5 nucleoside.

U6H3 in (rGC**G**GA**U**GCU)$_2$ has a broad line shape compared to the sharp lines for the nonexchangeable protons of U6 and also has a strong cross-peak to water. Both observations indicate that U6H3 is in rapid exchange with water and not in a GU wobble conformation. The sharp line width of the G3/3* imino proton (Figure 1), however, suggests that it is protected from exchange with water. In structures with the GU SHB4 conformation, G3/3*H1 is in the major groove underneath the G4/4* carbonyl oxygen. In this orientation, G3/3*H1 could be stabilized by a water bridged hydrogen bond to the G4/4* carbonyl oxygen. On the basis of functional group substitution (*68*) and NMR derived structures (*64*), water bridged hydrogen bonds have been proposed to help stabilize GNRA tetraloops.

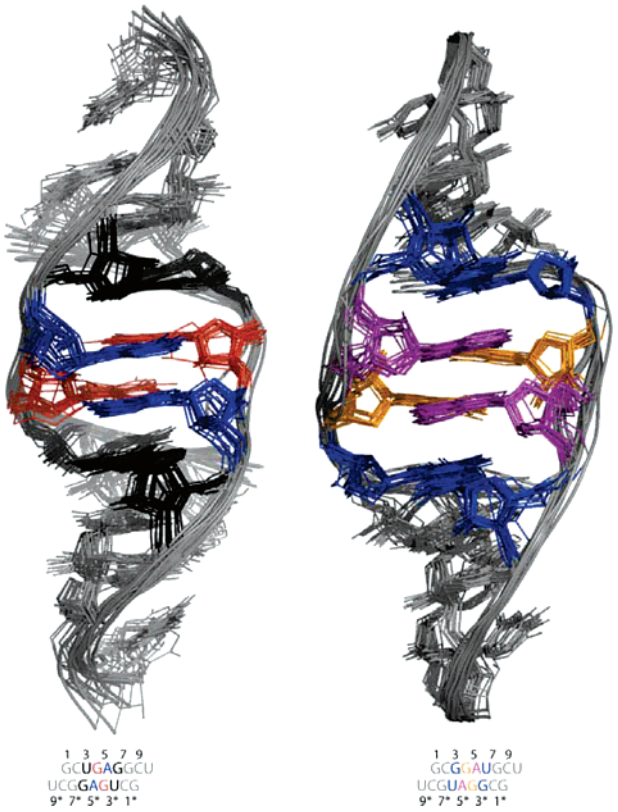

FIGURE 5: Minor groove view of a superposition of 15 selected structures of the 30 converged structures for (left) (rGC**U**G**A**G**G**CU)$_2$ (rmsd of 0.94 Å) and (right) (rGC**G**G**A**U**A**UGCU)$_2$ (rmsd of 0.78 Å). The color scheme is indicated in the sequence below each structure. Hydrogens and nonbridging oxygens have been removed for clarity.

The SHB4 conformation of (rGC**G**G**A**U**A**UGCU)$_2$ agrees best with the NMR data, and the ensuing discussions will focus on this model. The structure calculation statistics are provided in Tables S4 and S5 (Supporting Information). Figure 5 shows the superposition of 15 lowest energy structures for (rGC**U**G**A**G**G**CU)$_2$ and (rGC**G**G**A**U**A**UGCU)$_2$. Both structures are well defined because the all-atom pairwise rmsd's for the 30 calculated structures of (rGC**U**G**A**G**G**CU)$_2$ and (rGC**G**G**A**U**A**UGCU)$_2$ are 0.94 and 0.78 Å, respectively. Both internal loops form tandem sheared GA base pairs (trans Hoogsteen/Sugar edge AG) (*47*, *69*) as has been observed elsewhere (*11*, *27*).

DISCUSSION

A complete description of RNA structure−function relationships requires a mastery of the sequence dependence of the physicochemical properties of RNA. It is impractical to empirically determine these properties for every known RNA sequence. If the connections between sequence, stability, and structure can be deciphered, then rules for predicting structure and possibly function could be established. Although these linkages are largely known for fully paired Watson−Crick regions, internal loops still pose a challenge. This is largely due to the vast number of possible base pair and stacking configurations within internal loops. To provide benchmarks toward advancing the basic comprehension of RNA internal loops, the solution structures of two duplexes containing tandem GA pairs adjacent to GU pairs have been solved. Despite being thermodynamically destabilizing (Table 1), both loops adopt relatively ordered structures at low temperatures. The detailed structures of the loops consist of tandem sheared GA mismatches but with different hydrogen bonding patterns within the GU closing pairs. Presumably, the sequence dependent hydrogen bonding and stacking contribute to the disparate loop thermodynamic stabilities.

Thermodynamics of Internal Loops Containing Tandem GA Pairs. Table 1 shows the experimental free energy increments for all symmetric tandem GA internal loops of the type $^{5'XGAY3'}_{3'YAGX5'}$, where XY denotes GC/CG, AU/UA, or GU/UG base pairs. The loop free energies for tandem GA pairs span 4 kcal/mol at 37 °C. Of the six internal loops listed in Table 1, five have solution NMR structures (*23*, *27*, *37*) with one of those also having a crystal structure (*11*). The exception is the $^{5'AGAU3'}_{3'UAGA5'}$ internal loop, which is believed to form a sheared GA pair on the basis of a 1D NMR spectrum (*26*). The structural and thermodynamic results provide a database to probe structure−stability relationships. All of the loops have the $^{GA}_{AG}$ nearest neighbor. Thus, differences in stability and structure presumably reflect the differences in pairing within the adjacent base pairs and stacking between a GA pair and its nearest neighbor.

Five of the six internal loops in Table 1 contain two sheared GA pairs. Only $^{5'GGAC3'}_{3'CAGG5'}$ has imino GA pairs. It is the most stable internal loop at 37 °C, about 2 kcal/mol more stable than the next most stable internal loop, $^{5'CGAG3'}_{3'GAGC5'}$. The extra stability of $^{5'GGAC3'}_{3'CAGG5'}$ may reflect the smaller distortion from A-form-like structure for imino GA pairs compared to that of the sheared GA pairs. Stacking interactions within the loop and with the adjacent helix must also be important. Interestingly, the least stable internal loop is $^{5'GGAU3'}_{3'UAGG5'}$, which is about 1 kcal/mol less stable than the next least stable internal loop, $^{5'UGAA3'}_{3'AAGU5'}$. The other sheared GA internal loops have similar stabilities, averaging 0.1 ± 0.6 kcal/mol at 37 °C. Most of this variation can be attributed to the extra hydrogen bond in each closing pair of $^{5'CGAG3'}_{3'GAGC5'}$, relative to $^{5'UGAG3'}_{3'GAGU5'}$, $^{5'AGAU3'}_{3'UAGA5'}$, and $^{5'UGAA3'}_{3'AAGU5'}$. The extra hydrogen bonds are expected to enhance stability by about 1.3 kcal/mol (*19*). This suggests that interactions in the $^{5'GGAU3'}_{3'UAGG5'}$ internal loop differ from those in the other sheared GA structures. As discussed below, the instability of the $^{5'GGAU3'}_{3'UAGG5'}$ internal loop is accompanied by unexpected structural features in the closing GU pairs.

Context Dependent Structural Interactions. As outlined above and shown in Figures 2 and 3, the base pair stacking in (rGC**U**G**A**G**G**CU)$_2$ is established by several spectroscopic features, which result in the following stacking patterns: major groove, U3/G4/G4*/U3* and minor groove, G6/A5/A5*/G6* (Figures 6 and 7). Generally, the adenines overlap more than the guanines in the GA pairs, possibly because of the continuous spine of purine bases forming in the minor groove. In comparison, the observed NOE patterns shown in Figures 2 and 3 for (rGC**G**G**A**U**A**UGCU)$_2$ result in a major groove G3/G4/G4*/G3* stack and a minor groove U6/A5/A5*/U6* stack (Figures 6 and 7). In contrast to (rGC**U**G**A**G**G**CU)$_2$, the major groove of (rGC**G**G**A**U**A**UGCU)$_2$ is characterized by a continuous spine of guanines.

The internal loop of (rGC**U**G**A**G**G**CU)$_2$ is characterized by different functional groups in the major and minor grooves. Because of the U3/G4/G4*/U3* stacking pattern,

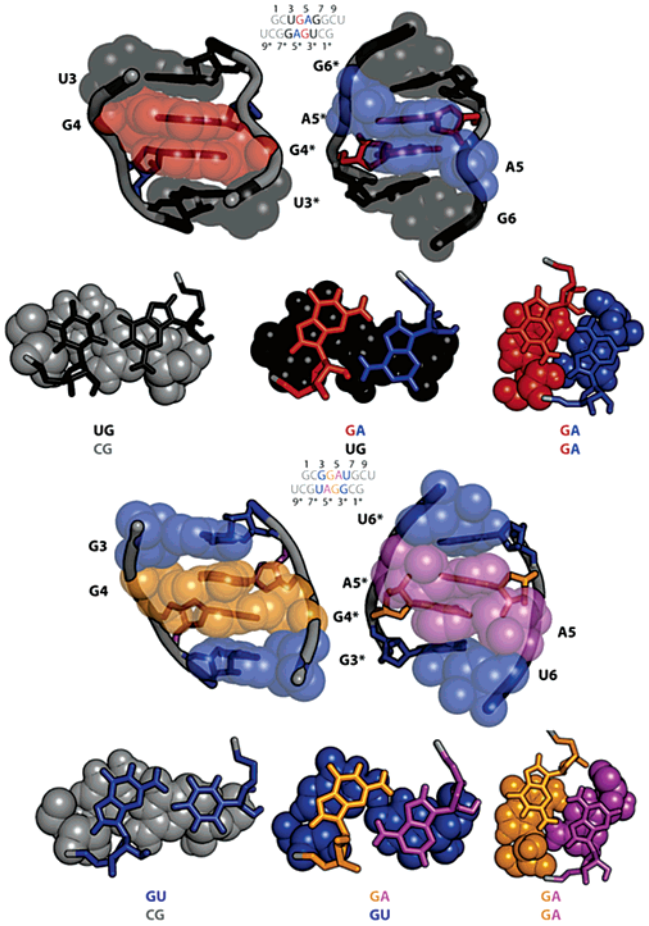

FIGURE 6: Major and minor groove view of the packing and stacking topology in the internal loops of (top) (rGC**U**GA**G**GCU)$_2$ and (bottom) (rGC**G**GA**U**GCU)$_2$. The loop bases are represented as transparent van der Waals spheres (front) and sticks (back). The color scheme is indicated in the sequence below each structure. Minimized average structures are shown. Hydrogens and nonbridging oxygen atoms have been removed to emphasize base ring overlap.

the major groove has four adjacent carbonyl groups and two adjacent N7 groups providing six hydrogen bond acceptors in the major groove (Figure 7). This arrangement could serve as a docking site for metal ions, protein side chains, or small molecules. Conversely, the minor groove is characterized by a tandem AA arrangement flanked by guanosines from the closing GU pairs. The stacking topology of the adenines results in the N1-C2-N3 base edges forming the minor groove. It has been suggested that this type of stacking pattern could provide the molecular determinants for the formation of A-minor or other tertiary interactions (*10*). As expected for a GU wobble pair, the G6/6* amino groups (denoted as H22 in Figure 7) protrude into the minor groove capping the side-by-side AA arrangement.

The G3/G4/G4*/G3* stacking pattern in the major groove of (rGC**G**GA**U**GCU)$_2$ places eight hydrogen bond acceptors in the major groove, two more than for (rGC**U**GA**G**GCU)$_2$ (Figure 7). The minor groove surface of (rGC**G**GA**U**GCU)$_2$ is similar to (rGC**U**GA**G**GCU)$_2$ locally because of the comparable side-by-side stacking arrangement of the adenines. As a result of the $\frac{5'UG}{3'GA}$ to $\frac{5'GG}{3'UA}$ nearest neighbor switch in going from (rGC**U**GA**G**GCU)$_2$ to (rGC**G**GA**U**GCU)$_2$, the tandem AA stack in the minor groove of (rGC**G**GA**U**GCU)$_2$ is capped by U6/6*O2 instead of G6/G6* amino groups.

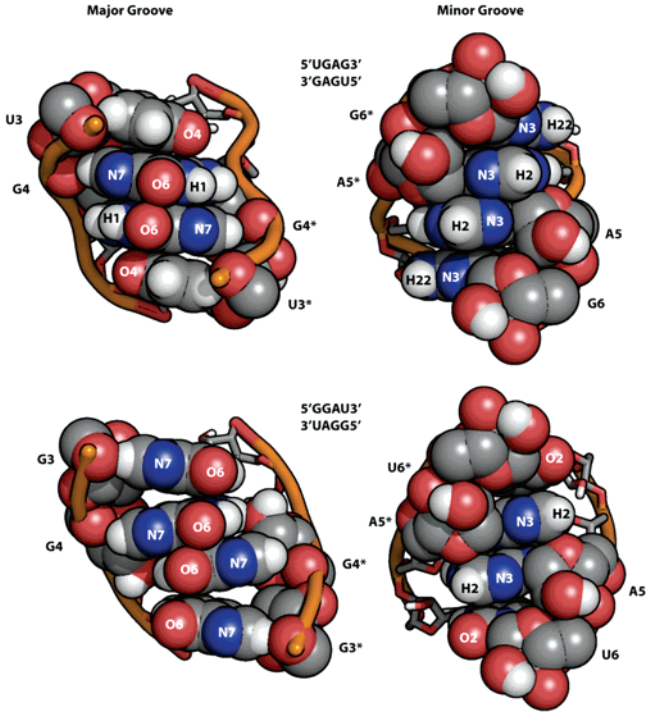

FIGURE 7: Major and minor groove functional group disposition in the internal loops of (top) (rGC**U**GA**G**GCU)$_2$ and (bottom) (rGC**G**GA**U**GCU)$_2$. Notice the differences in the number of hydrogen bond acceptors in the major grooves of (rGC**U**GA**G**GCU)$_2$ and (rGC**G**GA**U**GCU)$_2$. In both (rGC**U**GA**G**GCU)$_2$ and (rGC**G**GA**U**GCU)$_2$, the major grooves could provide a suitable molecular surface for ligand docking, tertiary, or quaternary interactions. The loop bases are represented as transparent van der Waals spheres (front) and sticks (back). The atom color code is C, light gray; N, blue; and O, red. The phosphate backbone is depicted as an orange tube.

Thus, the minor grooves of (rGC**U**GA**G**GCU)$_2$ and (rGC**G**GA**U**GCU)$_2$ are characterized by different hydrogen bond readouts, which offers different specificity for ligands or tertiary interactions.

Context Dependent Hydrogen-Bonding Interactions. Although explicit hydrogen bonds were not measured or restrained for the GA pairs, a number of hydrogen-bonding contacts are discernible from the structures (Figure 8). For (rGC**U**GA**G**GCU)$_2$, the GA sheared pairs are stabilized by putative hydrogen bond interactions involving G4/4* N3 to A5*/5 H61 (2.9 Å) and G4/4* H21 to A5*/5 N7 (2.0 Å), consistent with other examples of sheared GA pairs (*10*, *27*, *37*). Base to sugar−phosphate backbone hydrogen bonds are also observed. The detailed nature of these interactions is predicated on the G4 sugar pucker. In (rGC**U**GA**G**GCU)$_2$, the G4 ribose displayed intermediate (~5 Hz) H1′ to H2′ scalar couplings indicative of sampling between C2′ and C3′-endo sugar conformers (data not shown). As a result, the sugar pucker for G4 was restrained to cover both the C2′ and C3′-endo conformations. For all 50 calculated structures, the G4 sugar pucker was in the C3′-endo conformation. The C3′-endo conformation of G4 results in a G4/4* H22 to A5*/5 O2P hydrogen bond (2.5 Å). Similar hydrogen bonds have been observed in other sheared GA pairs and may add stability to the loops (*27*, *46*, *64*, *65*, *68*).

In (rGC**G**GA**U**GCU)$_2$, the base-to-base hydrogen bond interactions for the GA pairs are comparable to those in (rGC**U**GA**G**GCU)$_2$. The average distances between G4/4*

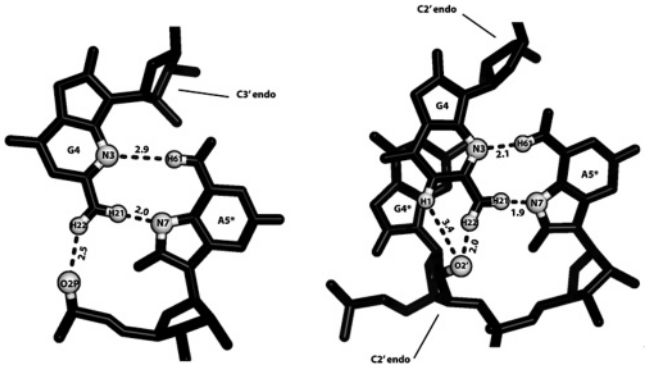

FIGURE 8: Putative hydrogen bond network in the GA sheared pairs of (left) (rGC**U**G**A**G**G**CU)$_2$ and (right) (rGC**G**G**A**U**A**GCU)$_2$. The donor and acceptor atoms are rendered as white spheres. The hydrogen bonds are shown as dashed lines connecting the atoms along with the average distance. The detailed hydrogen bond network is predicated on the G4 sugar. In (rGC**U**G**A**G**G**CU)$_2$, the G is displayed as C3′-endo (N) but is likely dynamic (see text for details). In (rGC**G**G**A**U**A**GCU)$_2$, the G4 sugar is restricted to the C2′-endo (S) conformation.

N3 to A5*/5 H61 and G4/4* H21 to A5*/5 N7 are 2.1 Å and 1.9 Å, respectively. Thus, the two internal loops form congruent sheared GA pairs stabilized by similar base-to-base hydrogen bond contacts.

Base-to-backbone hydrogen bonds in (rGC**G**G**A**U**A**GCU)$_2$ differ from those in (rGC**U**G**A**G**G**CU)$_2$, however. The G4 sugar pucker was restrained to the C2′-endo conformation on the basis of the >10 Hz G4 H1′ to H2′ cross-peak in the DQFCOSY spectrum (data not shown). As a result, the G4H22 to A5*O2P interaction observed in (rGC**U**G**A**G**G**CU)$_2$ is lost and supplanted with G4H1 to G4*O2′ (3.4 Å) and G4H21 to G4*O2′ (2.0 Å). A similar hydrogen bond interaction has been observed for the $^{5'UGAA3'}_{3'AAGU5'}$ internal loop, where the equivalent of G4 is in a C2′-endo conformation (*37*).

Context Dependent Structural Differences in the GU Closures. Perhaps the most striking difference between (rGC**U**G**A**G**G**CU)$_2$ and (rGC**G**G**A**U**A**GCU)$_2$ is the base pairing configuration of the closing/adjacent GU pairs. In (rGC**U**G**A**G**G**CU)$_2$, the closing GU pairs form a canonical wobble configuration. This is experimentally verified by the strong NOE between the imino resonances of G3/3* and U6*/6 and the upfield shift (~6.5 ppm) of the observable G6/6* amino protons (Figure 2), which is expected for a GU wobble pair because the G amino group is not in a hydrogen bond. The GU pairs in (rGC**G**G**A**U**A**GCU)$_2$, however, do not form a canonical wobble configuration. The GU pairs in (rGC**G**G**A**U**A**GCU)$_2$ are stabilized by a G3/3*H22 to U6*/6O4 single hydrogen bond (Figure 4). SantaLucia et al. observed an identical non-canonical GU pair in the 690 hairpin from *E. coli* 16S rRNA modeled with the NMR construct 5′-GGC**G**G**U**G**A**A**A**U**A**GCC-3′, where the bold nucleotides are G690 and U697, which are paired (*70*). In their NMR spectra, the G690H1 to U697H3 NOE is also missing, although the G690H1 resonance is as sharp as the stem iminos, and the U697H3 resonance is broadened by solvent exchange (*70*). The adjacent GA pair is also sheared. An analogous base pairing configuration for the 690 hairpin is found in the *E. coli* ribosome crystal structure (*71*). In the context of the ribosome, the 690 loop is involved in tertiary and quaternary contacts with the 790 hairpin loop and the

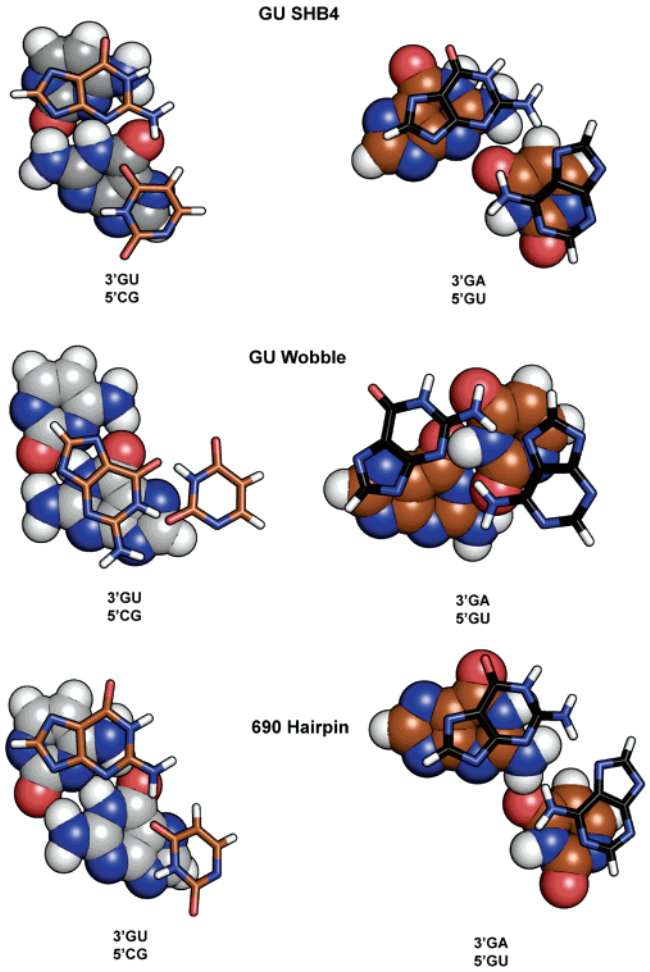

FIGURE 9: Comparison of the base-stacking differences for the $^{5'CG}_{3'GU}$ and $^{5'GG}_{3'UA}$ nearest neighbors in $^{5'GGAU3'}_{3'UAGG5'}$ modeled as an SHB4 pair (top) and wobble GU pair (middle). The NMR structure (*70*) of the *E. coli* 690 hairpin loop (bottom) is also shown. The base stacks are viewed down the helical axis with the lower base pair rendered as spheres and the upper base pair as sticks. Notice that in the SHB4 model, the GU pair overlaps more extensively with the adjacent GA pair compared to the corresponding bases in the GU wobble model. Only the base atoms are shown. The color scheme for the CG, GU, and GA pairs are the same for O (red), N (blue), and H (white) but differ for C (gray for CG, brown for GU, and black for GA pairs).

S11 protein. Evidently, the $^{5'GG}_{3'UA}$ nearest neighbor forms a noncanonical ordered structure that is not predicated on tertiary or quaternary interactions because similar structures are found for the $^{5'GG}_{3'UA}$ nearest neighbor in the context of an isolated hairpin and internal loop.

The different hydrogen-bonding patterns for (rGC**U**G**A**G**G**CU)$_2$ and (rGC**G**G**A**U**A**GCU)$_2$ suggest that a competition between stacking and hydrogen bonding determines the structure and that hydrogen bonding is not dominant. Figure 9 shows a comparison of the base overlap of the $^{5'CG}_{3'GU}$ and $^{5'GG}_{3'UA}$ nearest neighbor interactions in the NMR structure of (rGC**G**G**A**U**A**GCU)$_2$ modeled with SHB4 and GU wobble restraints. The $^{5'CG}_{3'GU}$ and $^{5'GG}_{3'UA}$ nearest neighbors taken from the NMR structure of the 690 hairpin loop of *E. coli* 16S rRNA are also shown for comparison (*70*). At the $^{5'CG}_{3'GU}$ step, the helix is overwound 3′ to the CG pair more in the wobble model compared to the SHB4 model. This results in a cross-strand stack of the Gs in the wobble conformation at the $^{5'CG}_{3'GU}$ step, which leaves the C unstacked on its 3′ side, and a

diminution of the G stacking at the $^{5'GG}_{3'UA}$ step compared to SHB4 (Figure 9). In the SHB4 structure, G3 is rotated toward the major groove, and U6 is shifted to the minor groove. This subtle perturbation of G3 and U6 provides more extensive overlap with the adjacent GA pair. Presumably, the more favorable stacking in the SHB4 structure compensates for the difference in hydrogen bonds compared to the GU wobble conformation. As shown in Figure 9, a similar stacking arrangement is observed for the $^{5'GG}_{3'UA}$ nearest neighbor in the 690 hairpin loop of *E. coli* 16S rRNA (*70*, *71*). Thus, the base stacking interactions for the $^{5'GG}_{3'UA}$ nearest neighbors are important determinants in fine-tuning the pairing configuration for this internal loop.

The two least stable tandem GA internal loops, $^{5'UGAA3'}_{3'AAGU5'}$ and $^{5'GGAU3'}_{3'UAGG5'}$ (Table 1), have less than optimal hydrogen bonds, given their hydrogen-bonding capacity. In $^{5'UGAA3'}_{3'AAGU5'}$, this is manifested as stretched UA pairs with distances of 2.8 and 3.0 Å for the AH61 to UO4 and AN1 to UH3 hydrogen bonds, respectively (*37*). Presumably, optimal hydrogen bonding would give less favorable free energy because of less favorable stacking. This implies that hydrogen bonds are not dominant in determining local geometry. Local geometry is a competition between stacking and hydrogen bonding. This type of competition has also been suggested for tandem GU pairs (*72*) and is found in molecular dynamics simulations of them (*41*). The results presented here provide more benchmarks for testing simulations (*39*, *42*).

Sugar Pucker Correlates with Stability. It has been suggested that the stability increment of the adjacent pair determines the sugar pucker of the G in tandem sheared GA pairs (*73*). For instance, in $^{5'CGAG3'}_{3'GAGC5'}$, the G in the tandem GA pair has a C3′-endo (N) conformation (*27*). Switching to $^{5'UGAA3'}_{3'AAGU5'}$ results in a G sugar pucker change to C2′-endo (S) (*37*). Moreover, in the $^{5'CGAA3'}_{3'GAGC5'}$ internal loop from the Neurospora VS ribozyme, the G that is 3′ to the CA pair is in the C2′-endo (S) conformation, whereas the G that is 3′ to the CG pair is in equilibrium between C3′-endo (N) and C2′-endo (S) (*73*). The results for $^{5'UGAG3'}_{3'GAGU5'}$ and $^{5'GGAU3'}_{3'UAGG5'}$ agree with this hypothesis (Table 1). Evidently, stacking and base-to-backbone hydrogen-bonding interactions are strong enough to affect sugar pucker as well as hydrogen bonding. This is not surprising because the free energy difference between C3′-endo and C2′-endo sugars is about 0.5 kcal/mol (*74*, *75*), whereas the base-to-base hydrogen bonds in RNA typically enhance stability by about 1 kcal/mol (*68*, *76*, *77*), and stacking interactions can contribute up to 2 kcal/mol to stability (*78*).

Correlation with Occurences in rRNA. It is probable that the structural polymorphism and thermodynamic stability of GU pairs are related to their phylogenetic occurrence because the $^{5'UG}_{3'GA}$ motif occurs more often than the $^{5'GG}_{3'UA}$ motif (*44*). In a survey of 30 of 23S and 20 of 16S rRNA secondary structures (*79*), the $^{5'UG}_{3'GA}$ motif occurs 394 times, whereas the $^{5'GG}_{3'UA}$ motif occurs only 47 times. The average number of appearances of the $^{5'UG}_{3'GA}$ motif per secondary structure for 23S and 16S rRNA are 11.4 and 2.8, respectively. Conversely, the average number of appearances of the $^{5'GG}_{3'UA}$ motif per secondary structure for 23S and 16S rRNA are 0.5 and 1.7, respectively. Evidently, there is selective pressure to maintain the $^{5'UG}_{3'GA}$ nearest neighbor interaction relative to $^{5'GG}_{3'UA}$, which may be related to the different stability and structure of $^{5'UG}_{3'GA}$ compared to $^{5'GG}_{3'UA}$. Of the 50 rRNA structures surveyed, the $^{5'UGAG3'}_{3'GAGU5'}$ internal loop occurs in the 23S rRNA subunits of *Micrococcus luteus*, *Listeria monocytogenes*, and *Lactococcus lactis*. Conversely, the $^{5'GGAU3'}_{3'UAGG5'}$ internal loop was not observed in the 50 structures. A $^{5'GGAU3'}_{3'UAGU5'}$ internal loop is found in the 5′ half of the *S. cerevisiae* 23S rRNA, however.

CONCLUSIONS

The local structure and stability of internal loops are highly sequence dependent, which presents a challenge to predicting these properties from sequence information alone. The structural polymorphism of tandem pairs depends upon several factors such as the type and number of hydrogen bonds, electrostatics, base stacking interactions, and the stability increment of the closing/adjacent base pair. A delicate balance between these factors determines local structure. This work and others provide theoreticians (*39*−*41*) with experimental structures and thermodynamic data to benchmark theories aimed at understanding the interplay between these components.

ACKNOWLEDGMENT

We thank past and present members of the Turner lab for helpful discussions regarding the manuscript. Special recognition goes to Anthony Furstoss for his help with assigning the (GC**G**GA**U**GCU)$_2$/cobalt hexamine spectra.

SUPPORTING INFORMATION AVAILABLE

Chemical shift assignments, distance restraints, structure calculation statistics, and SNOESY, natural abundance ^{1}H-^{15}N HSQC, ^{1}H-^{13}C HSQC, TOCSY, and ^{1}H-^{31}P HETCOR spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Gesteland, R. F., Cech, T. R., and Atkins, J. F. (2006) *The RNA World*, 3rd ed., Cold Spring Harbor Press, Cold Spring Harbor, NY.
2. Sevignani, C., Calin, G. A., Siracusa, L. D., and Croce, C. M. (2006) Mammalian microRNAs: a small world for fine-tuning gene expression, *Mamm. Genome 17*, 189−202.
3. Stark, A., Brennecke, J., Bushati, N., Russell, R. B., and Cohen, S. M. (2005) Animal microRNAs confer robustness to gene expression and have a significant impact on 3′ UTR evolution, *Cell 123*, 1133−1146.
4. Winkler, W. C. (2005) Riboswitches and the role of noncoding RNAs in bacterial metabolic control, *Curr. Opin. Chem. Biol. 9*, 594−602.
5. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure, *J. Mol. Biol. 288*, 911−940.
6. Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M., and Turner, D. H. (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure, *Proc. Natl. Acad. Sci. U.S.A. 101*, 7287−7292.
7. Xia, T., SantaLucia, J., Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs, *Biochemistry 37*, 14719−14735.

8. Washietl, S., Hofacker, I. L., Lukasser, M., Huttenhofer, A., and Stadler, P. F. (2005) Mapping of conserved RNA secondary structures predicts thousands of functional noncoding RNAs in the human genome, *Nat. Biotechnol. 23*, 1383−1390.
9. Uzilov, A. V., Keegan, J. M., and Mathews, D. H. (2006) Detection of non-coding RNAs on the basis of predicted secondary structure formation free energy change, *BMC Bioinf. 7*, 173.
10. Chen, G., Znosko, B. M., Kennedy, S. D., Krugh, T. R., and Turner, D. H. (2005) Solution structure of an RNA internal loop with three consecutive sheared GA pairs, *Biochemistry 44*, 2845−2856.
11. Jang, S. B., Baeyens, K., Jeong, M. S., SantaLucia, J., Jr., Turner, D., and Holbrook, S. R. (2004) Structures of two RNA octamers containing tandem G.A base pairs, *Acta Crystallogr., Sect. D 60*, 829−835.
12. Chen, G., Znosko, B. M., Jiao, X., and Turner, D. H. (2004) Factors affecting thermodynamic stabilities of RNA 3 × 3 internal loops, *Biochemistry 43*, 12865−12876.
13. Schroeder, S. J., Fountain, M. A., Kennedy, S. D., Lukavsky, P. J., Puglisi, J. D., Krugh, T. R., and Turner, D. H. (2003) Thermodynamic stability and structural features of the J4/5 loop in a *Pneumocystis carinii* group I intron, *Biochemistry 42*, 14184−14196.
14. Znosko, B. M., Burkard, M. E., Schroeder, S. J., Krugh, T. R., and Turner, D. H. (2002) Sheared Aanti. Aanti base pairs in a destabilizing 2 × 2 internal loop: the NMR structure of 5′(rGGCAAGCCU)2, *Biochemistry 41*, 14969−14977.
15. Schroeder, S. J., and Turner, D. H. (2001) Thermodynamic stabilities of internal loops with GU closing pairs in RNA, *Biochemistry 40*, 11509−11517.
16. Burkard, M. E., Xia, T., and Turner, D. H. (2001) Thermodynamics of RNA internal loops with a guanosine-guanosine pair adjacent to another noncanonical pair, *Biochemistry 40*, 2478−2483.
17. Schroeder, S. J., and Turner, D. H. (2000) Factors affecting the thermodynamic stability of small asymmetric internal loops in RNA, *Biochemistry 39*, 9257−9274.
18. Burkard, M. E., and Turner, D. H. (2000) NMR structures of r(GCAGGCGUGC)2 and determinants of stability for single guanosine-guanosine base pairs, *Biochemistry 39*, 11748−11762.
19. Schroeder, S. J., Burkard, M. E., and Turner, D. H. (1999) The energetics of small internal loops in RNA, *Biopolymers 52*, 157−167.
20. Kierzek, R., Burkard, M. E., and Turner, D. H. (1999) Thermodynamics of single mismatches in RNA duplexes, *Biochemistry 38*, 14214−14223.
21. Xia, T., McDowell, J. A., and Turner, D. H. (1997) Thermodynamics of nonsymmetric tandem mismatches adjacent to G·C base pairs in RNA, *Biochemistry 36*, 12486−12497.
22. Wu, M., SantaLucia, J., Jr., and Turner, D. H. (1997) Solution structure of (rGGCAGGCC)2 by two-dimensional NMR and the iterative relaxation matrix approach, *Biochemistry 36*, 4449−4460.
23. Wu, M., and Turner, D. H. (1996) Solution structure of (rGCG-GACGC)2 by two-dimensional NMR and the iterative relaxation matrix approach, *Biochemistry 35*, 9677−9689.
24. Schroeder, S., Kim, J., and Turner, D. H. (1996) G.A and U·U mismatches can stabilize RNA internal loops of three nucleotides, *Biochemistry 35*, 16105−16109.
25. Wu, M., McDowell, J. A., and Turner, D. H. (1995) A periodic table of symmetric tandem mismatches in RNA, *Biochemistry 34*, 3204−3211.
26. Walter, A. E., Wu, M., and Turner, D. H. (1994) The stability and structure of tandem GA mismatches in RNA depend on closing base pairs, *Biochemistry 33*, 11349−11354.
27. SantaLucia, J., Jr., and Turner, D. H. (1993) Structure of (rGGCGAGCC)$_2$ in solution from NMR and restrained molecular dynamics, *Biochemistry 32*, 12612−12623.
28. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1991) Stabilities of consecutive A·C, C·C, G·G, U·C, and U·U mismatches in RNA internal loops: Evidence for stable hydrogen-bonded U·U and C·C$^+$ pairs, *Biochemistry 30*, 8242−8251.
29. Peritz, A. E., Kierzek, R., Sugimoto, N., and Turner, D. H. (1991) Thermodynamic study of internal loops in oligoribonucleotides: symmetric loops are more stable than asymmetric loops, *Biochemistry 30*, 6428−6436.
30. Znosko, B. M., Kennedy, S. D., Wille, P. C., Krugh, T. R., and Turner, D. H. (2004) Structural features and thermodynamics of the J4/5 loop from the *Candida albicans* and *Candida dubliniensis* group I introns, *Biochemistry 43*, 15822−15837.
31. Morse, S. E., and Draper, D. E. (1995) Purine-purine mismatches in RNA helices: evidence for protonated G·A pairs and next-nearest neighbor effects, *Nucleic Acids Res. 23*, 302−306.
32. Butcher, S. E., Dieckmann, T., and Feigon, J. (1997) Solution structure of a GAAA tetraloop receptor RNA, *EMBO J. 16*, 7490−7499.
33. Butcher, S. E., Allain, F. H., and Feigon, J. (1999) Solution structure of the loop B domain from the hairpin ribozyme, *Nat. Struct. Biol. 6*, 212−216.
34. Bourdelat-Parks, B. N., and Wartell, R. M. (2005) Thermodynamics of RNA duplexes with tandem mismatches containing a uracil-uracil pair flanked by C·G/G·C or G·C/A·U closing base pairs, *Biochemistry 44*, 16710−16717.
35. Hoogstraten, C. G., Legault, P., and Pardi, A. (1998) NMR solution structure of the lead-dependent ribozyme: evidence for dynamics in RNA catalysis, *J. Mol. Biol. 284*, 337−350.
36. Campbell, D. O., and Legault, P. (2005) Nuclear magnetic resonance structure of the Varkud satellite ribozyme stem-loop V RNA and magnesium-ion binding from chemical-shift mapping, *Biochemistry 44*, 4157−4170.
37. Heus, H. A., Wijmenga, S. S., Hoppe, H., and Hilbers, C. W. (1997) The detailed structure of tandem G.A mismatched base-pair motifs in RNA duplexes is context dependent, *J. Mol. Biol. 271*, 147−158.
38. Chen, G., Kennedy, S. D., Qiao, J., Krugh, T. R., and Turner, D. H. (2006) An alternating sheared AA pair and elements of stability for a single sheared purine−purine pair flanked by sheared GA pairs in RNA, *Biochemistry 45*, 6889−6903.
39. Villescas-Diaz, G., and Zacharias, M. (2003) Sequence context dependence of tandem guanine:adenine mismatch conformations in RNA: a continuum solvent analysis, *Biophys. J. 85*, 416−425.
40. Sponer, J., Mokdad, A., Sponer, J. E., Spackova, N., Leszczynski, J., and Leontis, N. B. (2003) Unique tertiary and neighbor interactions determine conservation patterns of Cis Watson-Crick A/G base-pairs, *J. Mol. Biol. 330*, 967−978.
41. Pan, Y., Priyakumar, U. D., and MacKerell, A. D., Jr. (2005) Conformational determinants of tandem GU mismatches in RNA: insights from molecular dynamics simulations and quantum mechanical calculations, *Biochemistry 44*, 1433−1443.
42. Yildirim, I., and Turner, D. H. (2005) RNA challenges for computational chemists, *Biochemistry 44*, 13225−13234.
43. Heus, H. A., and Hilbers, C. W. (2003) Structures of non-canonical tandem base pairs in RNA helices: review, *Nucleosides, Nucleotides Nucleic Acids 22*, 559−571.
44. Gautheret, D., Konings, D., and Gutell, R. R. (1994) A major family of motifs involving G·A mismatches in ribosomal RNA, *J. Mol. Biol. 242*, 1−8.
45. Elgavish, T., Cannone, J. J., Lee, J. C., Harvey, S. C., and Gutell, R. R. (2001) AA.AG@helix.ends: A:A and A:G base-pairs at the ends of 16 S and 23 S rRNA helices, *J. Mol. Biol. 310*, 735−753.
46. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1990) Effects of GA mismatches on the structure and thermodynamics of RNA internal loops, *Biochemistry 29*, 8813−8819.
47. Leontis, N. B., Stombaugh, J., and Westhof, E. (2002) The non-Watson-Crick base pairs and their associated isostericity matrices, *Nucleic Acids Res. 30*, 3497−3531.
48. Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Tracz, D., Workman, C., Sweedler, D., Gonzalez, C., Scaringe, S., and Usman, N. (1995) Synthesis, deprotection, analysis and purification of RNA and ribozymes, *Nucleic Acids Res. 23*, 2677−2684.
49. Usman, N., Ogilvie, K. K., Jiang, M. Y., and Cedergren, R. J. (1987) Automated chemical synthesis of long oligoribonucleotides using 2′-O-silylated ribonucleoside 3′-O-phosphoramidites on a controlled-pore-glass support: Synthesis of a 43 nucleotide sequence similar to the 3′-half molecule of an Escherichia coli formylmethionine tRNA, *J. Am. Chem. Soc. 109*, 7845−7854.
50. McDowell, J. A., and Turner, D. H. (1996) Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: solution structure of (rGAGGUCUC)2 by two-dimensional NMR and simulated annealing, *Biochemistry 35*, 14077−14089.
51. Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974) Stability of ribonucleic acid double-stranded helices, *J. Mol. Biol. 86*, 843−853.
52. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277−293.

53. Goddard, T. D., and Kneller, D.G. (2004) *Sparky 3*, NMR assignment and integration software, University of California, San Francisco, CA.
54. Klosterman, P. S., Shah, S. A., and Steitz, T. A. (1999) Crystal structures of two plasmid copy control related RNA duplexes: An 18 base pair duplex at 1.20 Å resolution and a 19 base pair duplex at 1.55 Å resolution, *Biochemistry 38*, 14784−14792.
55. Gorenstein, D. G. (1984) *Phosphorus-31 NMR: Principles and Applications*, Academic Press, Orlando, FL.
56. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr., Sect. D 54*, 905−921.
57. DeLano, W. L. (2002) The PyMol Molecular Graphics System, DeLano Scientific, San Carlos, CA.
58. Lu, X. J., and Olson, W. K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures, *Nucleic Acids Res. 31*, 5108−5121.
59. Burkard, M. E., Kierzek, R., and Turner, D. H. (1999) Thermodynamics of unpaired terminal nucleotides on short RNA helixes correlates with stacking at helix termini in larger RNAs, *J. Mol. Biol. 290*, 967−982.
60. Longfellow, C. E., Kierzek, R., and Turner, D. H. (1990) Thermodynamic and spectroscopic study of bulge loops in oligoribonucleotides, *Biochemistry 29*, 278−285.
61. Varani, G., Aboulela, F., and Allain, F. H. T. (1996) NMR investigation of RNA structure, *Prog. Nucl. Magn. Reson. Spectrosc. 29*, 51−127.
62. Kieft, J. S., and Tinoco, I., Jr. (1997) Solution structure of a metal-binding site in the major groove of RNA complexed with cobalt (III) hexammine, *Structure 5*, 713−721.
63. Rudisser, S., and Tinoco, I., Jr. (2000) Solution structure of cobalt-(III)hexammine complexed to the GAAA tetraloop, and metal-ion binding to G·A mismatches, *J. Mol. Biol. 295*, 1211−1223.
64. Jucker, F. M., Heus, H. A., Yip, P. F., Moors, E. H., and Pardi, A. (1996) A network of heterogeneous hydrogen bonds in GNRA tetraloops, *J. Mol. Biol. 264*, 968−980.
65. Heus, H. A., and Pardi, A. (1991) Structural features that give rise to the unusual stability of RNA hairpins containing GNRA loops, *Science 253*, 191−194.
66. Wijmenga, S. S., Kruithof, M., and Hilbers, C. W. (1997) Analysis of 1H chemical shifts in DNA: assesment of reliability of 1H chemical shift calculations for use in structure refinement, *J. Biomol. NMR 10*, 337−350.
67. Cromsigt, J. A. M. T. C., Hilbers, C. W., and Wijmenga, S. S. (2001) Prediction of proton chemical shifts in RNA. Their use in structure refinement and validation, *J. Biomol. NMR 21*, 11−29.
68. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1992) Context dependence of hydrogen bond free energy revealed by substitutions in an RNA hairpin, *Science 256*, 217−219.
69. Lescoute, A., Leontis, N. B., Massire, C., and Westhof, E. (2005) Recurrent structural RNA motifs, isostericity matrices and sequence alignments, *Nucleic Acids Res. 33*, 2395−2409.
70. Morosyuk, S. V., Cunningham, P. R., and SantaLucia, J., Jr. (2001) Structure and function of the conserved 690 hairpin in Escherichia coli 16 S ribosomal RNA. II. NMR solution structure, *J. Mol. Biol. 307*, 197−211.
71. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. (2005) Structures of the bacterial ribosome at 3.5 A resolution, *Science 310*, 827−834.
72. Chen, X., McDowell, J. A., Kierzek, R., Krugh, T. R., and Turner, D. H. (2000) Nuclear magnetic resonance spectroscopy and molecular modeling reveal that different hydrogen bonding patterns are possible for G·U pairs: one hydrogen bond for each G·U pair in r(GGCGUGCC)(2) and two for each G·U pair in r(GAGUGCUC)(2), *Biochemistry 39*, 8970−8982.
73. Michiels, P. J., Schouten, C. H., Hilbers, C. W., and Heus, H. A. (2000) Structure of the ribozyme substrate hairpin of Neurospora VS RNA: a close look at the cleavage site, *RNA 6*, 1821−1832.
74. Brameld, K. A., and Goddard, W. A. (1999) Ab initio quantum mechanical study of the structures and energies for the pseudorotation of 5′-dehydroxy analogues of 2′-deoxyribose and ribose sugars, *J. Am. Chem. Soc. 121*, 985−993.
75. Foloppe, N., and MacKerell, A. D., Jr. (1999) Intrinsic conformational properties of deoxyribonucleosides: implicated role for cytosine in the equilibrium among the A, B, and Z forms of DNA, *Biophys. J. 76*, 3206−3218.
76. Turner, D. H., Sugimoto, N., Kierzek, R., and Dreiker, S. D. (1987) Free energy increments for hydrogen bonds in nucleic acid base pairs, *J. Am. Chem. Soc. 109*, 3783−3785.
77. SantaLucia, J., Kierzek, R., and Turner, D. H. (1991) Functional group substitutions as probes for hydrogen bonding between GA mismatches in RNA internal loops, *J. Am. Chem. Soc. 113*, 4313−4322.
78. Turner, D. H. (2000) Conformational Changes, in *Nucleic Acids: Structures, Properties, and Functions* (Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr., Eds.) pp 259−334, University Science Books, Sausalito, CA.
79. Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Muller, K. M., Pande, N., Shang, Z., Yu, N., and Gutell, R. R. (2002) The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs, *BMC Bioinf. 3*, 2.
80. Gralla, J., and Crothers, D. M. (1973) Free energy of imperfect nucleic acid helices. 3. Small internal loops resulting from mismatches, *J. Mol. Biol. 78*, 301−319.

BI061350M