BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2010,** *49,* 5817–5827 5817
DOI: 10.1021/bi100332r

RNA Internal Loops with Tandem AG Pairs: The Structure of the 5′GAGU/3′UGAG Loop Can Be Dramatically Different from Others, Including 5′AAGU/3′UGAA[†]

Nicholas B. Hammond,[‡] Blanton S. Tolbert,[§] Ryszard Kierzek,[∥] Douglas H. Turner,*[,‡] and Scott D. Kennedy[§]

[‡]*Department of Chemistry, University of Rochester, Rochester, New York 14627-0216,* [∥]*Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland, and* [§]*Department of Biochemistry and Biophysics, University of Rochester School of Medicine and Dentistry, Rochester, New York 14642*

Received March 4, 2010; Revised Manuscript Received May 14, 2010

ABSTRACT: Thermodynamic stabilities of 2 × 2 nucleotide tandem AG internal loops in RNA range from −1.3 to +3.4 kcal/mol at 37 °C and are not predicted well with a hydrogen-bonding model. To provide structural information to facilitate development of more sophisticated models for the sequence dependence of stability, we report the NMR solution structures of five RNA duplexes: (rGACGAGCGUCA)$_2$, (rGACUAGAGUCA)$_2$, (rGACAAGUGUCA)$_2$, (rGGUAGGCCA)$_2$, and (rGACGAGUGUCA)$_2$. The structures of these duplexes are compared to that of the previously solved (rGGCAGGCC)$_2$ (Wu, M., SantaLucia, J., Jr., and Turner, D. H. (1997) *Biochemistry 36*, 4449−4460). For loops bounded by Watson−Crick pairs, the AG and Watson−Crick pairs are all head-to-head imino-paired (*cis* Watson−Crick/Watson−Crick). The structures suggest that the sequence-dependent stability may reflect non-hydrogen-bonding interactions. Of the two loops bounded by G-U pairs, only the 5′UAGG/3′GGAU loop adopts canonical UG wobble pairing (*cis* Watson−Crick/Watson−Crick), with AG pairs that are only weakly imino-paired. Strikingly, the 5′GAGU/3′UGAG loop has two distinct duplex conformations, the major of which has both guanosine residues (G4 and G6 in (rGACGAGUGUCA)$_2$) in a *syn* glycosidic bond conformation and forming a sheared GG pair (G4-G6*, GG *trans* Watson−Crick/Hoogsteen), both uracils (U7 and U7*) flipped out of the helix, and an AA pair (A5-A5*) in a dynamic or stacked conformation. These structures provide benchmarks for computational investigations into interactions responsible for the unexpected differences in loop free energies and structure.

Understanding RNA's functions in a plethora of cellular processes, including catalysis (*1*−*3*), and in selective alteration of those processes (*4*−*7*) begs further investigation into the connections between sequence, structure, and function. Elucidation of secondary and three-dimensional structure, however, cannot keep up with the rapid acquisition of primary sequence. With the assistance of experimental data, such as thermodynamics, NMR spectra, and chemical and enzymatic mapping, computational predictions can greatly increase the speed for discovery of structure (*8*−*15*). Still, the combination of these methods falls short of absolute accuracy due to an incomplete grasp of the various interactions involved in RNA secondary and three-dimensional structure formation. Thus, further investigations of folding interactions are needed.

The symmetric 2 × 2 nucleotide internal loop is a motif with a wide range of thermodynamic stabilities that are not explained by simple models (*16*−*24*). For example, 2 × 2 internal loops with the motif 5′AG/3′GA have free energy increments that range from −1.3 to +3.4 kcal/mol at 37 °C depending on the adjacent canonical base pair (Table 1). This spread corresponds to an ∼2000-fold range in equilibrium constant for folding. Structural models will provide a good starting point for theoreticians to develop explanations for the range of stabilities, but such structural information for these loops is lacking. Therefore, NMR structures were determined for five duplexes with the 5′AG/3′GA motif (see Table 1 for sequences and abbreviations).

At 37 °C, the order of thermodynamic stabilities of symmetric AG internal loops closed by canonical base pairs is 5′GAGC/3′CGAG > 5′CAGG/3′GGAC > 5′UAGA/3′AGAU > 5′AAGU/3′UGAA ≥ 5′GAGU/3′UGAG > 5′UAGG/3′GGAU (Table 1). Only the loops bounded by GC or CG base pairs are stabilizing, i.e., have a negative ΔG°_{loop}, and the range of $\Delta\Delta G^\circ_{loop}$ exhibited is not that expected from a simple hydrogen-bonding model. For example, the 5′UAGA/3′AGAU loop is 2.5 kcal/mol more stable than the 5′UAGG/3′GGAU loop at 37 °C, amounting to an ∼50-fold increase in binding constant. Watson−Crick pairs provide more stable interactions, but if one presumes the expected base pairing and number of hydrogen bonds, then the difference in stabilities is not explained.

More subtle differences are also observed. For example, 5′GAGC/3′CGAG is ∼0.6 kcal/mol more stable than 5′CAGG/3′GGAC, but NMR data indicate that both loops have *cis* Watson−Crick/Watson−Crick (imino) AG pairs adjacent to Watson−Crick GC pairs (Table 1). Similarly, the 5′UAGA/3′AGAU loop is ∼0.8 kcal/mol more stable than the 5′AAGU/3′UGAA loop despite both apparently having identical pairing motifs and thus the same number of hydrogen bonds (*17*).

Comparisons of the thermodynamics for 5′AG/3′GA internal loops with those of 5′GA/3′AG internal loops with identical closing base pairs are also interesting. For instance, the 5′GGAC/3′CAGG loop is 1.3 kcal/mol more stable than the 5′GAGC/3′CGAG loop, even though both have imino-paired GA pairs on

[†]This work was supported by NIH Grant GM22939 (to D.H.T.)
*Corresponding author. Phone: 585-275-3207. Fax: 585-276-0205. E-mail: turner@chem.rochester.edu.

Published on Web 05/20/2010

Table 1: Thermodynamics and Structural Features of 2 × 2 Nucleotide Internal Loops with GA or AG Pairs[a]

AG sequence for NMR	loop abbreviation	measured $\Delta G^\circ_{37,loop}$ in 1 M NaCl (kcal/mol)	pairing and pucker	comparable GA loop	measured $\Delta G^\circ_{37,loop}$ in 1 M NaCl (kcal/mol)	pairing and pucker
5′-GACGAGCGUCA ACUGCGAGCAG-5′	5′GAGC/3′CGAG	−1.3[b,f] (±0.2)	*cis*-WC/WC (imino), C3′-endo	5′GGAC/3′CAGG	−2.6[b,f] (±0.1)	*cis*-WC/WC (imino), C3′-endo
5′-GGCAGGCCA ACCGGACGG-3′	5′CAGG/3′GGAC	−0.7[c,d] (±0.2)	*cis*-WC/WC (imino), C3′-endo	5′CGAG/3′GAGC	−0.7[c,f] (±0.3)	*trans*-Hoogsteen/sugar edge (shear), C3′-endo
5′-GACUAGAGUCA ACUGAGAUCAG-5′	5′UAGA/3′AGAU	+0.9[e] (±0.3)	*cis*-WC/WC (imino), C3′-endo	5′UGAA/3′AAGU	+0.7[b,f] (±0.1)	*trans*-Hoogsteen/sugar edge (shear), C2′-endo
5′-GACAAGUGUCA ACUGUGAACAG-5′	5′AAGU/3′UGAA	+1.7[e] (±0.3)	*cis*-WC/WC (imino), C3′-endo	5′AGAU/3′UAGA	+0.3[b,f] (±0.1)	*trans*-Hoogsteen/sugar edge (shear), ND
5′-GACGAGUGUCA ACUGUGAGCAG-5′	5′GAGU/3′UGAG	+2.6[e] (±0.6), [+1.3][g] (±0.4)	GG *trans*-WC/Hoogsteen, C2′-endo	5′GGAU/3′UAGG	+1.8[b,e] (±0.1)	*trans*-Hoogsteen/sugar edge (shear), C2′-endo
5′-GGUAGGCCA ACCGGAUGG-5′	5′UAGG/3′GGAU	+3.4[e] (±0.8), [+3.4][g] (±0.4)	*cis*-WC/WC (imino), C3′-endo	5′UGAG/3′GAGU	+0.1[b,e] (±0.5)	*trans*-Hoogsteen/sugar edge (shear), C3′/C2′-endo

[a]Measured free energy is calculated by the equation (*49*) $\Delta G^\circ_{37,loop} = \Delta G^\circ_{37,duplex\ with\ loop} - (\Delta G^\circ_{37,duplex\ without\ loop} - \Delta G^\circ_{37,interrupted\ base\ pair})$. Pairing uses the nomenclature of ref *12* and is based on NMR results in refs *17*, (*21*−*24*), and *50* and this work. Sugar puckers are for the A in AG loops and for the G in GA loops, which are from refs (*21*−*24*) and *50* and this work. All sugar puckers are based on coupling constants. [b]Reference *17*. [c]Reference *47*. [d]Reference *48*. [e]Reference *18*. [f]Reference *9*. [g]Values in square brackets are calculated using measured values for duplex with tandem AG pairs (*18*) and nearest neighbor parameters for stacks of canonical pairs in measured duplex. This avoids correcting for tandem GU nearest neighbor in measured core duplex that lacks tandem AG pairs.

the basis of 1D NMR spectra (*17*). Similarly, the 5′UGAG/3′GAGU loop is 3.3 kcal/mol more stable at 37 °C than 5′UAGG/3′GGAU, although in this case the GA pairing changes from *trans* Hoogsteen/sugar edge (sheared) to *cis* Watson−Crick/Watson−Crick (imino) (Table 1). In contrast, the stabilities of 5′CAGG/3′GGAC and 5′CGAG/3′GAGC are identical at 37 °C, which is also essentially true for 5′UAGA/3′AGAU and 5′UGAA/3′AAGU at 37 °C. There is also a structural change in these cases; the 5′GA/3′AG loops have sheared GA pairs, while the 5′AG/3′GA loops have imino AG pairs (Table 1). Consideration of possible hydrogen-bonding patterns and backbone contortions indicates that while 5′GA/3′AG internal loops can adopt either sheared or imino-pairing conformations, 5′AG/3′GA internal loops flanked by Watson−Crick pairs are constrained to be only imino-paired (Table 1) as predicted by Gautheret et al. (*25*).

Here, we report the solution structures of 5′AG/3′GA internal loops closed by GC, UA, AU, GU, and UG pairs and compare them to the solution structure for the 5′CAGG/3′GGAC loop (*23*). The structures of the AG pairs bound by Watson−Crick closing pairs are all head-to-head imino-paired (*cis* Watson−Crick/Watson−Crick) (Table 1). As such, the number and types of AG hydrogen bonds cannot account for the stability differences for loops closed by Watson−Crick pairs.

Of AG loops bound by GU or UG pairs, only 5′UAGG/3′GGAU adopts wobble UG closing pairs and apparently weak imino AG pairs in the loop. For 5′GAGU/3′UGAG, there are two duplex conformations. These were resolved by bromination of C-8 on G6 of $(GACGAGUGUCA)_2$, which precludes the *anti* conformation of G6 (*26*−*28*). The 5′GAGU/3′UGAG loop's major conformation is novel, with both of the loop guanines, G4 and G6, having a *syn* glycosidic bond and U7 flipped out of the helix. The loop guanines form a GG pair (G4-G6*, G·G *trans* Watson−Crick/Hoogsteen) that has been reported in the ribosome (*29*, *30*) and in an ATP/AMP-binding RNA aptamer (*31*, *32*), but never with both guanines in the *syn* conformation. These solution structures provide benchmarks for theoreticians exploring the underlying causes of the sequence dependence of loop stability and structure.

MATERIALS AND METHODS

RNA Preparation and Purification. Oligonucleotides, rGACGAGCGUCA, rGACGAGUGUCA, rGACAAGUGUCA, rGACUAGAGUCA, and rGGUAGGCCA, were purchased from Dharmacon RNA Technologies or IDT. Oligonucleotides with a modified G, rGACGA^{Br}GUGUCA, rGAC^{Br}GAGUGUCA, and rGACGA^{Me}GUGUCA, were synthesized as previously described (*33*, *34*). "NMR buffer" is 10 mM sodium phosphate with 0.5 mM Na_2EDTA and 80 mM NaCl at pH 6.1, which has been filter-sterilized with Corning 0.22 μm PES filters. Samples for NMR spectra were dissolved in RNase-free water, dialyzed against NMR buffer for 48 h at 4 °C, dried down, and resuspended with RNase-free 90% water/10% D_2O, in a volume equal to that removed from dialysis. For spectra in D_2O, oligonucleotides were lyophilized and resuspended in 99.96% D_2O three times and then lyophilized and resuspended in 99.996% D_2O from Cambridge Isotopes. The single strand concentrations for rGACGAGCGUCA, rGACGAGUGUCA, rGACAAGUGUCA, rGACUAGAGUCA, and rGGUAGGCCA were each ~2.0 mM, unless otherwise noted. Samples in NMR tubes were incubated in a water bath at 80 °C for 5 min and then allowed to slowly cool to 4 °C over a course of ~30 min for annealing of the duplex.

NMR Spectroscopy. NMR spectra were taken on Varian Inova spectrometers at 500 or 600 MHz. The 1D imino proton spectra were taken using an S pulse for excitation (*35*). Except for $(rGGUAGGCCA)_2$, NOESY spectra in 90% H_2O/10% D_2O were acquired at 0 or 1 °C with mixing times of 75 and 150 ms. For $(rGGUAGGCCA)_2$, 125 and 150 ms spectra were acquired at 1 °C, and an additional 150 ms spectrum was taken at −3 °C. Conditions for NOESY spectra recorded in D_2O are given in Supporting Information.

Measurements of scalar couplings were derived from TOCSY, DQ-COSY, and $^{31}P-^{1}H$ HETCOR spectra. The 2D clean-TOCSY spectra (*36*) were acquired at short and medium mixing times of approximately 13 and 36 ms, respectively, with wrapping in the f_1 dimension for high resolution in the sugar proton region. DQ-COSY spectra with the same resolution as the TOCSY spectra were acquired for 5′GAGC/3′CGAG and 5′UAGG/3′GGAU. Measurement of proton−phosphorus scalar couplings and assignment of H3′ resonances were aided by $^{31}P-^{1}H$ HETCOR spectra. One-dimensional ^{31}P spectra were also acquired. Verification of peak assignments was provided by natural abundance $^{13}C-^{1}H$ heteronuclear single-quantum coherence (HSQC) spectra acquired for all duplexes. NMRPipe (*37*) was used for data processing, and Sparky (*38*) was used for peak assignments and integration. Assignments are listed in Tables S1−S5 in Supporting Information.

Restraint Generation. Distance restraints were generated from cross-peaks in 75 ms mixing time NOESY spectra using $(1/r)^6$ scaling. Cross-peaks from H5-H6 (2.45 Å) and H1′-H2′ (2.75 Å) in the Watson−Crick stems were used for reference volumes. Watson−Crick hydrogen bond restraints were applied between bases not adjacent to AG pairs as indicated by imino proton cross-peaks in NOESY spectra. Dihedral angle restraints were determined based on sugar proton and phosphorus scalar couplings taken from TOCSY, DQ-COSY, NOESY, and $^{31}P-^{1}H$ HETCOR spectra. Strong H3′-H4′ peaks and the absence of H1′-H2′ peaks in the TOCSY spectra indicated C3′-endo sugar puckers ($\delta \sim 81°$). H4′-H5′/H5″ *J*-couplings less than 2 Hz indicated γ was not in the trans or g^- conformation, and so the γ dihedral angle was restrained to g^+ ($\gamma \sim 60°$). In cases where J(H4′-H5′/H5″) > 7 Hz, γ was restrained to be trans ($\gamma \sim 180°$). $^{31}P(n+1)$-H3′(n) *J*-couplings greater than 14 Hz indicated $\varepsilon \sim -115°$ (excluding g^+). Weak ^{31}P-H5′/H5″ cross-peaks in $^{1}H-^{31}P$ HETCOR spectra (*J*-coupling < 6 Hz) indicated β in the trans conformation (~165°). Couplings within the stem residues were within typical A-form ranges for all duplexes. Consequently, backbone dihedrals in the stems were restrained to A-form values: α (−65 ± 90°), β (165 ± 75°), γ (60 ± 60°), ε (−115 ± 125°), and ζ (−70 ± 90°) as defined previously (*39*). Near the loops, angles β, γ, δ, and ε were restrained to A-form values if indicated by the data. For all duplexes except 5′GAGU/3′UGAG and 5′UAGG/3′GGAU, J(H4′-H5′/H5″) for the loop G was greater than 7 Hz, so γ was taken to be trans ($\gamma \sim 180°$), although assignments for H5′ and H5″ were not stereospecific. No cross-peak for G6P or for G5H4′ could be identified for 5′UAGG/3′GGAU, so γ and ε were not restrained. For all duplexes, α and ζ were not restrained although no phosphorus shifts were observed outside of a 1 ppm range, except for A5P of 5′GAGU/3′UGAG. Glycosidic bonds were set to be *anti* (χ = 255 ± 85°) for all residues not exhibiting large H8/H6-H1′ NOE cross-peaks. Only G4 and G6 of 5′GAGU/3′UGAG had large

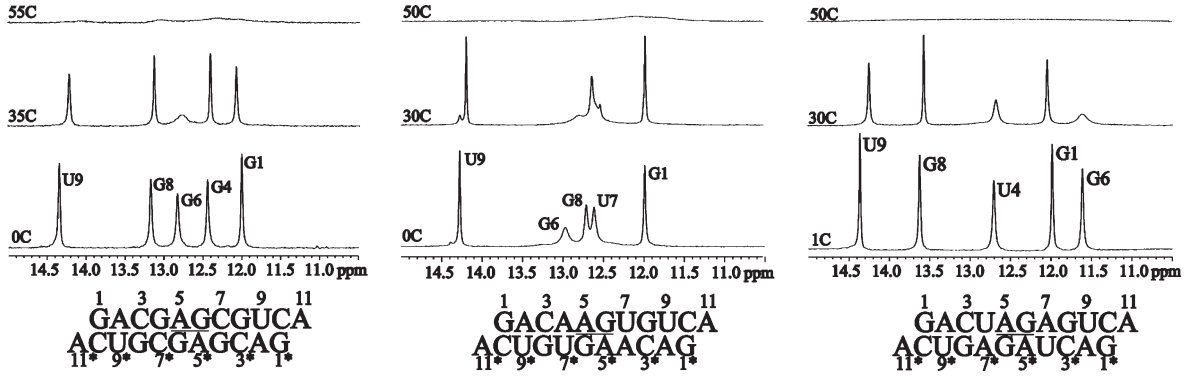

FIGURE 1: 1D NMR spectra of the imino proton regions at various temperatures for (rGACGAGCGUCA)$_2$, (rGACAAGUGUCA)$_2$, and (rGACUAGAGUCA)$_2$ (left to right) at ~2 mM strand concentration. Spectra were taken in 10 mM sodium phosphate buffer with 0.5 mM Na_2EDTA and 80 mM NaCl at pH 6.1.

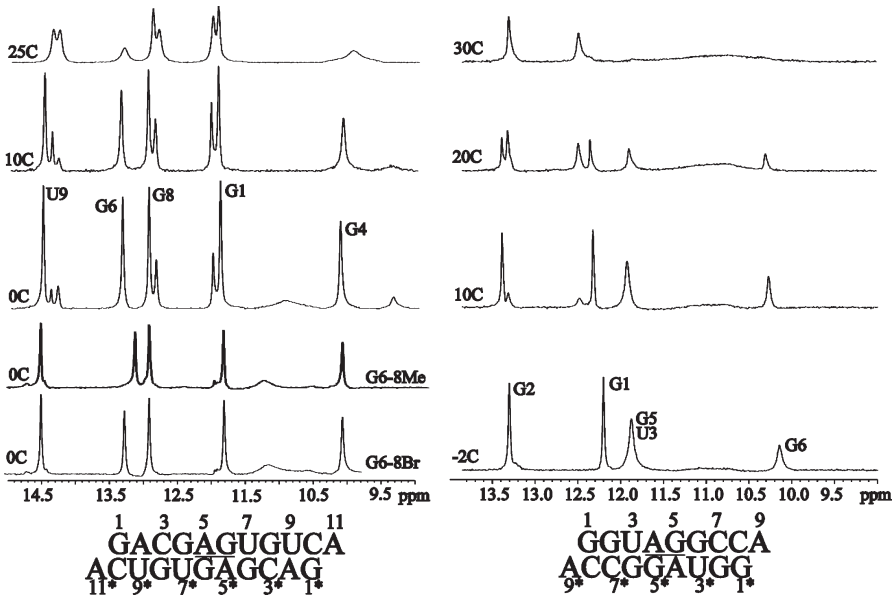

FIGURE 2: 1D NMR spectra of the imino proton regions at various temperatures for 1 mM (rGACGAGUGUCA)$_2$ (left, with 5′GA^{Br}GU/3′U^{Br}GAG at bottom and 5′GA^{Me}GU/3′U^{Me}GAG second from bottom, both at 0 °C) and 1 mM (rGGUAGGCCA)$_2$ (right). Spectra were taken in 10 mM sodium phosphate buffer with 0.5 mM Na_2EDTA and 80 mM NaCl at pH 6.1.

H8-H1′ cross-peaks. Dihedral angles in the loop and adjacent residues of 5′GAGU/3′UGAG are unusual and are discussed in Results.

Structure Calculation. Simulated annealing and molecular dynamics calculations were carried out using distance and dihedral angle restraints generated from NMR data. The structures were calculated with the following protocol using implicit solvent: (1) high temperature dynamics at 5000 K in torsion angle space for 4 ps with NOE and dihedral scale factors of 150 kcal/mol Å^2 and 25 kcal/mol rad^2, respectively; (2) simulated annealing in torsion angle space for 40 ps with slow cooling from 2000 to 0 K (40000 steps) with NOE and dihedral scale factors of 75 kcal/mol Å^2 and 100 kcal/mol rad^2, respectively; (3) simulated annealing for 40 ps in Cartesian space with slow cooling from 1000 to 0 K (40000 steps) with the NOE and dihedral angle scale factors constant at 75 kcal/mol Å^2 and 100 kcal/mol rad^2, respectively; the van der Waals factor was linearly increased from 1 to 4; and (4) Powell energy minimization was applied with full van der Waals and electrostatic terms. A total of 40 structures were calculated in this way from randomized initial atom velocities using the program CNS version 1.2 (*40*). Ten structures with the lowest total energies and without distance violations (> 0.2 Å) were subjected to an additional 100 ps of restrained MD using the program AMBER (version 9, ff99 force field). In this protocol, the system was heated to 600 K followed by slow cooling to 0 K over 100000 steps using NOE and dihedral scale factors of 20 kcal/mol Å^2 and 30 kcal/mol rad^2, respectively, and a generalized-Born implicit solvent model (*41*). Distance and dihedral restraints used in the CNS calculation were also used in the AMBER calculation. The structures of 5′AAGU/3′UGAA, 5′GAGC/3′CGAG, 5′UAGA/3′AGAU, and 5′UAGG/3′GGAU are deposited with the RCSB Protein Data Bank with ID codes 2KXZ, 2KY0, 2KY1, and 2KY2, respectively.

RESULTS

Exchangeable Proton Spectra. The 1D NMR spectra of the imino proton region for (rGACAAGUGUCA)$_2$ (denoted 5′AAGU/3′UGAA), (rGACUAGAGUCA)$_2$ (denoted 5′UAGA/3′AGAU), (rGACGAGCGUCA)$_2$ (denoted 5′GAGC/3′CGAG) (Figure 1), and (rGGUAGGCCA)$_2$ (denoted 5′UAGG/3′GGAU) (Figure 2) indicate that these duplexes have essentially a single conformation. For each of these duplexes, five peaks are observed between 10 and 14.5 ppm at ~0 °C, corresponding to G imino protons in the loops and the expected Watson−Crick pairs, though

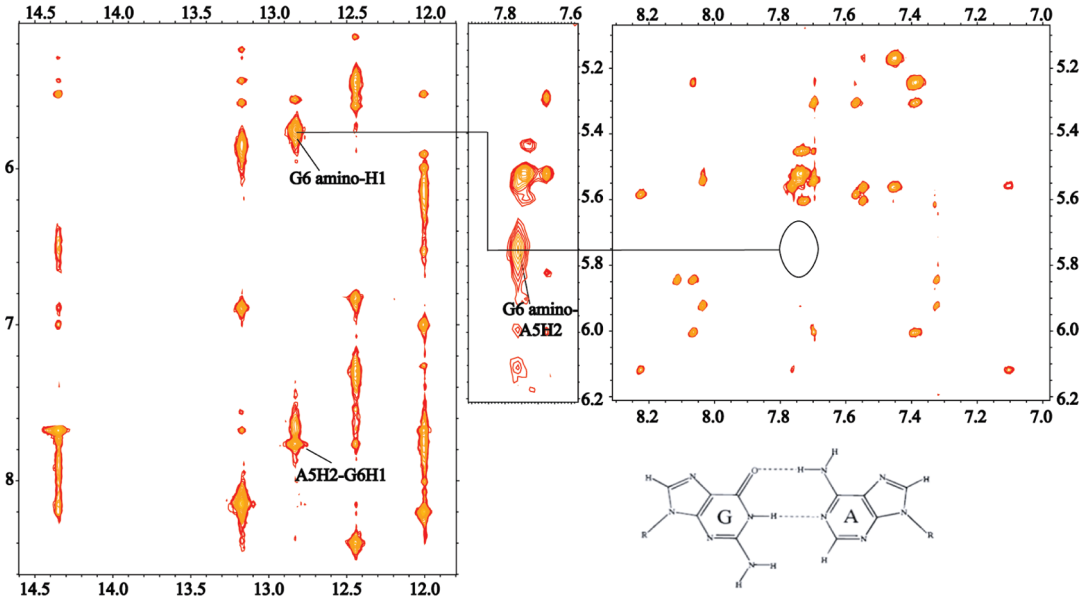

FIGURE 3: NOESY spectra of (rGACGAGCGUCA)$_2$ in H_2O at 1 °C with 150 ms mixing time (left and middle) and in D_2O at 15 °C with 400 ms mixing time (right). The black line connects the G6 amino-H1 cross-peak to the G6 amino-A5H2 cross-peak in the same spectrum and then continues on to indicate the lack of a cross-peak in D_2O, confirming the identity of the G6 amino-A5H2 cross-peak. Also of note is the A5H2-G6H1 cross-peak (left). These confirm the imino pairing in the AG loop (shown bottom, right) as these cross-peaks would not be seen in a sheared confirmation.

the G5 and U3 imino peaks of 5′UAGG/3′GGAU are overlapped (Figure 2).

The 2D SNOESY spectra for 5′AAGU/3′UGAA, 5′UAGA/3′AGAU, 5′GAGC/3′CGAG, and 5′UAGG/3′GGAU confirm the presence of the expected Watson−Crick and imino AG (*cis* Watson−Crick/Watson−Crick) pairs for each of these duplexes. The imino proton region of the 2D NOESY spectrum for 5′GAGC/3′CGAG is shown in Figure 3 and is representative of the 2D NOESY spectra for 5′AAGU/3′UGAA, 5′UAGA/3′AGAU, and 5′UAGG/3′GGAU (see Supporting Information). The imino peaks for Watson−Crick pairs were confirmed by cross-strand NOEs to cytosine amino protons for GC pairs and adenine H2s for AU pairs. The H2 protons of the loop adenines (A5H2 and A5*H2, where * represents the second strand) have strong cross-peaks to the imino and amino protons of the cross-strand loop guanosines (G6* and G6) (Figure 3). This eliminates the possibility of a sheared AG pair, because neither of these cross-peaks would be seen if the pair was sheared. The imino protons in the loops (G6H1 in Figure 1, G5H1 in 5′UAGG/3′GGAU in Figure 2) have downfield chemical shifts, which are indicative of relatively stable, hydrogen-bonded G imino protons.

In 5′AAGU/3′UGAA and 5′UAGA/3′AGAU, U7H3 and U4H3, respectively, are shifted upfield (12.6 and 12.7 ppm) relative to other Watson−Crick AU pairs in the duplexes studied, which average 14.4 ppm (Figure 1). The ^{15}N bound to U7H3 and U4H3 have their resonances at ~161 ppm, which is consistent with expectations for a UA pair (Supporting Information). The U7H3 proton is exchanging rapidly with water, which suggests weak hydrogen bonding in this closing base pair. Interestingly, the 5′AAGU/3′UGAA loop is the least thermodynamically stable of all loops closed by Watson−Crick base pairs.

The chemical shifts of G6H1 and U3H3 of 5′UAGG/3′GGAU are fairly typical of GU pairs (Figure 2), and a strong NOE cross-peak between these protons in 2D spectra is consistent with the U3 and G6 closing pair being in a GU wobble (*cis* Watson−Crick/Watson−Crick) conformation. The AG hydrogen bonding is apparently not always formed, as suggested by a relatively weak G5H1-A4H2 cross-peak compared to the duplexes with Watson−Crick closing pairs and by rapid exchange with water, as indicated by an exchange cross-peak (Figure S2 in Supporting Information). The chemical shift of the G5 imino resonance in the AG pair is nearly identical to that observed for the imino AG pair in 5′CAGG/3′GGAC (*23*). Diffuse intensity observed at ~11.0 ppm in the 1D spectrum of 5′UAGG/3′GGAU is probably due to a small population of hairpin loops. This alternate conformation becomes more evident at higher temperatures (Figure 2).

The 1D spectrum for (rGACGAGUGUCA)$_2$ (denoted 5′GAGU/3′UGAG) has 11 peaks between 9 and 15 ppm, indicating multiple conformations (Figure 2). Spectra measured after decreasing strand concentration and heating the sample to 70 °C followed by rapid cooling (to favor hairpin) or after increasing strand concentration and reannealing slowly (to favor duplex) demonstrated that there was a minor concentration of hairpin. In addition, there is at least one minor conformation of the duplex (Figure 4). The minor duplex conformation is upward of 30% of the total duplex population, however, and gives rise to a significant number of major-to-minor conformation exchange cross-peaks in the 2D spectra of 5′GAGU/3′UGAG. Replacement of the G4 or G6 H8 protons with bromine resulted in a marked reduction in minor conformation peaks, with bromination of G6H8 yielding five sharp imino resonances between 10 and 15 ppm at the same chemical shifts as the peaks of the major conformation (Figure 2). Bromination at the H8 position forces glycosidic bonds to adopt a *syn* conformation (*26*−*28*, *33*). Thus, the spectra in Figure 2 suggest that the major conformation of the 5′GAGU/3′UGAG loop has G6 in the *syn* conformation. Most of the analysis of this motif used spectra obtained from the construct with 8-bromoguanine for G6, (rGACGA^{Br}GUGUCA)$_2$ (denoted 5′GA^{Br}GU/3′U^{Br}GAG). Assignment of the imino protons was confirmed by a $^{15}N-^{1}H$ HSQC spectrum of 5′GAGU/3′UGAG (see Supporting Information). The lack of a clear U7 imino resonance and the absence of any NOE cross-peaks to it suggest that the expected U7-G4 closing wobble pair is not formed. The downfield shift of

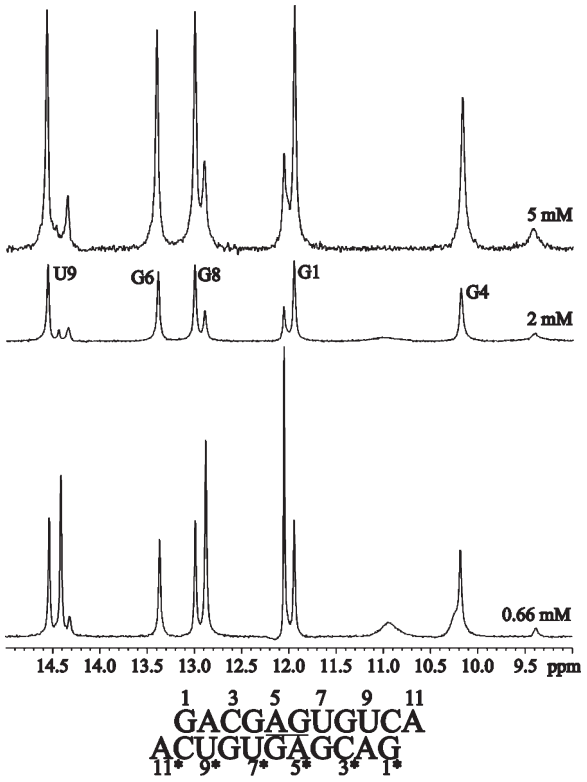

FIGURE 4: 1D NMR spectra of the imino proton region of (rGACG-AGUGUCA)$_2$ at 0 °C at 0.66, 2, and 5 mM strand concentration (bottom to top). The size of the spectrum at 2 mM is reduced to allow easier comparisons to chemical shifts in the 5 mM spectrum.

the G6 imino indicates involvement in a hydrogen bond, but no G6H1-A5H2 NOE cross-peak is observed. Thus, formation of a GA imino pair in the loop is not indicated. A strong G4*H8 - G6H1 cross-peak (Figure 5), however, suggests a G6/G4* *trans* Watson−Crick/Hoogsteen pair, where the asterisk denotes a cross-strand interaction. This is supported by the relatively upfield shift and rapid water exchange of G4H1, suggesting lack of hydrogen bonding for this proton. A G6H1-G8H1 cross-peak suggests U7 is not between G6 and G8.

A similar construct with a methyl group instead of a bromine in the H8 position of G6 was synthesized (denoted 5′GA^{Me}GU/3′U^{Me}GAG). Features in the exchangeable proton spectra were essentially the same as for 5′GA^{Br}GU/3′U^{Br}GAG (Figure 2).

Nonexchangeable Proton Spectra. The 2D NOESY spectra of nonexchangeable protons for 5′AAGU/3′UGAA, 5′UAGA/3′AGAU, 5′GAGC/3′CGAG, and 5′UAGG/3′GGAU exhibit NOE patterns mostly typical of A-form RNA, even in the loop regions (Supporting Information). For instance, loop adenine H2 protons show two weak or medium H1′ cross-peaks rather than the strong cross-peaks typically observed for sheared GA pairs. TOCSY spectra indicate weak H1′-H2′ and strong H3′-H4′ scalar coupling typical of C3′-endo sugar puckers, and ^{31}P−^{1}H correlation spectra indicate strong H3′-P and weak H5′/H5″-P scalar coupling in both Watson−Crick stems and AG loops, indicative of A-form backbone dihedrals throughout. However, J(H4′-H5′) > 7 Hz for the G in the AG imino pairs indicates that the γ dihedral angle is not in the typical g$^+$ conformation, and chemical shifts of the phosphorus, H3′, H4′, H5′, and H5″ nuclei of this G exhibit a common pattern which is distinct from

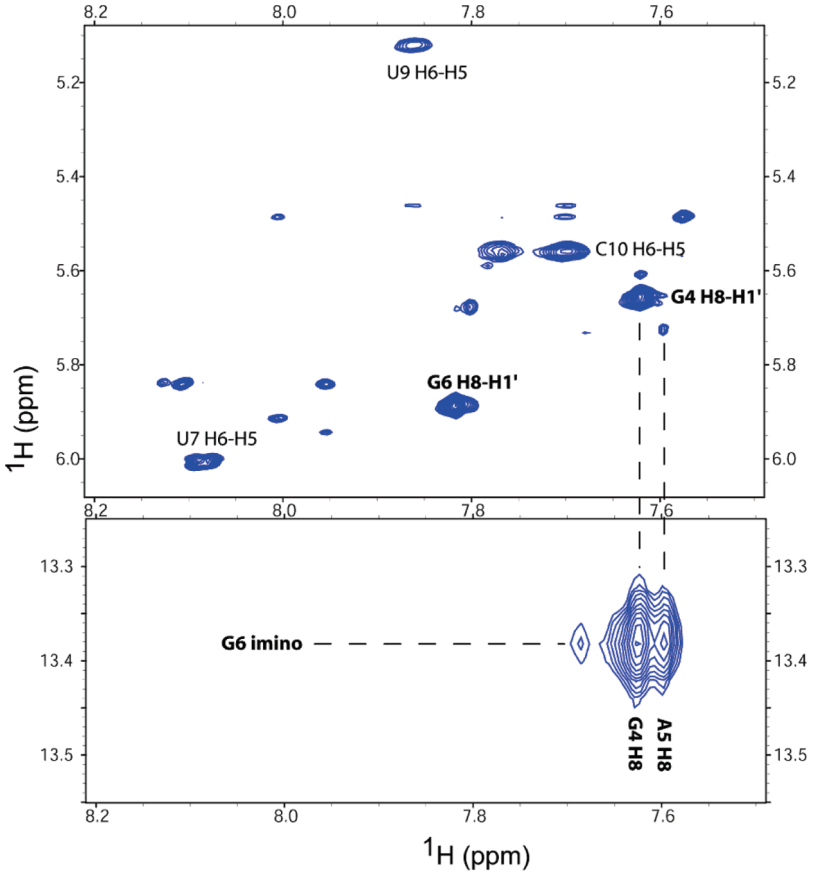

FIGURE 5: 2D NOESY spectra of (rGACGAGUGUCA)$_2$ showing large H8-H1′ cross-peaks for G4 and G6 indicating *syn* conformation and G6H1-G4*H8 indicating G6-G4* pair.

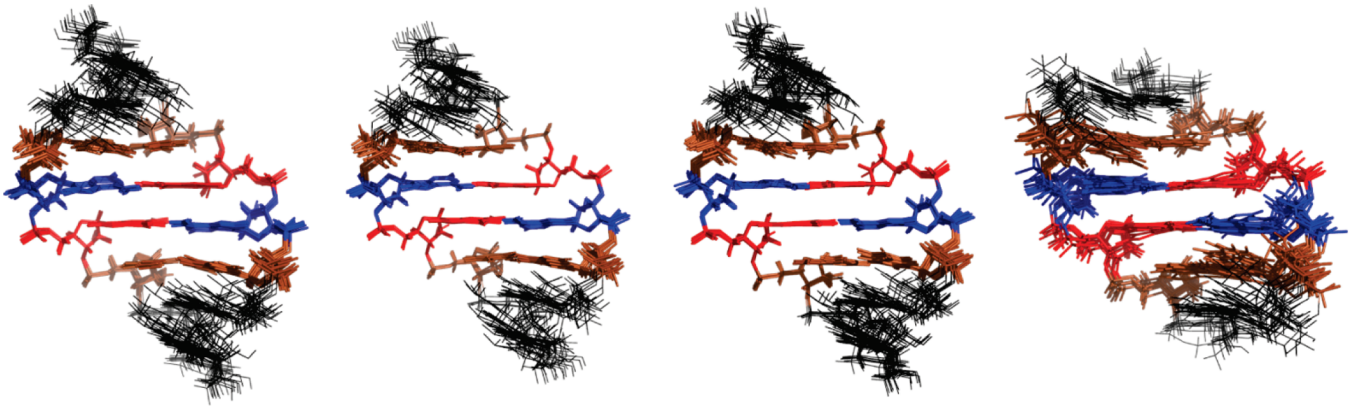

FIGURE 6: Major groove view of the lowest energy structures with no NOE distance violations for (left to right) (rGACGAGCGUCA)$_2$, (rGACAAGUGUCA)$_2$, (rGACUAGAGUCA)$_2$, and (rGGUAGGCCA)$_2$ with the terminal base pairs and dangling adenosines removed. The loop adenosines and guanines are colored blue and red, respectively, while closing base pairs are in brown and all other base pairs are in black.

equivalent nuclei in Watson−Crick stems. For 5′UAGG/3′GGAU, broad resonances for G5H1′, H2′, H3′, and H8, as well as for G6H8, and the absence of cross-peaks for G6P (possibly due to a broad G5H3′) provide further evidence of the dynamic nature of this loop.

The 2D NOESY spectrum for 5′GAGU/3′UGAG, which represents two conformations, exhibits large G4H8-H1′ and G6H8-H1′ cross-peaks, indicating that both G4 and G6 have a *syn* glycosidic bond in the major conformation (Figure 5). Furthermore, 2D NOESY spectra of the G6 8-bromo duplex, 5′GA^{Br}GU/3′U^{Br}GAG, retain the large G4H8-H1′ cross-peak seen in spectra of the unbrominated duplex, indicating that the glycosidic bond of G4 retains the *syn* conformation. The 2D NOESY spectrum of the duplex, 5′GA^{Me}GU/3′U^{Me}GAG, shows large G4H8-H1′ and G6H8Me-H1′ cross-peaks (Supporting Information). Thus, both G4 and G6 of 5′GA^{Me}GU/3′U^{Me}GAG have *syn* glycosidic bonds. Other features in spectra of the 5′GA^{Br}GU/3′U^{Br}GAG duplex suggest U7 is out of the helix. These include lack of a U7H6-G6H2′ cross-peak, medium cross-peaks for G8H8-G6H1′ and G8H8-G6H2′, a medium cross-peak for G8H8-G6H8Me (observed in 5′GA^{Me}GU/3′U^{Me}GAG; see Supporting Information), and an extremely weak or absent G8H8-U7H6 cross-peak. The ribose moieties for residues 4−7 all exhibit J(H1′-H2′) ≥ 6 Hz and J(H3′-H4′) < 2 Hz, indicating that these residues have a C2′-endo sugar pucker. Also of note is the chemical shift of A5P which is more than 1 ppm downfield from those in canonical pairs.

Information about the structure of the minor conformation of 5′GAGU/3′UGAG can be extracted from 2D NOESY spectra. In the minor conformation, G4 and G6 no longer exhibit large H1′-H8 cross-peaks, suggesting these bases are not in *syn* conformations. Also, the minor conformation chemical shifts of U7H5 and H6 are substantially upfield relative to their shifts in the major conformation, suggesting that this base is now stacked with other bases rather than being in an extrahelical orientation (Supporting Information Figure S6).

Structure Modeling. Structures were modeled by restrained molecular dynamics and simulated annealing as described in Materials and Methods. Low energy models of 5′GAGC/3′CGAG, 5′AAGU/3′UGAA, 5′UAGA/3′AGAU, and 5′UAGG/3′GGAU without violations from NMR data are overlaid in Figure 6. The stems and closing base pairs of all four loops form *cis* Watson−Crick/Watson−Crick pairs, and the loops are imino-paired (AG *cis* Watson−Crick/Watson−Crick). The purine−purine pairing widens

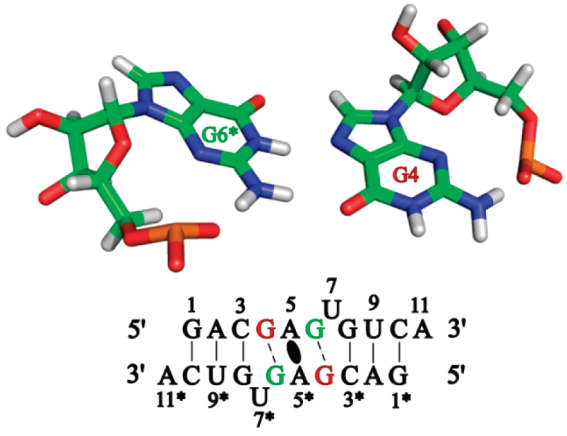

FIGURE 7: GG pair (*trans* Watson−Crick/Hoogsteen) found between G4 and G6* of (rGACGAGUGUCA)$_2$ (5′GAGU/3′UGAG) (top) and schematic of the hydrogen-bonding network of the entire (rGACGAGUGUCA)$_2$ duplex (bottom).

the duplex from an average cross-strand C1′−C1′ distance of 10.0 ± 0.1 Å for the Watson−Crick base pairs to an average of 12.5 ± 0.1 Å for the imino AG pairs. In addition, the γ dihedral angle between the two loop residues is forced into the *trans* conformation instead of the A-form g$^+$ conformation. In the models, this γ switch is accompanied by a change in the α angle from −71° to +140° (crankshaft conformation), although there is no NMR observable to verify this. Interestingly, if the NMR restraints were removed and this γ dihedral angle was forced to be g$^+$ at the beginning of a 25 ns simulation with the AMBER99 force field, then it rapidly switched to trans and returned to g$^+$ only rarely (see Supporting Information). Thus, AMBER99 predicts this structural feature well. The structural flexibility suggested by models for 5′UAGG/3′GGAU (Figure 6) is consistent with NMR spectra as described above.

A schematic of pairing in 5′GAGU/3′UGAG as inferred from NMR data is presented at the bottom of Figure 7, showing the G4-G6* base pair, an A5-A5* interaction, and U7 flipped out of the helix. The G4-G6* pair with both guanines in the *syn* conformation and forming a GG *trans* Watson−Crick/Hoogsteen pair is illustrated at the top of Figure 7. The A5-A5* interaction is represented in the schematic, though the exact conformation is unclear from the spectra. In particular, the NMR spectra are consistent with either a sheared A5-A5* pair in rapid exchange or stacking of A5 on A5*.

Chemical Shift Analysis. Some unusual chemical shifts can be compared to predictions by the program NUCHEMICS (*42*).

The average shift of an adenine H2 in an AU pair is 7.67 ppm for the structures presented here, which is also the average shift of all adenine H2s reported to the Biological Magnetic Resonance Bank (*43*). In contrast, A4H2 in 5′AAGU/3′UGAA and A7H2 in 5′UAGA/3′AGAU are shifted upfield to 6.68 and 6.63 ppm, respectively. NUCHEMICS overpredicts the upfield shift of these AH2 protons by about 0.4 ppm, at 6.24 ± 0.09 and 6.27 ± 0.14 ppm, respectively. Only two adenine H2s in AU pairs and surrounded by canonical pairs in the BMRB database have larger upfield shifts (6.4 and 6.49 ppm) than these two protons. The upfield shifts in 5′AAGU/3′UGAA and 5′UAGA/3′AGAU support the positioning of these adenine H2s in our models, where they are stacked between the pyrimidine rings of the 5′ cross-strand and 3′ intrastrand purines (Figure 8).

As shown in Table 2, the four AG imino duplexes studied here have an ~0.4 ppm upfield shift of the H2′ in the residue that is 5′ of the AG loop when compared to the chemical shifts of the H2′ protons of Watson−Crick pairs flanked by Watson−Crick pairs. Examination of spectra for three other loops with imino *cis* Watson−Crick/Watson−Crick tandem AG or GA pairing revealed similar chemical shifts (Table 2). This shift might be explained by the necessary widening of the duplex to accommodate the purine−purine pair. NUCHEMICS (*42*) underpredicts these chemical shifts by an average of 0.4 ppm.

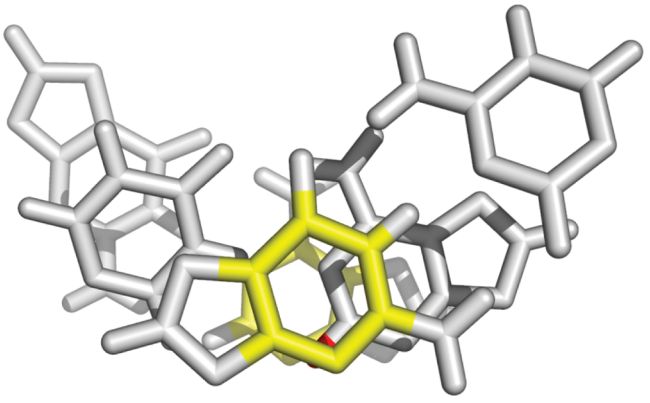

FIGURE 8: Top view of A4H2 (in red) in (rGACAAGUGUCA)$_2$ (5′AAGU/3′UGAA) showing the proton stacked between the pyrimidine rings (yellow) of the 5′ cross-strand G6* (top) and the 3′ intrastrand A5 (bottom).

DISCUSSION

Much remains to be discovered about the interactions determining RNA structure and energetics (*15*). For example, differences in free energy increments of 2 × 2 nucleotide RNA internal loops with tandem AG base pairs are not fully explained by counting hydrogen bonds. To provide structural benchmarks for theoretical calculations of interactions, we report the NMR solution structures of five tandem AG internal loops and compare them to the previously solved 5′CAGG/3′GGAC loop structure (*23*). These structures will help to provide insight into the subtle connections between sequence and thermodynamics. The structures will also contribute to prediction of 3D structures by homology modeling. A search of RNA FRABASE (*52*) revealed no 3D structures of a natural RNA with a 5′AG/3′GA internal loop. The 5′AAGU/3′UGAA loop, however, occurs in the secondary structure of pre-miR-890 RNA (*44*), and a similar loop, 5′GAGG/3′CGAC, occurs near the 3′ end of the HIV genome (*45*). Interestingly, 5′AG/3′GA represents less than 0.5% of 2 × 2 nt internal loops in a database of 1899 secondary structures whereas 5′GA/3′AG loops comprise 20% of the database (*53*).

Solution structures of the loop regions of 5′GAGC/3′CGAG, 5′AAGU/3′UGAA, 5′CAGG/3′GGAC, and 5′UAGA/3′AGAU revealed that each of the loops is wholly *cis* Watson−Crick/Watson−Crick and imino GA paired (Table 1 and Figure 6). This suggests that a subtle interaction between stacked base pairs contributes to the difference in stabilities of comparable loops.

Table 2: Chemical Shifts of H2′ Protons in Loop vs Stem[a]

sequence	avg measured δ of H2′ in WC pair not terminal or 5′ of AG or GA loop	measured δ of H2′ 5′ of AG or GA loop	NUCHEMICS prediction of δ of H2′ 5′ of AG or GA loop
5′-GACGAGCGUCA ACUGCGAGCAG-5′	4.41 ± 0.13	3.97	4.28 ± 0.14
5′-GGCAGGCC[b] CCGGACGG-3′	4.40 ± 0.14	4.06	4.64
5′-GACUAGAGUCA ACUGAGAUCAG-5′	4.41 ± 0.18	4.16	4.57 ± 0.02
5′-GACAAGUGUCA ACUGUGAACAG-5′	4.44 ± 0.14	4.15	4.63 ± 0.05
5′-GCGGACGC[c] CGCAGGCG-5′	4.36 ± 0.11	3.74	4.44
5′-GGUAGGCCA ACCGGAUGG-5′	4.43 ± 0.15	3.78	4.02 ± 0.10
5′-GGiCGAiGCCA[d] ACCiGAGiCGG-5′	4.37 ± 0.11	4.15	N/A
avg	4.41 ± 0.14	4.00 ± 0.17	4.43 ± 0.22

[a]Comparison of H2′ proton chemical shifts (δ in ppm) in residues 5′ of tandem AG or GA imino pairs with H2′ shifts of all other nonterminal Watson−Crick (WC) paired residues and with NUCHEMICS prediction of chemical shift. [b]Reference *23*. [c]Reference *22*. [d]Reference *51*.

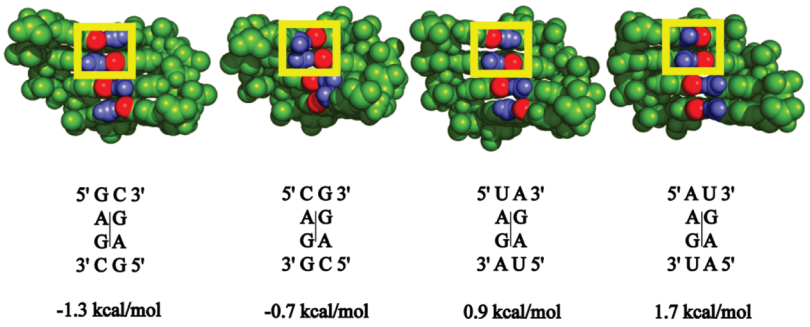

FIGURE 9: In the 5′GAGC/3′CGAG loop (left), viewed from the major groove, G4O6 (red) can be seen hydrogen bonding to the cross-strand amino of C7 (blue) and stacking above the amino of A5 (blue). Similarly, G6O6 has a pattern of hydrogen bonding to the A5 amino and stacking underneath an amino group. A similar pattern is observed for 5′UAGA/3′AGAU (second from right). When the sequence is changed to 5′CAGG/3′GGAC (second from left) or 5′AAGU/3′UGAA (right), this favorable electrostatic interaction is replaced with an unfavorable amino stacking on amino and oxygen stacking on oxygen interaction. The O-amino−O-amino interaction could add stability to 5′GAGC/3′CGAG and 5′UAGA/3′AGAU that is lost when the sequences are changed to 5′CAGG/3′GGAC and 5′AAGU/3′UGAA. The loop sequences and ΔG°_{loop} are below their corresponding model.

Because the number of hydrogen bonds is identical, the average 0.7 kcal/mol greater stability of 5′GAGC/3′CGAG relative to 5′CAGG/3′GGAC and of 5′UAGA/3′AGAU relative to 5′AAGU/3′UGAA (Table 1) must be due to other interactions in or near the loop. The structures shown in Figure 9 suggest a possible source of the sequence dependence of stability when the tandem AG loops are closed by Watson−Crick pairs. The spatial arrangement of the amino and carbonyl partial charges in the major groove appears to be more electrostatically favorable for 5′GAGC/3′CGAG and 5′UAGA/3′AGAU than for 5′CAGG/3′GGAC and 5′AAGU/3′UGAA. As shown in the yellow boxed regions in Figure 9, in the 5′GAGC/3′CGAG and 5′UAGA/3′AGAU motifs, each major groove amino group is stacked on a carbonyl group, whereas in the 5′CAGG/3′GGAC and 5′AAGU/3′UGAA motifs, each amino group is stacked on an amino group and each carbonyl group is stacked on a carbonyl group. The same structural difference may also contribute to the 2.9 kcal/mol difference in stability between the 5′GAGC/3′CGAG and 5′AAGU/3′UGAA motifs. A transition from 5′GAGC/3′CGAG to 5′AAGU/3′UGAA involves a loss of two hydrogen bonds, but on average 2 × 2 nucleotide symmetric internal loops without AG or GA pairs that are closed by GC are only 1.7 kcal/mol more stable than those closed by AU (*16*).

In Table 1, the 5′GA/3′AG motif is thermodynamically either more or equally stable to the 5′AG/3′GA motif. There are probably several reasons for this. One is that the 5′GA/3′AG motif can form either sheared or imino pairs in order to provide a backbone conformation able to accommodate a Watson−Crick pair 5′ of the A (*25*). Thus the 5′GA/3′AG motif has two options for maximizing base stacking interactions (*46*). The sheared bases are also able to make hydrogen bonds to the opposite backbone (*21*, *24*, *50*). Furthermore, it has been suggested for the 5′GA/3′AG motif that interactions of a non-hydrogen-bonded nonplanar guanosine amino group with a carbonyl group stacked on it contributes to stability (*46*). In the tandem 5′AG/3′GA motif, the loop guanosine's amino hydrogens are not close to a group with partial negative charge. The closest approach is ~3.8 Å, to the oxygen of a carbonyl group on the same strand.

The structures of the two 2 × 2 nucleotide AG symmetric loops expected to be closed by GU pairs provide a remarkable contrast. The structure of 5′UAGG/3′GGAU has the expected UG wobble and AG imino pairs. Rapid exchange of the AG imino proton with water suggests its hydrogen bond is weak or not formed at all times. Several resonances in the loop show broadening indicative of dynamic sampling of alternate conformations, as is also evident from the overlap of structures in Figure 6. The high degree of flexibility in this loop correlates with the lowest stability. In contrast, the structure of 5′GAGU/3′UGAG is quite different. First, as demonstrated in the imino proton spectra in Figures 2 and 4, the duplex exists in two conformations. In the major conformation, G4 and G6 have *syn* glycosidic bonds (Figure 7). The presence of the minor conformation at ~25−30% prevented an accurate determination of the major structure of this duplex. Bromination or methylation of carbon-8 in G6 of $(rGACGAGUGUCA)_2$ eliminated the minor conformation (Figure 2).

NMR spectra of 5′GA^{Br}GU/3′U^{Br}GAG reveal that the major conformation is novel. G4 and G6 pair with the cross-strand G6* and G4* residues, respectively, in G-G N7-imino pairs with all four G residues' glycosidic bonds in the *syn* conformation (Figure 7). This pairing forces U7 and U7* to flip out of the loop. The adenines, A5 and A5*, apparently interact in a dynamic way, though it cannot be determined whether they form a sheared pair or stack. The structural features of $(GACGAGUGUCA)_2$ are quite different from those of a shorter duplex with the 5′GAGU/3′UGAG motif, $(GCGAGUGC)_2$ (*18*). A 1D spectrum of $(GCGAGUGC)_2$ has five major peaks with chemical shifts consistent with wobble GU and imino AG pairs. This is a dramatic non-nearest neighbor structural effect. The results suggest that the shorter duplex may be more flexible, which allows interactions in the 5′GAGU/3′UGAG motif to dominate, whereas the Watson−Crick and 3′ dangling end interactions dominate in the longer duplex and provide local restraints that make the structure with wobble GU and imino AG pairs less favorable than the observed structure. In fact, NOE and chemical shifts of the minor conformation observed for $(GACGAGUGUCA)_2$ (Supporting Information Figure S6) are also consistent with a structure with wobble GU and imino AG pairs, further highlighting the thermodynamic similarity of these two structures. The results suggest that homology modeling of RNA three-dimensional structure may have to consider interactions beyond nearest neighbor.

CONCLUSIONS

The thermodynamic increments for symmetric 2 × 2 nucleotide 5′AG/3′GA internal loops have been fully mapped (*9*, *10*, *17*, *18*, *47*, *48*), but contributions leading to the observed order of stabilities are not understood. To determine the extent to

which hydrogen bonding, electrostatics, and base-stacking influence the local structure, theoreticians will need to develop more accurate models for these systems. The structures reported here provide necessary benchmarks to facilitate expansion of knowledge regarding the roles of various interactions in determining local RNA structure and stability. The major structure of $(GACGAGUGUCA)_2$ is novel and suggests that homology modeling and energetic calculations of RNA may need to consider interactions beyond nearest neighbors.

ACKNOWLEDGMENT

We thank Prof. Susan Schroeder for comments on the manuscript.

SUPPORTING INFORMATION AVAILABLE

Tables of chemical shift assignments, distance and dihedral angle restraints, structure calculation statistics and $^{1}H-^{31}P$ HETCOR, natural abundance $^{1}H-^{15}N$ HSQC, TOCSY, SNOESY, and NOESY spectra, and results of an unrestrained molecular dynamics simulation of 5′GAGC/3′CGAG. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982) Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell 31*, 147–157.
2. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell 35*, 849–857.
3. Gesteland, R. F., Cech, T. R., and Atkins, J. F. (2006) The RNA World, 3rd ed., Cold Spring Harbor Press, Cold Spring Harbor, NY.
4. Kota, J., Chivukula, R. R., O'Donnell, K. A., Wentzel, E. A., Montgomery, C. L., Hwang, H., Chang, T., Vivekanandan, P., Torbenson, M., Clark, K. R., Mendell, J. R., and Mendell, J. T. (2009) Therapeutic microRNA delivery suppresses tumorigenesis in a murine liver cancer model. *Cell 137*, 1005–1017.
5. Sul, J., Wu, C. K., Jochems, J., Lee, M. T., Kim, T. K., Peritz, T., Buckley, P., Cappelleri, D. J., Maronski, M., Kim, M., Kumar, V., Meaney, D., Kim, J., and Eberwine, J. (2009) Transcriptome transfer produces a predictable cellular phenotype. *Proc. Natl. Acad. Sci. U.S.A. 106*, 7624–7629.
6. John, M., Constien, R., Akinc, A., Goldberg, M., Moon, Y. A., Spranger, M., Hadwiger, P., Soutschek, J., Vornlocher, H. P., Manoharan, M., Stoffel, M., Langer, R., Anderson, D. G., Horton, J. D., Koteliansky, V., and Bumcrot, D. (2007) Effective RNAi-mediated gene silencing without interruption of the endogenous microRNA pathway. *Nature 449*, 745–747.
7. Childs, J. L., Disney, M. D., and Turner, D. H. (2002) Oligonucleotide directed misfolding of RNA inhibits *Candida albicans* group I intron splicing. *Proc. Natl. Acad. Sci. U.S.A. 99*, 11091–11096.
8. 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.
9. 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.
10. 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.
11. Hart, J. M., Kennedy, S. D., Mathews, D. H., and Turner, D. H. (2008) NMR-assisted prediction of RNA secondary structure: Identification of a probable pseudoknot in the coding region of an R2 retrotransposon. *J. Am. Chem. Soc. 130*, 10233–10239.
12. Stombaugh, J., Zirbel, C. L., Westhof, E., and Leontis, N. B. (2009) Frequencey and isostericity of RNA base pairs. *Nucleic Acids Res. 37*, 2294–2312.
13. Deigan, K. E., Li, T. W., Mathews, D. H., and Weeks, K. M. (2009) Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci. U.S.A. 106*, 97–102.
14. Parisien, M., and Major, F. (2008) The MC-Sym pipeline infers RNA structure from sequence data. *Nature 452*, 51–55.
15. Yildirim, I., and Turner, D. H. (2005) RNA challenges for computational chemists. *Biochemistry 44*, 13225–13234.
16. Christiansen, M. E., and Znosko, B. M. (2008) Thermodynamic characterization of the complete set of sequence symmetric tandem mismatches in RNA and an improved model for predicting the free energy contribution of sequence asymmetric tandem mismatches. *Biochemistry 47*, 4329–4336.
17. 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.
18. Schroeder, S. J., and Turner, D. H. (2001) Thermodynamic stabilities of internal loops with GU closing pairs in RNA. *Biochemistry 40*, 11509–11517.
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. 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.
21. Heus, H. A., Wijmenga, S. S., Hoppe, H., and Hilbers, C. W. (1997) The detailed structure of tandem G·A mismatched basepair motifs in RNA duplexes is context dependent. *J. Mol. Biol. 271*, 147–158.
22. Wu, M., and Turner, D. H. (1996) Solution structure of $(rGCGGACGC)_2$ by two-dimensional NMR and the iterative relaxation matrix approach. *Biochemistry 35*, 9677–9689.
23. 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.
24. Tolbert, B. S., Kennedy, S. D., Schroeder, S. J., Krugh, T. R., and Turner, D. H. (2006) NMR structures of $(rGCUGAGGCU)_2$ and $(rGCGGAUGCU)_2$: Probing the structural features that shape the thermodynamic stability of GA pairs. *Biochemistry 46*, 1511–1522.
25. Gautheret, D. F., Konings, D., and Gutell, R. R. (1994) A major family of motifs involving GA mismatches in ribosomal RNA. *J. Mol. Biol. 242*, 1–8.
26. Tavale, S. S., and Sobell, H. M. (1970) Crystal and molecular structure of 8-bromoguanosine and 8-bromoadenosine, two purine nucleosides in the *syn* conformation. *J. Mol. Biol. 48*, 109–123.
27. Michelson, A. M., Monny, C., and Kapuler, A. M. (1970) Poly 8-bromoguanylic acid. *Biochim. Biophys. Acta 217*, 7–17.
28. Ikehara, M., Uesugi, S., and Yoshida, K. (1972) Studies on the conformation of purine nucleosides and their 5′ phosphates. *Biochemistry 11*, 830–836.
29. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science 11*, 905–920.
30. Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Wimberly, B. T., and Ramakrishnan, V. (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature 407*, 340–348.
31. Dieckmann, T., Suzuki, E., Nakamura, G. K., and Feigon, J. (1996) Solution structure of an ATP-binding RNA aptamer reveals a novel fold. *RNA 2*, 628–640.
32. Jiang, F., Fiala, R., Live, D., Kumar, R, A., and Patel, D. J. (1996) RNA folding topology and intermolecular contacts in the AMP-RNA aptamer complex. *Biochemistry 35*, 13250–13266.
33. Proctor, D. J., Kierzek, E., Kierzek, R., and Bevilacqua, P. C. (2003) Restricting the conformational heterogeneity of RNA by specific incorporation of 8-bromoguanosine. *J. Am. Chem. Soc. 125*, 2390–2391.
34. Xu, Y., Ikeda, R., and Sugiyama, H. (2003) 8-methylguanosine: A powerful Z-DNA stabilizer. *J. Am. Chem. Soc. 125*, 13519–13524.
35. Smallcombe, S. H. (1993) Solvent suppression with symmetrically-shifted pulses. *J. Am. Chem. Soc. 115*, 4776–4785.
36. Cavanagh., J., and Rance, M. (1992) Suppression of cross-relaxation effects in TOCSY spectra via a modified dipsi-2 mixing sequence. *J. Magn. Reson. 96*, 670–678.
37. 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.
38. Goddard, T. D., and Kneller, D. G. (2004) Sparky 3, NMR assignment and integration software, University of California, San Francisco, CA.

39. Richardson, J. S., Schneider, B., Murray, L., Kapral, G., Immormino, R., Headd, J., Richardson, D., Ham, D., Hershkovits, E., Williams, L., Keating, K., Pyle, A., Micallef, D., Westbrook, J., and Berman, H. (2008) RNA backbone: Consensus all-angle conformers and modular string nomenclature (an RNA Ontology Consortium contribution). *RNA 14*, 465–481.
40. 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.
41. Case, D. A., Darden, T. A., Cheatham, III, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S. Roitberg, A., Seabra, G., Wong, K. F., Paesani, K. F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P, Mathews, D. H., Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) AMBER 9, University of California, San Francisco.
42. Wijmenga, S. S., Kruithof, M., and Hilbers, C. W. (1997) Analysis of ^{1}H chemical shifts in DNA: Assessment of reliability of ^{1}H chemical shift calculations for use in structure refinement. *J. Biomol. NMR 10*, 337–350.
43. Ulrich, E. L., Akutsu, H., Doreleijers, J. F., Harano, Y., Ioannidis, Y. E., Lin, J., Livny, M., Mading, S., Maziuk, D., Miller, Z., Nakatani, E., Schulte, C. F., Tolmie, D. E., Wenger, R. K., Yao, H., and Markley, J. L. (2007) BioMagResBank. *Nucleic Acids Res. 36*, D402–D408 (DOI: 10.1093/nar/gkm957).
44. Sun, G., Yan, J., Nolther, K., Feng, J., Li, H., Sarkis, D. A., Sommer, S. S., and Rossi, J. J. (2009) SNPs in human miRNA genes affect biogenesis and function. *RNA 15*, 1640–1651.
45. Watts, J. M., Dang, K. K., Gorelick, R. J., Leonard, C. W., Bess, J. W., Jr., Swanstrom, R., Burch, C. L., and Weeks, K. M. (2009) Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature 460*, 711–719.
46. Yildirim, I., Stern, H. A., Sponer, J., Spackova, N., and Turner, D. H. (2009) Effects of restrained sampling space and nonplanar amino groups on free-energy predictions for RNA with imino and sheared tandem GA base pairs flanked by GC, CG, iGiC or iCiG base pairs. *J. Chem. Theory Comput. 5*, 2088–2100.
47. 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.
48. Xia, T. B., McDowell, J. A., and Turner, D. H. (1997) Thermodynamics of nonysmmetric tandem mismatches adjacent to G-C base pairs in RNA. *Biochemistry 36*, 12486–12497.
49. 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.
50. SantaLucia, J., Jr., and Turner, D. H. (1993) Structure of $(rGGCGAGCC)_2$ in solution from NMR and restrained molecular dynamics. *Biochemistry 47*, 12612–12623.
51. Chen, G., Kierzek, R., Yildirim, I., Krugh, T. R., Turner, D. H., and Kennedy, S. D. (2007) Stacking effects on local structure in RNA: Changes in the structure of tandem GA pairs when flanking GC pairs are replaced by isoG-isoC pairs. *J. Phys. Chem. B 111*, 6718–6727.
52. Popenda, M., Blazewicz, M., Szachniuk, M., and Adaniak, R. W. (2008) RNA FRABASE version 1.0: An engine with a database to search for the three dimensional fragments within RNA structures. *Nucleic Acids Res. 36*, D386–D391.
53. Christiansen, M. E., and Znosko, B. M. (2009) Thermodynamic characterization of tandem mismatches found in naturally occurring RNA. *Nucleic Acids Res. 37*, 4696–4706.