Solution Structure of an RNA Internal Loop with Three Consecutive Sheared GA Pairs[†,‡]

Gang Chen,[§] Brent M. Znosko,[§,||] Scott D. Kennedy,[⊥] Thomas R. Krugh,[§] and Douglas H. Turner*[,§,#]

Department of Chemistry, University of Rochester, Rochester, New York 14627, Department of Biochemistry and Biophysics, and Center for Human Genetics and Molecular Pediatric Disease and Department of Pediatrics, School of Medicine and Dentistry, University of Rochester, Rochester, New York 14642

Received September 7, 2004; Revised Manuscript Received November 23, 2004

ABSTRACT: Internal loops in RNA are important for folding and function. Many folding motifs are internal loops containing GA base pairs, which are usually thermodynamically stabilizing, i.e., contribute favorable free energy to folding. Understanding the sequence dependence of folding stability and structure in terms of molecular interactions, such as hydrogen bonding and base stacking, will provide a foundation for predicting stability and structure. Here, we report the NMR structure of the oligonucleotide duplex, $^{5'GGUGGAGGCU3'}_{3'PCCGAAGCCG5'}$ (P = purine), containing an unusually stable and relatively abundant internal loop, $^{5'GGA3'}_{3'AAG5'}$. This loop contains three consecutive sheared GA pairs (trans Hoogsteen/Sugar edge AG) with separate stacks of three G's and three A's in a row. The thermodynamic consequences of various nucleotide substitutions are also reported. Significant destabilization of ∼2 kcal/mol at 37 °C is found for substitution of the middle GA with AA to form $^{5'GAA3'}_{3'AAG5'}$. This destabilization correlates with a unique base stacking and hydrogen-bonding network within the $^{5'GGA3'}_{3'AAG5'}$ loop. Interestingly, the motifs, $^{5'UG3'}_{3'GA5'}$ and $^{5'UG3'}_{3'AA5'}$, have stability similar to $^{5'CG3'}_{3'GA5'}$ even though UG and UA pairs are usually less stable than CG pairs. Consecutive sheared GA pairs in the $^{5'GGA3'}_{3'AAG5'}$ loop are preorganized for potential tertiary interactions and ligand binding.

RNA can fold into different secondary and tertiary structures for various functions. The increasing number of NMR and crystal structures of RNA and RNA−protein complexes have revealed new RNA-folding motifs (*1*−*3*). Many such motifs are internal loops that contain noncanonical base pairs such as GA and UU, which are usually thermodynamically stabilizing, i.e., contribute favorable free energy to folding (*4*). Predicting the sequence dependence of folding stability and structure for such loops is at an early stage, however (*5*−*8*). Understanding this sequence dependence in terms of detailed molecular recognition interactions, such as hydrogen bonding and base stacking, will provide a foundation for structural genomics of RNA. Here, we report the NMR structure of an oligonucleotide containing an unusually stable internal loop, $^{5'GGA3'}_{3'AAG5'}$, which is found in the signal recognition particle (SRP)[1] (*9*) and small subunit (SSU) and large subunit (LSU) ribosomal RNAs (*10*−*12*). The thermodynamic consequences of various nucleotide substitutions are also reported. The oligonucleotide sequence studied mimics an internal loop found in the SRP of the archaebacterium, *Thermococcus celer* (Figure 1) (*9*). It forms part of the Alu domain, which homology modeling (*13*, *14*) suggests may form extensive tertiary contacts that allow it to substitute for the SRP9/14 protein found in eukaryotic SRPs (*15*−*17*). The Alu domain is involved in the elongation arrest of nascent peptide chains once the S domain binds the signal sequence of a protein. Elongation arrest is thought to allow efficient targeting by providing a time window for the translocation of signal sequence (*15*, *17*). NMR spectra of the $^{5'UGGAG3'}_{3'GAAGC5'}$ internal loop indicate that it is preorganized, which may facilitate formation of tertiary contacts important for function.

MATERIALS AND METHODS

Oligoribonucleotide Synthesis and Purification. Oligoribonucleotides were synthesized on an Applied Biosystems 392 DNA/RNA synthesizer using the phosphoramidite method (*18*, *19*). CPG support and phosphoramidites were acquired from Glen Research or ChemGenes. Base-protecting groups and CPG support were removed by incubation in 3:1 (v/v) ammonia/ethanol at 55 °C overnight (*20*). A disposable filter column was used to separate the oligonucleotides from solid support. Removal of the silyl-protecting groups on the 2′ hydroxyls was achieved by incubation in 9:1 (v/v) TEA·

[†] This work was supported by NIH Grant GM22939 (to D.H.T.).
[‡] Protein Data Bank entry 1XV0.
[§] Department of Chemistry.
[||] Current address: Department of Chemistry, Saint Louis University, St. Louis, MO 63103.
[⊥] Department of Biochemistry and Biophysics.
[#] Center for Human Genetics and Molecular Pediatric Disease and Department of Pediatrics.
* To whom correspondence should be addressed. Telephone: (585) 275-3207. Fax: (585) 276-0205. E-mail: turner@chem.rochester.edu.

[1] Abbreviations: C_T, total concentration of all strands of oligonucleotides in solution; LSU, large subunit; N, any nucleotide including A, C, G, or U; P, purine; R, any G or A; SRP, signal recognition particle; SSU, small subunit; T_M, melting temperature in kelvin; T_m, melting temperature in degrees Celsius.

10.1021/bi048079y CCC: $30.25
Published on Web 02/08/2005

FIGURE 1: (Top) Secondary structure of part of helix 5 of *T. celer* SRP RNA (*9*), where the nucleotides in bold are studied by NMR here. (Bottom) Duplexes used for NMR structure determination (with numbering) and for thermodynamic studies. The value to the right of each duplex is the free-energy increment in kcal/mol at 37 °C for formation of the underlined internal loop in 1 M NaCl.

3HF (triethylamine trihydrofluoride)/DMF (*N*,*N*-dimethyl formamide) at 55 °C for ~2 h, followed by 1-butanol precipitation. The sample was lyophilized before redissolving in 5 mM ammonium bicarbonate at pH 7. The solution was loaded onto a Waters Sep-Pak C18 chromatography column to remove excess salts. The oligonucleotide was purified by TLC on a preparative Baker Si500F silica gel plate (20 × 20 cm, 500 μm thick), with the running solution of 55:35:10 (v/v/v) 1-propanol/ammonia/water. For G-rich sequences, the running solution was 50:40:10 (v/v/v) 1-propanol/ammonia/water. The product was identified by UV shadowing and scraped from the plate. RNA was extracted from the silica with distilled water. The Sep-Pak procedure was repeated to desalt the sample. The mass of all oligonucleotides was verified by ESI MS with a Hewlett−Packard 1100 LC/MS Chemstation. Purities were checked by reverse-phase HPLC or analytical TLC on a Baker Si500F silica gel plate (250 μm thick), and all were greater than 95% pure.

UV Melting Experiments and Thermodynamics. Concentrations of single-stranded oligonucleotides were calculated from the absorbance at 280 nm at 80 °C and extinction coefficients predicted from those of dinucleotide monophosphates and nucleosides (*21*). Purine (P) was assumed to be the same as adenine for approximation of extinction coefficients. Small mixing errors for nonself-complementary duplexes do not affect thermodynamic measurements appreciably (*22*). Oligonucleotides were lyophilized and redissolved in 1.0 M NaCl, 20 mM sodium cacodylate, and 0.5 mM disodium EDTA at pH 7.0. Curves of absorbance at 280 nm versus temperature were acquired using a heating rate of 1 °C/min with a Beckman Coulter DU640C spectrophotometer having a high performance temperature controller cooled with flowing water.

Melting curves were fit to a two-state model with MeltWin, assuming linear sloping baselines and temperature-independent $\Delta H°$ and $\Delta S°$ (*23*−*25*). Additionally, the temperature at which half the strands are in duplex, T_M, at total strand concentration, C_T, was used to calculate thermodynamic parameters according to (*26*)

$$T_M^{-1} = (R/\Delta H°)\ln(C_T/4) + (\Delta S°/\Delta H°) \qquad (1)$$

The equation $\Delta G°_{37} = \Delta H° - (310.15)\Delta S°$ was used to calculate the free energy change at 37 °C (310.15 K).

NMR Sample Preparation. Sample preparation was similar to Znosko et al. (*27*) with minor modification. The 300 μL sample was dialyzed overnight against 1 L of distilled water in a Gibco Life Technologies microdialysis system with a 1000 MW cutoff Spectro−Por #7 dialysis membrane and a Rainin Dynamax peristaltic pump. The sample was then lyophilized and redissolved in 80 mM NaCl, 10 mM sodium phosphate, and 0.5 mM Na_2EDTA at pH = 5.9. Total volume was 300 μL with 90:10 (v:v) H_2O/D_2O for exchangeable proton spectra. D_2O exchange was performed with lyophilization 3 times from 99.96% D_2O and finally dissolved in 300 μL of 99.996% D_2O (Cambridge Isotope Laboratories). The total duplex concentration was ~2 mM.

NMR Spectroscopy. Unless otherwise noted, all exchangeable and nonexchangeable proton spectra were acquired on a Varian Inova 500 MHz spectrometer. One-dimensional imino proton spectra were acquired with an S pulse sequence with a sweep width of 12 kHz and temperatures ranging from 0 to 55 °C. SNOESY spectra were recorded with a 150 ms mixing time at 0, 5, 30, and 40 °C. The spectra were acquired with 2k complex points and a spectral width of 10 kHz. A total of 300 FIDs were accumulated with 72 scans per FID and recycle delay of 1.2 s. The FELIX (2000) software package (Molecular Simulations Inc.) was used to process 2D spectra. Proton spectra were referenced to H_2O or HDO at a known temperature-dependent chemical shift relative to 3-(trimethylsilyl)tetradeutero sodium propionate (TSP).

NOESY spectra of samples in D_2O were acquired at 30 °C with 100, 200, 300, 400, and 600 ms mixing times. The spectra were acquired with 2k complex points and a spectral width of 4200 Hz. A total of 360 FIDs were acquired with 32 scans per FID and recycle delay of 2.8 s. TOCSY spectra were acquired at 30 °C with 8, 20, and 40 ms mixing times. The spectra were acquired with a spectral width of 4000 Hz with 2k complex points or with a spectral width of 1700 Hz with 1k complex points. A total of 256−512 FIDs were acquired with 24−64 scans per FID. The natural abundance ^1H-^{13}C HMQC spectrum was acquired with a 5000 Hz spectral width for proton and 15 000 Hz spectral width for carbon. The HMQC spectrum consisted of 60 increments each having 1k complex points with 512 scans per FID. The natural abundance ^1H-^{13}C HSQC spectrum was acquired with a 6000 Hz spectral width for proton and 9000 Hz spectral width for carbon on a Varian Inova 600 MHz spectrometer. The HSQC spectrum consisted of 64 increments each having 1k complex points with 700 scans per FID. The inversion recovery method was used to measure T_1 relaxation times for base protons. A 1D ^{31}P spectrum of the D_2O sample was acquired at 30 °C on a Bruker Avance 400 MHz spectrometer. The ^1H-^{31}P HETCOR spectrum was acquired with a 1400 Hz spectral width for proton and 2000 Hz for ^{31}P. The ^{31}P chemical shifts were referenced to the phosphate buffer at 0 ppm.

Restraint Generation. NOESY cross-peak volumes were integrated with FELIX (2000). Distance restraints were generated from 100 and 200 ms mixing time NOE volumes at 30 °C with the average volume of all pyrimidine (C and U) H5−H6 cross peaks as a reference (2.45 Å). All other distances were calculated from the two-spin approximation, scaling by $1/r^6$. The upper and lower bounds were set to

Table 1: Thermodynamic Parameters for Duplex Formation in 1 M NaCl at pH 7 unless Otherwise Noted

	T_M^{-1} versus ln(C_T/4) plots (eq 1)				average of melt curve fits			
sequences	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$-\Delta G°_{37}$ (kcal/mol)	T_m[a] (°C)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	$-\Delta G°_{37}$ (kcal/mol)	T_m[a] (°C)
GGU**GGA**GGCU[b]	94.3 ± 8.2	261.2 ± 24.5	13.26 ± 0.57	60.8	94.5 ± 2.4	261.9 ± 7.2	13.27 ± 0.25	60.8
PCCG**AAG**CCG								
GGU**GGA**GGCU	90.5 ± 2.1	248.7 ± 6.3	13.37 ± 0.14	62.4	95.0 ± 2.3	262.2 ± 7.2	13.66 ± 0.13	62.2
PCCG**AAG**CCG								
GGU**GGA**GGCU	96.3 ± 4.6	271.9 ± 14.2	12.02 ± 0.25	55.7	93.8 ± 5.7	264.0 ± 17.5	11.89 ± 0.26	55.8
PCCG**AAG**UCG								
GGU**GGA**GGCU	92.7 ± 2.2	258.0 ± 6.6	12.64 ± 0.13	58.9	93.2 ± 3.9	259.7 ± 12.0	12.67 ± 0.23	58.9
PCCA**AAG**CCG								
GGC**GGA**GGCU	88.9 ± 6.2	242.3 ± 18.5	13.75 ± 0.48	64.4	87.1 ± 6.9	236.9 ± 20.3	13.63 ± 0.64	64.5
PCCG**AAG**CCG								
GGC**GGA**GGCU	89.9 ± 6.2	249.9 ± 18.8	12.40 ± 0.40	58.7	88.1 ± 5.4	244.3 ± 16.3	12.31 ± 0.35	58.8
PCCG**AAG**UCG								
GGU**GAA**GGCU	85.1 ± 5.8	239.4 ± 17.8	10.88 ± 0.30	53.7	95.7 ± 6.1	271.8 ± 19.0	11.38 ± 0.30	53.6
PCCA**AAG**CCG								
GGU**GAA**GGCU	84.2 ± 6.1	235.7 ± 18.6	11.12 ± 0.32	54.9	86.5 ± 5.4	242.8 ± 16.9	11.23 ± 0.25	54.8
PCCG**AAG**CCG								
GGC**GAA**GGCU	81.2 ± 7.0	223.4 ± 21.2	11.92 ± 0.47	59.1	77.8 ± 5.8	213.1 ± 17.7	11.76 ± 0.36	59.3
PCCG**AAG**CCG								
GGU**GAA**GGCU	81.2 ± 5.2	230.8 ± 16.2	9.65 ± 0.20	49.4	79.1 ± 6.8	224.2 ± 21.2	9.59 ± 0.22	49.5
PCCG**AAG**UCG								
GGC**GAA**GGCU	72.8 ± 5.3	202.6 ± 16.6	9.97 ± 0.26	52.4	69.8 ± 5.1	193.2 ± 15.8	9.88 ± 0.22	52.6
PCCG**AAG**UCG								
GGUGGCU	75.6 ± 3.9	205.0 ± 11.7	12.05 ± 0.28	61.4	80.8 ± 2.3	220.5 ± 7.0	12.39 ± 0.17	61.3
PCCGCCG	(68.8)[c]	(182.3)[c]	(12.17)[c]	(65.2)[c]				

[a] At C_T = 0.1 mM. [b] Values listed on the bottom are measured in 150 mM KCl, 10 mM $MgCl_2$, and 20 mM sodium cacodylate at pH 7. [c] Parameters listed in parentheses are calculated from the nearest-neighbor model in refs *25*, *32*, *35*, *37*, and *38*.

±30% of each distance to allow for potential errors because of spin diffusion, spectrum noise, and distortions of phasing and baseline. Restraints of overlapped peaks were loosened or discarded. For some weak peaks and SNOESY peaks, loosened upper bounds were used. No restraints from H5′ or H5″ were used. A total of 197 interproton distance restraints (101 intranucleotide and 96 internucleotide) (Tables S1 and S2 in the Supporting Information), 15 hydrogen-bond restraints (limiting proton and hydrogen-bond acceptor distances within 1.8−2.5 Å) for 5 Watson−Crick GC pairs, and 102 dihedral angle restraints were included. No hydrogen-bond restraints were used within the loop and GU pair. Because ^{31}P chemical shifts are within ±0.7 ppm, which is typical for A-form RNA (Figure S1 in the Supporting Information), backbone dihedral restraints were used. Dihedral angles of residues in the Watson−Crick stems and GU pair were loosely restrained: α (0 ± 120°), β (180 ± 30°), γ (60 ± 30°), δ (85 ± 30°), ϵ (−140 ± 40°), ξ (0 ± 120°), and χ (−170 ± 40°). For loop residues (G4, G5, A6, G14, A15, and A16), only the δ and χ dihedral angles were restrained. Dihedral angle constraints for δ were obtained from TOCSY and DQF−COSY spectra. G5 was restrained to be the C2′-*endo* conformation with δ (160 ± 30°); A6, G14, and the two 3′-dangling residues, U10 and P20, were restrained to cover both the C2′-*endo* and C3′-*endo* conformations with δ (122.5 ± 67.5°). G4, A15, and A16 were restrained to be the C3′-*endo* conformation with δ (85 ± 30°). Glycosidic bond dihedral angles, χ's, were loosely restrained (−120 ± 90°) because there was no indication of a syn conformation, which would be indicated by a strong intraresidue H1′-H8/H6 cross peak comparable in intensity to the H5-H6 cross peaks.

Structure Calculation. NMR-restrained molecular dynamics and energy minimization was done with the Discover 98 package on a Silicon Graphics computer. An A-form like RNA starting structure was generated with the Biopolymer module of Insight II (2000). Partial charges of the purine 3′-dangling end are the same as previously reported (*28*). The AMBER 95 force field (*29*) was used with addition of flat-bottom restraint pseudopotentials, with force constants of 25 kcal/(mol Å^2) for NOE distance restraints and 50 kcal/(mol rad^2) for torsion-angle restraints and with a maximum force of 1000 kcal/mol. Group-based summation with an 18 Å cutoff was used for calculating van der Waals interactions. The cell-multipole method, with a distance-dependent dielectric constant (ϵ = 2r), was used for calculating electrostatic interactions. The progression of the structure simulation was (*27*, *30*) (1) covalent bond energy, NOE, and dihedral angle constraints scaled to 100%, van der Waals and electrostatic interactions scaled to 0 and 1%, respectively; (2) 500 steps of steepest descent energy minimization; (3) 4 ps of restrained molecular dynamics with 1 fs time steps at 1000 K; (4) 2 ps rMD at 900 K; (5) 2 ps rMD at 800 K; (6) increase van der Waals and electrostatic interactions to 33%; (7) 2 ps rMD at 700 K; (8) increase van der Waals and electrostatic interactions to 67%; (9) 2 ps rMD at 600 K; (10) increase van der Waals and electrostatic interactions to 100%; (11) 2 ps rMD at 500 K; (12) 2 ps rMD at 400 K; (13) 2 ps rMD at 300 K; and (14) 40 000 steps of conjugate-gradient energy minimization. Only one starting structure was used because the rMD at 1000 K effectively randomizes the structure. A total of 33 of 40 modeled structures were very similar and selected for analysis.

RESULTS

Thermodynamics. Thermodynamic parameters for several duplexes and internal loops at 1 M NaCl are listed in Tables 1 and 2, respectively. Throughout the paper, each top strand is written from 5′ to 3′ as going from left to right. Results for $\frac{\text{GGUGGAGGCU}}{\text{PCCGAAGCCG}}$ in 1 M NaCl and in 0.15 M KCl, 10 mM Mg^{2+} were the same within experimental error (Table 1).

Table 2: Thermodynamic Parameters for Internal Loop Formation in 1 M NaCl at pH 7[a]

sequence	$\Delta G^\circ_{37,loop}$ (kcal/mol)	ΔH°_{loop} (kcal/mol)	ΔS°_{loop} (eu)
GGU**GGA**GGCU PCCG**AAG**CCG	−2.62 ± 0.78 (−0.17)	−24.3 ± 12.4	−69.7 ± 37.5
GGU**GGA**GGCU[b] PCCG**AAG**UCG	−2.53 ± 0.65 (0.67)	−27.6 ± 10.7	−80.9 ± 32.7
GGU**GGA**GGCU[b] PCCA**AAG**CCG	−2.27 ± 0.59 (−0.17)	−23.9 ± 9.7	−69.5 ± 29.5
GGC**GGA**GGCU[b] PCCG**AAG**CCG	−2.20 ± 0.77 (−1.01)	−16.6 ± 11.5	−46.4 ± 34.9
GGC**GGA**GGCU[b] PCCG**AAG**UCG	−2.00 ± 0.77 (−0.17)	−18.9 ± 11.8	−54.5 ± 35.9
GGU**GAA**GGCU[b] PCCA**AAG**CCG	−0.51 ± 0.60 (1.05)	−16.3 ± 9.6	−50.9 ± 29.5
GGU**GAA**GGCU PCCG**AAG**CCG	−0.48 ± 0.57 (1.05)	−14.2 ± 11.1	−44.2 ± 34.0
GGC**GAA**GGCU[b] PCCG**AAG**CCG	−0.37 ± 0.76 (0.21)	−8.9 ± 11.9	−27.5 ± 36.4
GGU**GAA**GGCU[b] PCCG**AAG**UCG	−0.16 ± 0.63 (1.89)	−12.5 ± 11.0	−39.8 ± 33.6
GGC**GAA**GGCU[b] PCCG**AAG**UCG	0.43 ± 0.65 (1.05)	−1.8 ± 11.4	−7.2 ± 34.8

[a] Experimental errors for ΔG°_{37}, ΔH°, and ΔS° for the canonical stems are estimated as 4, 12, and 13.5%, respectively, according to ref *25*. There is less error in comparisons between these sequences because the stems are either identical or different by only one or two base pairs. Values in parentheses are predicted according to ref *5*. [b] Nearest-neighbor parameters (*25*, *32*) are used to estimate the difference of one or two base pairs compared to the measured core helix in Table 1.

Thermodynamic parameters for formation of the internal loops (Table 2) are calculated according to the following equation (*31*):

$$\Delta G^\circ_{37,loop} = \Delta G^\circ_{37(duplex\ with\ loop)} - \Delta G^\circ_{37(duplex\ without\ loop)} + \Delta G^\circ_{37(interrupted\ base\ stack)} \quad (2a)$$

For example,

$$\Delta G^\circ_{37}\left(\begin{smallmatrix}5'UGGAG3'\\3'GAAGC5'\end{smallmatrix}\right) = \Delta G^\circ_{37}\left(\begin{smallmatrix}5'GGUGGAGGCU3'\\3'PCCGAAGCCG5'\end{smallmatrix}\right) - \Delta G^\circ_{37}\left(\begin{smallmatrix}5'GGUGGCU3'\\3'PCCGCCG5'\end{smallmatrix}\right) + \Delta G^\circ_{37}\left(\begin{smallmatrix}5'UG3'\\3'GC5'\end{smallmatrix}\right) \quad (2b)$$

Here, $\Delta G^\circ_{37}\left(\begin{smallmatrix}5'GGUGGAGGCU3'\\3'PCCGAAGCCG5'\end{smallmatrix}\right)$ is the measured value of the duplex containing the internal loop (Table 1), $\Delta G^\circ_{37}\left(\begin{smallmatrix}5'GGUGGCU3'\\3'PCCGCCG5'\end{smallmatrix}\right)$ is the measured value of the duplex without the loop (Table 1), and $\Delta G^\circ_{37}\left(\begin{smallmatrix}5'UG3'\\3'GC5'\end{smallmatrix}\right)$ is the free energy increment for the nearest-neighbor base-stack interaction interrupted by the internal loop (*25*, *32*). Identical calculations can be done for ΔH°_{loop} and ΔS°_{loop}. All of the thermodynamic parameters used in this calculation are derived from T_M^{-1} versus $\ln(C_T/4)$ plots (eq 1). All of the ΔH° values from T_M^{-1} versus $\ln(C_T/4)$ plots and from the average of the fits of melting curves to two-state transitions agree within 11% (Table 1), suggesting that the two-state model is a good approximation for these transitions. Optical melting studies of the individual strands for $\begin{smallmatrix}GGUGGAGGCU\\PCCGAAGCCG\end{smallmatrix}$ in 1 M NaCl showed that the free energies for the formation of homoduplexes are at least 2 kcal/mol less favorable at 37 °C than for the heteroduplex (data not shown). Moreover, the NMR spectra for $\begin{smallmatrix}GGUGGAGGCU\\PCCGAAGCCG\end{smallmatrix}$ reveal the expected duplex structure, and the 1D imino proton spectra as a function of the temperature provide no indication of significant concentrations of intermediate states (see below and Figure 2).

Exchangeable Proton Assignments. NMR spectra were measured for the duplex, $\begin{smallmatrix}GGUGGAGGCU\\PCCGAAGCCG\end{smallmatrix}$, which mimics the sequence formed in helix 5 of *T. celer* SRP (Figure 1). The

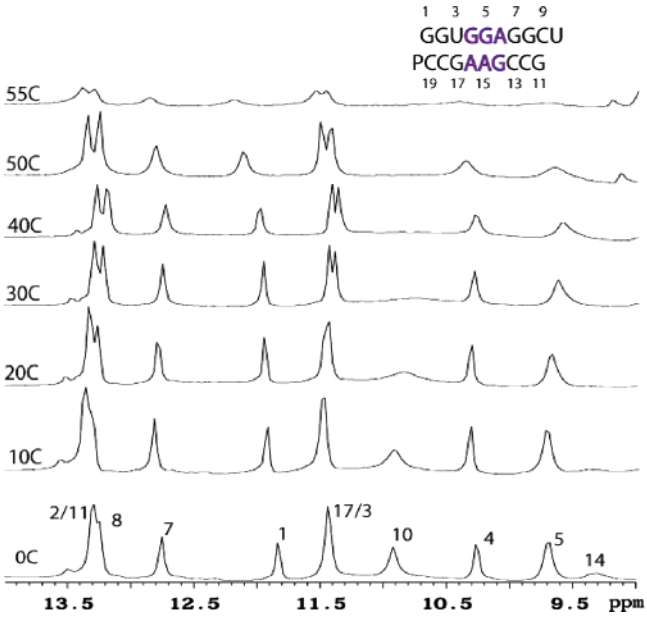

FIGURE 2: One-dimensional imino ^{1}H NMR spectra of $\begin{smallmatrix}GGUGGAGGCU\\PCCGAAGCCG\end{smallmatrix}$ at various temperatures. Imino proton resonances are assigned in the bottom spectrum.

1D imino proton spectra at variable temperatures (Figure 2) reveal that, with the possible exception of G14, the internal loop structure melts at about the same temperature as the canonical base pairs. The weak peak resonating at 13.5 ppm is probably due to a minor conformation involving G2 and/or G11 imino protons (Figure 2). Imino proton spectra of each individual strand show no peak at 13.5 ppm. The SNOESY spectrum (Figure S2) confirms the secondary structure and helps assign G imino, C amino, and C H5/H6 protons. G and A amino proton peaks are usually not observable because of intermediate exchange by rotation. Some unassigned peaks in the SNOESY spectrum are probably due to G or A amino protons in this uniquely

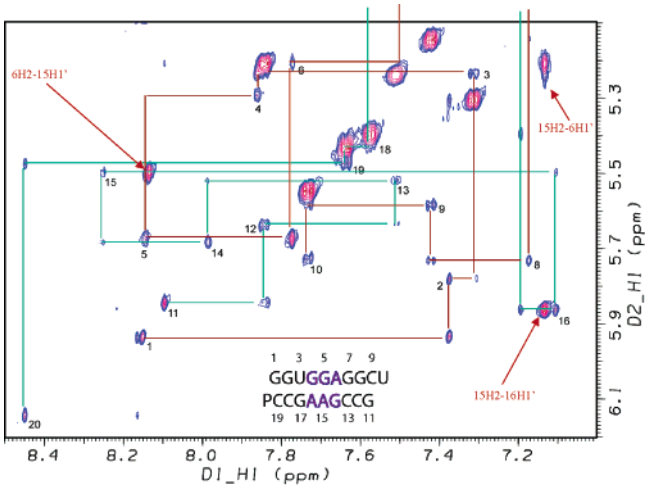

FIGURE 3: (H8/H6/H2)-(H1′/H5) region of the 400 ms mixing time NOESY spectrum of $\frac{\text{GGUGGAGGCU}}{\text{PCCGAAGCCG}}$ at 30 °C. Brown and green lines trace the NOESY walks for the top and bottom strands, respectively.

structured loop. U3H3 and G17H1 are overlapped at 5 °C but separate above 30 °C (Figure 2). G4H1 was assigned by the weak NOE to U3H3. G5H1 and G14H1 were assigned following the cross peaks G4H1-G5H1 and G5H1-G14H1. The imino proton of the 3′-dangling U10 is observable at 10.9 ppm, probably because of the slower exchange rate with water at pH 5.9 compared with the pH 6.5−6.8 used in previous NMR studies (*27*, *28*).

Nonexchangeable Proton Assignments. NMR resonances were assigned essentially as described previously (*27*, *28*, *33*, *34*). The (H8/H6/H2)-(H1′/H5) "NOESY walk" region of the 400 ms NOESY spectrum at 30 °C is shown in Figure 3. The 3′-dangling P20 with characteristic downfield chemical shifts of H8/H6/H2/H1′ helps initiate the "walk" (*28*, *35*). The G7H1′ and G17H1′ were shifted upfield to 4.21 and 4.75 ppm, respectively, and were confirmed by a natural abundance ^{1}H-^{13}C HSQC spectrum (Figure S3). Strong cross peaks resulting from H5-H6 (∼2.45 Å), H1′-H2′ (2.8−3.0 Å), and H1′-H3′ (3.5−4.0 Å) protons are important for assignments. Five characteristic H5-H6 cross peaks in the NOESY walk region and corresponding spin diffusion cross peaks to C amino protons and G imino protons in the SNOESY spectrum confirmed the formation of 5 Watson−Crick GC pairs within different nearest-neighbor environments. The H2′ and H3′ peaks were assigned from strong cross peaks to H1′ resonances and were reconfirmed by the strong to medium sequential cross peaks (*n*)H8/H6 − (*n* − 1)H2′/H3′. TOCSY (Figure S4 in the Supporting Information), DQF−COSY (data not shown), and ^{1}H-^{31}P HETCOR (data not shown) spectra were also used to facilitate sugar proton assignments. The assignment of adenine H2 protons, which usually relax slowly because they are relatively spatially isolated, was facilitated by measurement of base proton T_1 relaxation times (data not shown). A natural abundance ^{1}H-^{13}C HMQC spectrum (Figure S5 in the Supporting Information) confirmed the AH2 assignments. All of the assignments are listed in Table S3 in the Supporting Information.

Structural Features. The structure of the $\frac{\text{GGUGGAGGCU}}{\text{PCCGAAGCCG}}$ duplex was modeled from distance and dihedral angle restraints derived from the NMR spectra. The superposition

FIGURE 4: Superposition of 30 low-energy structures derived from restrained molecular dynamics. The kink in the sugar−phosphate backbone is between G14 and A15. The backbone between G5 and A6 is sometimes kinked but is more dynamic. A total of 3 of 33 selected structures were omitted because the backbones between G14 and A15 were not well-overlapped.

of 30 structures shown in Figure 4 demonstrates that the structure is well-defined. The average root-mean-square deviation of all selected 33 structures to the average structure for heavy atoms and all atoms are 0.66 ± 0.10 and 0.69 ± 0.10 Å, respectively. The average number of distance restraint violations is 4 ± 2. Only 6 of 33 selected structures have violations greater than 0.1 Å. Violations that were >0.1 Å are only for G14H8-G14H2′ (NMR distance of ∼2.7 Å), which is due to the G14 sugar pucker being C3′-*endo* in modeled structures, instead of being populated between C3′-*endo* and C2′-*endo* conformations as indicated by intermediate scalar coupling of ∼5 Hz for G14 H1′-H2′ (Figure S4 in the Supporting Information). No distance or dihedral-angle restraint violations were greater than 0.20 Å or 2°, respectively.

The structure of the $\frac{\text{UGGAG}}{\text{GAAGC}}$ loop contains a standard GU wobble pair (cis Watson−Crick/Watson−Crick) and three sheared GA pairs [trans Hoogsteen/Sugar edge AG (*36*)]. Within the three consecutive sheared GA pairs, intra- and cross-strand stacking are adopted, resulting in the stacking pattern in the loop: G4/G5/G14 and A6/A15/A16 (Figures 5 and 6). Detailed hydrogen bonds within the loop are shown in Figure 7 and Table 3. The 3′-dangling ends, U10 and P20, stack against the stem terminal base pairs.

Thermodynamic Stabilization by 3′-Dangling Ends Correlates with the Modeled NMR Structure. The 3′-dangling P and uridine stabilize the duplex by 1.3 and 1.1 kcal/mol, respectively, at 37 °C in 1 M NaCl buffer (*35*, *37*, *38*), which is favorable for NMR studies at 30 °C in NMR buffer (*27*, *28*, *39*, *40*). A preliminary TOCSY spectrum for the duplex without 3′-dangling ends showed more H5-H6 cross peaks than expected, suggesting dynamics between multiple con-

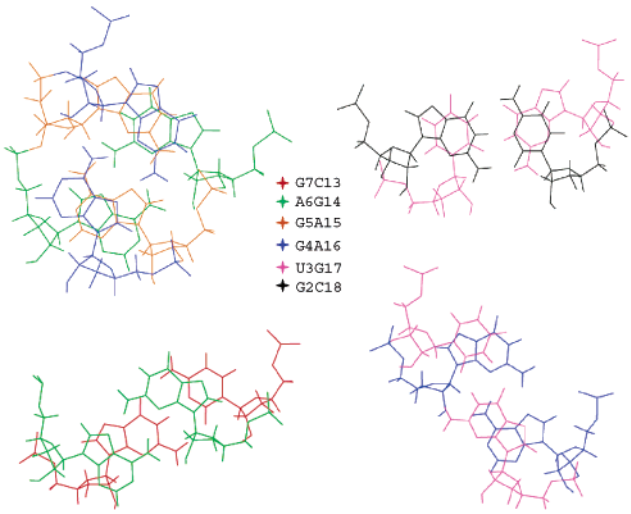

FIGURE 5: (Top left) Stacking pattern of three consecutive sheared GA pairs with G4A16 (blue), G5A15 (brown), and A6G14 (green). A6G14 is closest to the viewer. (Top right) Stacking of $\frac{GU}{CG}$ with U3G17 (pink) closer to the viewer and G2C18 in black. (Bottom left) Stacking of $\frac{CG}{GA}$ with G7C13 (red) closer to the viewer and A6G14 in green. (Bottom right) Stacking of $\frac{UG}{GA}$ with G4A16 closer to the viewer and U3G17 in pink.

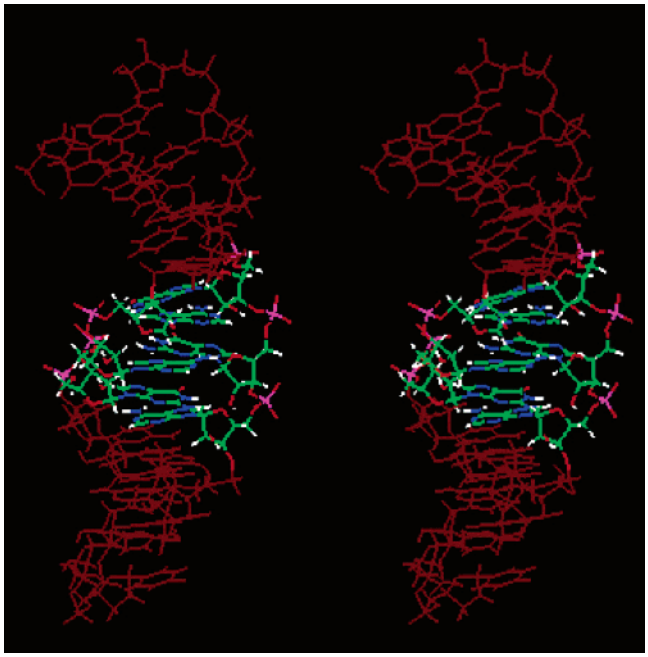

FIGURE 6: Stereoview of one of the modeled low-energy structures. GC and GU pairs and 3′-dangling end nucleotides are in dark red. The 3′-dangling ends U and P are at the bottom and top, respectively. The color code for loop residues is red for oxygen, blue for nitrogen, green for carbon, white for hydrogen, and pink for phosphorus.

formations. The U10 3′-dangling end stacked on C9 with favorable overlap of the U10 carbonyl group, O4, and C9 amino groups, which is similar to previous NMR (*27*, *28*, *37*) and X-ray structures (*37*, *41*−*43*). P20 is capping the stem terminal base pair G1C19 by both intra- and cross-strand stacking, which is similar to that observed for adenine 3′-dangling ends (*37*, *39*, *40*).

GU Wobble Pair Correlates with NMR Spectra. The wobble pair of U3G17 is confirmed by (1) medium cross peaks of G2H1-U3H3 and G2H1-G17H1 (Figure S2 in the Supporting Information), which are still seen at 30 °C and bring U3H3 and G17H1 close; (2) a weak cross-strand cross peak of G17H1-U3H2′ (Figure S2 in the Supporting Information); (3) U3H3 and G17H1 resonances, which are seen

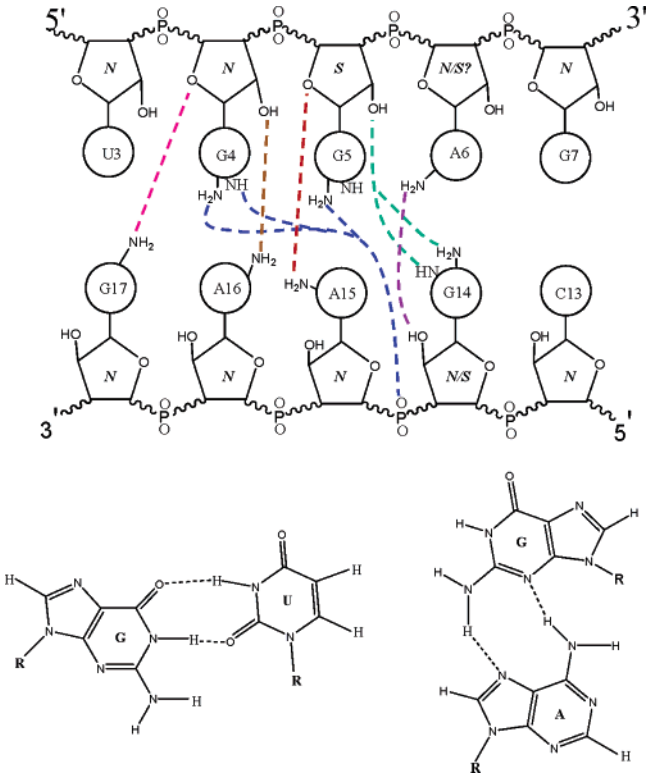

FIGURE 7: Scheme of potential hydrogen-bond network involving backbone groups within the loop (top). A distance shorter than 3.2 Å between hydrogen and the hydrogen-bond acceptor was used as the only criterion for a hydrogen bond (*54*). Sugars are labeled N and S for C3′-*endo* and C2′-*endo* conformations, respectively. The sugar pucker for A6 appears to be dynamic based on the shapes of the A6H1′ cross peaks in NOESY spectra. Base−base hydrogen bonds for standard GU wobble and sheared GA pairs are shown on the bottom.

Table 3: Potential Hydrogen Bonds and Distances within the Internal Loop, $\frac{GGA}{AAG}$

hydrogen-bond acceptor	hydrogen-bond donor	fraction occurrence[a]	average of hydrogen-bond distance (Å)[b]
U3 carbonyl	G17 imino	1.00	1.86 ± 0.03
G17 carbonyl	U3 imino	1.00	2.01 ± 0.02
G4 O4′	G17 amino	0.97	2.57 ± 0.37
G4 2′ hydroxyl	A16 amino	0.97	1.93 ± 0.24
G4 N3	A16 amino	0.97	2.16 ± 0.05
A15 OP	G4 imino	0.88	2.28 ± 0.27
A15 OP	G4 amino	0.88	2.80 ± 0.37
A16 N7	G4 amino	0.97	2.19 ± 0.30
G5 O4′	A15 amino	0.58	2.71 ± 0.29
G5 2′ hydroxyl	G14 imino	0.73	2.39 ± 0.41
G5 2′ hydroxyl	G14 amino	0.94	2.35 ± 0.45
G5 N3	A15 amino	0.52	2.74 ± 0.33
A15 OP	G5 amino	0.88	1.86 ± 0.08
A15 N7	G5 amino	0.97	2.07 ± 0.29
A6 N7	G14 amino	0.94	1.96 ± 0.03
G14 2′ hydroxyl	A6 amino	0.73	2.47 ± 0.41
G14 N3	A6 amino	0.48	2.23 ± 0.06

[a] Fraction of 33 structures that had hydrogen and the hydrogen-bond acceptor within 3.2 Å, thus indicating a hydrogen bond (*54*). [b] Distance from hydrogen atom to acceptor atom.

up to 50 °C in 1D imino proton spectra (Figure 2), suggesting that they are both tightly hydrogen-bonded; (4) the SNOESY spectrum at 30 °C (data not shown), suggesting that both U3H3 and G17H1 show an NOE to a pair of unassigned amino protons, which we speculate as G17 amino protons. On the basis of the relative intensities of these pairs of cross peaks at 30 °C, G17H1 was assigned further upfield than U3H3; and (5) The U3H3-G17H1 cross peak, which is observable at 40 °C (Figure S6 in the Supporting Information). One of the G17 amino protons is potentially hydrogen-bonded to the G4O4′ as shown in modeled structures (Figure 7 and Table 3), which is in agreement with the observed pair of chemical shifts (6.37 and 5.42 ppm) for G17 amino protons (Figure S2 in the Supporting Information).

Three Consecutive Sheared GA Pairs Correlate with NMR Spectra. Sheared pairs of G4A16 and G14A6 are confirmed by (1) medium cross-strand cross peaks of G7H1-G14H1′ and G17H1-G4H1′ (Figure S2 in the Supporting Information); (2) medium cross peaks of A16H2-G17H1′ and A6H2-G7H1′ (data not shown), which bring the Watson−Crick edges of A16 and A6 to the minor groove; and (3) upfield chemical shifts of G7H1′ and G17H1′ to 4.21 and 4.75 ppm (Figure S3 in the Supporting Information), respectively. An upfield chemical shift for the H1′ of the residue 3′ to a sheared purine−purine pair is common (*27*, *28*, *44*−*51*). This upfield chemical shift may be explained by empirical modeling of the ring current of adenosine above the H1′ proton of the 3′ residue, which is in agreement with quantum mechanical calculations and molecular modeling (*52*, *53*).

Cross-strand stacking between two sheared GA pairs of G5A15 and A6G14 is confirmed by the medium strength cross peaks, A15H2-A6H1′ and A6H2-A15H1′ (Figure 3), and G5H1-G14H2′. Several weak cross peaks of A15H2-A6H8, A15H2-A16H2 (data not shown), and G5H1-G14H1 (Figure S2 in the Supporting Information) also indicate cross-strand stacking of the adjacent sheared GA pairs. G4H2′ is above the aromatic ring of G5, which correlates with its unusual upfield chemical shift of 3.61 ppm. A15H2 has a relatively upfield chemical shift of 7.13 ppm, which also correlates with the stacking pattern shown in Figure 5 because A15H2 experiences ring currents from A16 and A6 bases.

Interestingly, at various temperatures, the G4H1 peak is always sharper than that of G14H1 (Figure 2), which is consistent with the modeled structure: more extensive stacking of U3/G4/G5 than C13/G14/G5 (Figure 5) and the stronger "cross-step" hydrogen bond expected between G4H1 and a nonbridging phosphate oxygen between G14 and A15 as compared to that between G14H1 and the G5 2′-hydroxyl oxygen (*54*) (Figure 7 and Table 3). The dynamic character of G14 is also shown by the intermediate scalar coupling of H1′-H2′ (∼5 Hz) in TOCSY (Figure S4 in the Supporting Information) and DQF−COSY (data not shown) spectra, which indicate that the sugar pucker of G14 is populated in both C2′-*endo* and C3′-*endo* conformations. Broad doublet peaks for A6H1′ in NOESY and a missing cross peak for G5H3′-A6P in a HETCOR spectrum (data not shown) also indicate that the G14A6 region is dynamic on the NMR time scale.

DISCUSSION

The constantly expanding databases of RNA sequences contain a huge amount of information. Understanding relationships between sequence, energetics, structure, and function can facilitate rapid extraction of this information. The internal loop is a common RNA motif where such relationships are not fully understood. The results of this study provide insights into relationships between sequence, energetics, and structure for an unusually stable 3 × 3 internal loop, $\frac{GGA}{AAG}$, which is relatively abundant in known secondary structures and likely important for function (*5*, *9*, *12*−*14*, *55*−*57*).

Molecular Recognition and Thermodynamics. The comparison between measured and predicted (in parentheses) values of $\Delta G^{\circ}_{37,loop}$ in Table 2 shows that $\frac{GGA}{AAG}$ and $\frac{GAA}{AAG}$ internal loops are more stable than expected. The NMR structure presented here is for the natural loop, $\frac{UGGAG}{GAAGC}$. Surprisingly, replacement of the U with C destabilizes this loop by 0.4 kcal/mol at 37 °C (Figure 1 and Table 2). This compares with a predicted stabilization of ∼0.8 kcal/mol, assuming a GU pair closing a loop is equivalent to an AU pair (*4*, *5*, *8*). This correlates with more extensive stacking of $\frac{UG}{GA}$ than $\frac{CG}{GA}$ (Figure 5) and a hydrogen bond between the G17 amino group and G4O4′ (Figure 7 and Table 3) in the NMR structure. Enhanced stability for a $\frac{UG}{GA}$ nearest-neighbor interaction is also suggested by the observation that replacement of the internal loop closing C in $\frac{GGUGGAGGCU}{PCCGAAGCCG}$ with a closing U destabilizes the loop by only 0.1 kcal/mol, rather than the predicted ∼0.8 kcal/mol (*4*, *5*, *8*). The subtle difference between the two closing base pairs for the $\frac{GGA}{AAG}$ loop might be due to the sequence asymmetry of the internal loop, e.g., different sugar pucker for G4 and G14 (Figure 7) and/or the different canonical stems in the NMR construct, $\frac{GGUGGAGGCU}{PCCGAAGCCG}$, i.e., different stacking interactions for $\frac{GU}{CG}$ and the CG replacement $\frac{GC}{CG}$ compared with $\frac{CC}{GG}$ and the UG replacement $\frac{CU}{GG}$ (*24*, *25*, *32*, *58*). As shown in Figure 5, the $\frac{GUGGA}{CGAAG}$ region is characterized by an extensive overlap of bases. In sequence symmetric 2 × 2 loops, which have two identical nearest-neighbor interactions involving closing canonical pairs with sheared GA pairs, the $\frac{UG}{GA}$ motif is also somewhat more stable than expected. In particular, $\frac{UGAG}{GAGU}$ (0.1 kcal/mol) is more stable than $\frac{UGAA}{AAGU}$ (0.7 kcal/mol) (*4*, *8*, *59*) but less stable than $\frac{CGAG}{GAGC}$ (−0.7 kcal/mol) (*4*, *8*, *45*). Evidently, the $\frac{UG}{GA}$ motif is unusually stable, which correlates with its frequent occurrence in ribosomal RNA (*59*, *60*).

Interestingly, an A^+C wobble pair (cis Watson−Crick/Watson−Crick), with AN1 protonated is isosteric with a GU wobble pair, and geometric compatibility has previously been reported between an A^+C wobble pair and sheared GA pair in a 3 × 3 internal loop, $\frac{CCGAG}{GAAGC}$ (*50*, *51*). A similar stacking pattern is seen for an AC wobble pair adjacent to a sheared AA pair (shown in bold) within the J4/5 loop, $\frac{GCAAC}{CAAAG}$, of a group I intron in complex with both the 5′ and 3′ exons (*61*).

Replacement of the loop middle G with A in $\frac{GGUGGAGGCU}{PCCGAAGCCG}$ and several variants destabilizes the loop by ∼2 kcal/mol at 37 °C (Figure 1 and Table 2), which is a significantly larger effect than predicted by the bonus parameter of −1.2 kcal/mol for a middle GA pair adjacent to at least one non-

pyrimidine-pyrimidine pair and ~ −1 kcal/mol for a loop-terminal GA pair (*4*, *5*). Extra stabilization by a GA middle pair was also found in other 3 × 3 loops with three consecutive GA pairs (*5*). This dramatic stabilization correlates with favorable stacking and hydrogen bonding within this internal loop (Figures 5 and 7 and Table 3). In particular, a central AA pair is not expected to form the many hydrogen bonds shown for the central GA pair in Figure 7.

As indicated in Figure 7, the G4 sugar adopts a predominant C3′-*endo* (N) sugar pucker, while G5 adopts a predominant C2′-*endo* (S) sugar pucker ($J_{H1'-H2'}$ ~ 8 Hz). The G14 sugar populates both sugar puckers as indicated by H1′-H2′ coupling of ~5 Hz. The base-backbone hydrogen bonds in the NMR structure of the loop $\frac{UGGAG}{GAAGC}$ include G4 (2′-hydroxyl)-A16 (amino), G5 (2′-hydroxyl)-G14 (amino/imino), G5 (amino)-A15 (nonbridging oxygen), and A6 (amino)-G14 (2′-hydroxyl) (Figure 7 and Table 3). These hydrogen-bonding patterns are similar to previously reported guanosine sugar-pucker-dependent hydrogen-bonding patterns of tandem sheared GA pairs (*50*). G imino and amino groups are close in space; thus, they could potentially share the same hydrogen-bond acceptors from the opposite backbone, perhaps mediated by water. Potential hydrogen bonds of G5O4′-A15 amino and a "cross-step" hydrogen bond of G4 amino/imino-A15 nonbridging oxygen are also observed in this structure. The formation of a "cross-step" hydrogen bond possibly stabilizes the backbone kink between G14 and A15. Similar hydrogen bonding was observed in a crystal structure of a related loop, $\frac{CGGAAG}{GAAGGC}$ (*62*).

Thermodynamic studies of 2 × 2 loops provide the following stability increments at 37 °C (*32*, *63*): $\frac{GAAG}{CAAC}$, 0.5 kcal/mol; $\frac{GAAG}{CAGC}$, −0.2 kcal/mol; $\frac{GGAG}{CAAC}$, −0.3 kcal/mol; and $\frac{GGAG}{CAGC}$, −1.7 kcal/mol. These values also suggest the importance of hydrogen bonding for the stabilities of tandem GA pairs (*63*). On the basis of isostericity matrixes (*36*), a sheared AA pair is expected to stack similar to a GA pair, although a GA pair is expected to have more hydrogen bonds. Consistent with these expectations, previous structural and thermodynamic studies showed that the destabilizing 2 × 2 loop $\frac{CAAG}{GAAC}$ (1.2 kcal/mol) with two sheared AA pairs has similar stacking geometries but fewer hydrogen bonds than $\frac{CGAG}{GAGC}$ (−0.7 kcal/mol) with two sheared GA pairs (*4*, *27*, *45*, *64*). The large effects associated with changing a GA to AA pair in 2 × 2 and 3 × 3 loops provide benchmarks for testing approaches such as free-energy perturbation and molecular dynamics (*65*, *66*), which attempt to make predictions of the sequence dependence of stability and structure. Understanding the interactions determining the sequence dependence of stability can facilitate prediction of secondary and three-dimensional structures from sequence (*32*).

Dynamics. In the $\frac{UGGAG}{GAAGC}$ loop, $\frac{UG}{GA}$ is more rigid than $\frac{CG}{GA}$ because of the stacking and hydrogen bonding discussed above. Significantly lower occurrence of hydrogen bonds of G (N3)-A (amino) (Figure 7) for G14A6 and G5A15 reflect dynamics as compared to G4A16 (Table 3). The missing cross peak of G5H3′-A6P suggests dynamics of the backbone, which is similar to those observed previously in the loop $\frac{CCGAG}{GAAGC}$ (*50*). Different orientations and detailed hydrogen-bonding patterns were also observed in different duplexes in the crystal structure of the loop, $\frac{CGAG}{GAGC}$, which had 2.5 duplexes in an asymmetric unit. This also suggests intrinsic dynamic features for sheared GA pairs in a $\frac{CG}{GA}$ motif in an internal loop (*64*). NMR studies of CGNRAG tetra-loop hairpins also reveal dynamics for the GA in a $\frac{CG}{GA}$ motif (*49*). Evidently, $\frac{CG}{GA}$ is an inherently dynamic motif.

Sequence- and Context-Dependent Structures of GA Pairs. The NMR structure of the $\frac{GGA}{AAG}$ internal loop provides insights useful for modeling of local three-dimensional structure. In general, formation of sheared GA pairs is accompanied by narrowing of the backbone 5′ of A and 3′ of G (*67*). This narrowing prohibits formation of a Watson−Crick pair 5′ of the A in a sheared GA pair. Thus, single GA pairs flanked by Watson−Crick pairs have never been observed in a sheared conformation but usually are observed in an imino conformation (cis Watson−Crick/Watson−Crick) (*55*, *68*) and less frequently in an A^+G cis Watson−Crick/Hoogsteen conformation (*69*, *70*). Previously, sheared GA's in size symmetric internal loops have only been observed as sequence symmetric pairs, i.e., as $\frac{GA}{AG}$ (*45*, *48*, *64*) or $\frac{GGAA}{AAGG}$ (*62*). The only previous structure with a 3 × 3 internal loop of consecutive GA pairs is in the crystal structure of *Deinococcus radiodurans* LSU rRNA (*57*). Helix 68 of that structure contains a $\frac{UGGAG}{AAAGC}$ loop that has only one sheared GA pair (shown in bold), with the rest of the loop bases not forming hydrogen-bonded pairs. The NMR structure of the $\frac{UGGAG}{GAAGC}$ loop reported here shows that it is possible to have an odd number of consecutive sheared GA pairs. It is possible that the UG closing base pair is a requirement for this motif. Alternatively, nonnearest-neighbor effects, the complex packing interactions, and/or difficulties in solving a large crystal structure may explain the structure reported for the $\frac{UGGAG}{AAAGC}$ loop in *D. radiodurans* (*57*).

While this is the first report of an odd number of consecutive sheared GA pairs in a size symmetric (i.e., 3 × 3) internal loop containing only GA pairs, this motif has been previously observed in crystals having the loop, $\frac{GGAU}{AAGU}$ (*55*, *56*), and in size asymmetric internal loops with a kink-turn motif (*71*, *72*). Biochemical and biophysical studies of the kink-turn motif, e.g., $\frac{G\mathbf{GGA}\ \ GC}{C\mathbf{AAG}AAGCG}$ (kt-7), suggest that the three consecutive sheared GA pairs (shown in bold) in the asymmetric internal loop are structured in solution even in the absence of Mg^{2+} and protein (*73*). This is consistent with the similar thermodynamic stabilities of $\frac{GGUGGAGGCU}{PCCGAAGCCG}$ in 1 M NaCl and in 0.15 M KCl with 10 mM Mg^{2+} (Table 1). Formation of the three sheared GA pair motif in $\frac{G\mathbf{GGA}\ \ GC}{C\mathbf{AAG}AAGCG}$ (kt-7) (*71*) instead of the imino hydrogen-bonded GA pairs usually found for the $\frac{GG}{CA}$ motif (*74*, *75*) indicates the favorable thermodynamics and function of three consecutive sheared GA pairs as a preorganized tertiary and protein-binding site. This is similar to the formation of a sheared GA pair in a $\frac{GG}{CA}$ motif in thermodynamically stable and functionally important GGNRAC tetraloops (*1*−*3*). The sequence $\frac{U\mathbf{GGA}A\ \ CG}{G\mathbf{AAG}GACG}$ (kt-58) with four potential GA pairs has only three sheared GA pairs (shown in bold) to facilitate the kink-turn (*71*). The sequences, $\frac{U\mathbf{GGA}\ \ C}{G\mathbf{AAG}GUAG}$ (kt-46) and $\frac{U\mathbf{GGA}\ \ G}{G\mathbf{AAG}UUUC}$ (kt-78), are well-conserved to have three GA pairs and a UG closing base pair in LSU rRNA (*12*, *71*, *72*). Interestingly, the asymmetric 3 × 4 internal loop

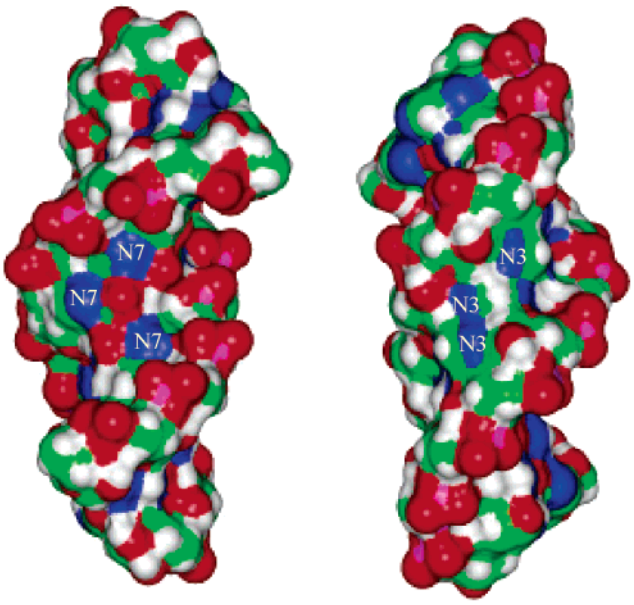

FIGURE 8: Solvent-accessible surface of major (left) and minor (right) grooves of the internal loop as defined by a 1.4 Å radius spherical probe. The color code is the same as that of loop residues in Figure 6. Note that the three consecutive GA pairs expose three N7-carbonyl group pairs from G's in the deep major groove and three N3-CH2 group pairs from A's in the shallow minor groove.

$\frac{\text{A GAC C}}{\text{UAGGAG}}$ in helix 5 of human SRP RNA, which is at a position similar to that of the loop $\frac{\text{UGGAG}}{\text{GAAGC}}$ in *T. celer* SRP (*9*), is predicted to have a kink-turn structure because it has the motif of tandem GA pairs in a size asymmetric loop (*71*). This loop is very important for the Alu domain folding as suggested by a low-resolution crystal structure and biochemical methods (*16*).

Functional Implications. Smooth N1-C2-N3 edges of three A's in the shallow minor groove of $\frac{\text{UGGAG}}{\text{GAAGC}}$ (Figure 8) provide a helix-packing site with an A-minor motif (*1*−*3*, *55*, *71*, *76*, *77*) and a potential protein-binding site (*16*, *71*, *78*). Interestingly, there is a 4 × 4 loop $\frac{\text{UGGAAG}}{\text{GAAGGC}}$ in helix 5 of *T. celer* SRP RNA (Figure 1) (*9*), which might also be involved in tertiary interactions facilitating the folding of the Alu domain.

A search of known RNA secondary structures containing 1079 3 × 3 internal loops (*5*, *10*−*12*) reveals that ∼20% of 3 × 3 loops have a middle GA pair and ∼4% have a $\frac{\text{GGA}}{\text{AAG}}$ motif. If 3 × 3 loop sequences were random, then only 12.5% would have a middle GA pair and ∼0.05% would be $\frac{\text{GGA}}{\text{AAG}}$. In eukaryal SSU rRNA, the $\frac{\text{GGA}}{\text{AAG}}$ motif is well-conserved in helix 44, with some variations to $\frac{\text{GGA}}{\text{AAA}}$, $\frac{\text{GA}}{\text{AG}}$, or $\frac{\text{GA}}{\text{AA}}$. The corresponding helix packing partner, helix 13 [according to the crystal structure of bacterial 30S rRNA of *T. thermophilus* (*55*)] is well-conserved as $\frac{\text{GGGAGC}}{\text{CCAUCG}}$ (*12*). Occurrences of the $\frac{\text{UGGAG}}{\text{GAAGC}}$ loop are found in helix 41a of 16S rRNA from bacteria including *Allochromatium vinosum*, *Haemophilus influenzae*, *Heliobacterium chorum*, *Legionella pneumophila*, *Leptonema illini*, *Listeria monocytogenes*, *Methylobacterium sp.*, *Methylococcus capsulatus*, *Petrotoga miotherma*, *Pseudomonas aeruginosa*, and *Thermomicrobium roseum* (*12*). Helix 41a is involved in tertiary interactions with helix 41 (*55*). In helix 42, a well-structured 4 × 4 loop, $\frac{\text{CGGAUU}}{\text{GAAGUA}}$ (conserved in SSU rRNA (*12*)), with three consecutive sheared GA pairs and a UU pair, is involved in tertiary contacts with helix 30, which has a conserved $\frac{\text{GGA}}{\text{CAU}}$ as shown in the crystal structure of 30S rRNA of *T. thermophilus* (*55*). In LSU rRNA, $\frac{\text{UGGAU}}{\text{AAAGU}}$ is also well-conserved in helix 2 (*12*) and involved in tertiary contacts with helix 26 (*56*, *57*, *77*). As shown in Figure 9, the superposition of the three consecutive sheared GA pairs in $\frac{\text{GUGGAUU}}{\text{UAAAGUA}}$ from helix 2 and $\frac{\text{G}\mathbf{\underline{GGA}}\text{ GC}}{\text{C}\mathbf{\underline{AAG}}\text{AAGCG}}$ (kt-7) of *H. marismortui* LSU rRNA to one of the NMR structures determined for $\frac{\text{UGGAG}}{\text{GAAGC}}$ shows that the $\frac{\text{GGA}}{\text{AAG}}$ motif is largely independent of sequence and structural contexts. Evidently, motifs with the potential for three consecutive sheared GA pairs are energetically stable and provide likely sites for tertiary interactions such as helix packing. Other $\frac{\text{GGA}}{\text{AAG}}$ motifs in the crystals mentioned above are similar.

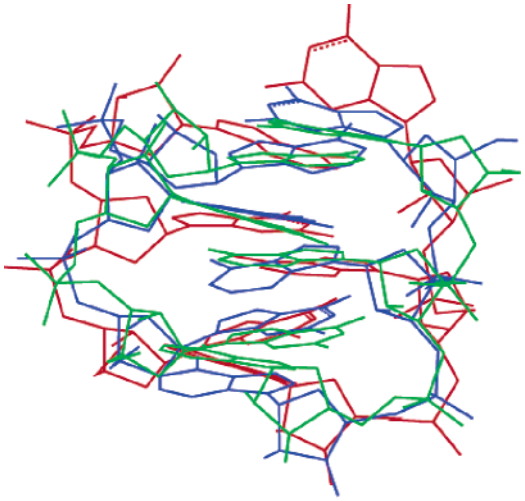

FIGURE 9: Superposition of three consecutive sheared GA pairs from one of the low-energy NMR structures of $\frac{\text{UGGAG}}{\text{GAAGC}}$ (green) and from $\frac{\text{GUGGAUU}}{\text{UAAAGUA}}$ (blue) in helix 2 and $\frac{\text{G}\mathbf{\underline{GGA}}\text{ GC}}{\text{C}\mathbf{\underline{AAG}}\text{AAGCG}}$ (kink-turn 7) (red) of the crystal structure of LSU rRNA (*56*).

Both GA imino and GU wobble pairs have their G amino groups projected into the minor groove, which is ideal for an A-minor motif interaction (*71*, *76*, *77*). A single GA pair at position −5 in the P1 helix of a *Candida albicans* group I intron is a recognition element for helix docking, and biochemical studies show that GU can substitute for this GA pair (*79*, *80*). In the *Tetrahymena thermophila* group I intron, there is a GU pair at position −5 of P1 and binding measurements suggest that it may also favor helix docking relative to substitution with a GC pair (*81*). The internal loop, J5b/5c, in the *C. albicans* intron has the sequence, $\frac{\text{UGGGAG}}{\text{GAAA C}}$, which can potentially form sheared GA pairs. Thus, this may provide another example of an A-minor motif tertiary interaction. Packing between an imino GA pair or a wobble GU pair and consecutive sheared GA pairs might be commonly used in the folding of many RNAs. Thus, improved understanding of the sequence dependence of their energetics may facilitate prediction of three-dimensional as well as secondary structure.

In the $\frac{\text{GGUGGAGGCU}}{\text{PCCGAAGCCG}}$ duplex studied here, three G's in the loop as well as U3 are stacked in a row, exposing carbonyl groups and GN7's with highly negative potentials in the deep major groove (Figures 6 and 8), thus also providing potential metal ion (*78*, *82*−*85*), protein (*71*), and small molecule (*86*,

87) binding sites. The lack of significant dynamics indicates that this binding site is preorganized and thus would favor tight binding.

ACKNOWLEDGMENT

We thank Dr. S. K. Sur for the help with 1D ^{31}P and ^{1}H-^{31}P HETCOR spectra.

SUPPORTING INFORMATION AVAILABLE

A table listing chemical-shift assignments, a table of NMR restraints, and figures of 1D ^{31}P, SNOESY, ^{1}H-^{13}C HSQC, TOCSY, and ^{1}H-^{13}C HMQC spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Batey, R. T., Rambo, R. P., and Doudna, J. A. (1999) Tertiary motifs in RNA structure and folding, *Angew. Chem., Int. Ed. 38*, 2327−2343.
2. Moore, P. B. (1999) Structural motifs in RNA, *Annu. Rev. Biochem. 68*, 287−300.
3. Leontis, N. B., and Westhof, E. (2003) Analysis of RNA motifs, *Curr. Opin. Struct. Biol. 13*, 300−308.
4. 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.
5. 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.
6. Schroeder, S. J., Burkard, M. E., and Turner, D. H. (1999) The energetics of small internal loops in RNA, *Biopolymers 52*, 157−167.
7. Schroeder, S. J., and Turner, D. H. (2000) Factors affecting the thermodynamic stability of small asymmetric internal loops in RNA, *Biochemistry 39*, 9257−9274.
8. Schroeder, S. J., and Turner, D. H. (2001) Thermodynamic stabilities of internal loops with GU closing pairs in RNA, *Biochemistry 40*, 11509−11517.
9. Kaine, B. P. (1990) Structure of the archaebacterial 7S RNA molecule, *Mol. Gen. Genet. 221*, 315−321.
10. Gutell, R. R., Gray, M. W., and Schnare, M. N. (1993) A compilation of large subunit (23S- and 23S-like) ribosomal RNA structures, *Nucleic Acids Res. 21*, 3055−3074.
11. Gutell, R. R. (1994) Collection of small subunit (16S- and 16S-like) ribosomal RNA structures, *Nucleic Acids Res. 22*, 3502−3507.
12. 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 Bioinformatics 3*, 2.
13. Eichler, J., and Moll, R. (2001) The signal recognition particle of Archaea, *Trends Microbiol. 9*, 130−136.
14. Zwieb, C., and Eichler, J. (2002) Getting on target: The archaeal signal recognition particle, *Archaea 1*, 27−34.
15. Keenan, R. J., Freymann, D. M., Stroud, R. M., and Walter, P. (2001) The signal recognition particle, *Annu. Rev. Biochem. 70*, 755−775.
16. Weichenrieder, O., Wild, K., Strub, K., and Cusack, S. (2000) Structure and assembly of the Alu domain of the mammalian signal recognition particle, *Nature 408*, 167−173.
17. Halic, M., Becker, T., Pool, M. R., Spahn, C. M. T., Grassucci, R. A., Frank, J., and Beckmann, R. (2004) Structure of the signal recognition particle interacting with the elongation-arrested ribosome, *Nature 427*, 808−814.
18. 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.
19. 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.
20. Stawinski, J., Stromberg, R., Thelin, M., and Westman, E. (1988) Evaluation of the use of the *tert*-butyldimethylsilyl group for 2′-protection in RNA: Synthesis via the H-phosphonate approach, *Nucleosides Nucleotides 7*, 779−782.
21. Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., pp 597, CRC Press, Cleveland, OH.
22. Peritz, A. E., Kierzek, R., Sugimoto, N., and Turner, D. H. (1991) Thermodynamic study of internal loops in oligoribonucleotides: Symmetrical loops are more stable than asymmetric loops, *Biochemistry 30*, 6428−6436.
23. Petersheim, M., and Turner, D. H. (1983) Base-stacking and base-pairing contributions to helix stability: Thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp, *Biochemistry 22*, 256−263.
24. 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.
25. 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.
26. 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.
27. Znosko, B. M., Burkard, M. E., Schroeder, S. J., Krugh, T. R., and Turner, D. H. (2002) Sheared A_{anti}•A_{anti} base pairs in a destabilizing 2 × 2 internal loop: The NMR structure of 5′(rGGCAAGCCU)$_2$, *Biochemistry 41*, 14969−14977.
28. Znosko, B. M., Burkard, M. E., Krugh, T. R., and Turner, D. H. (2002) Molecular recognition in purine-rich internal loops: Thermodynamic, structural, and dynamic consequences of purine for adenine substitutions in 5 ′(rGGCAAGCCU)$_2$, *Biochemistry 41*, 14978−14987.
29. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A 2nd generation force field for the simulation of proteins, nucleic acids, and organic molecules, *J. Am. Chem. Soc. 117*, 5179−5197.
30. 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.
31. Gralla, J., and Crothers, D. M. (1973) Free energy of imperfect nucleic acid helices. III. Small internal loops resulting from mismatches, *J. Mol. Biol. 78*, 301−319.
32. 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.
33. Varani, G., and Tinoco, I. (1991) RNA structure and NMR spectroscopy, *Q. Rev. Biophys. 24*, 479−532.
34. Varani, G., Aboulela, F., and Allain, F. H. T. (1996) NMR investigation of RNA structure, *Prog. Nucl. Magn. Reson. Spectrosc. 29*, 51−127.
35. Znosko, B. M., Kennedy, S. D., Wille, P. C., Krugh, T. R., and Turner, D. H. Structural features and thermodynamics of the J4/5 loop from the *Candida albicans* and *Candida dubliniensis* group I introns, *Biochemistry 43*, 15822−15837.
36. 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.
37. 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.
38. 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.
39. Zhang, H., Fountain, M. A., and Krugh, T. R. (2001) Structural characterization of a six-nucleotide RNA hairpin loop found in *Escherichia coli*, r(UUAAGU), *Biochemistry 40*, 9879−9886.

40. Yuan, Y. Q., Kerwood, D. J., Paoletti, A. C., Shubsda, M. F., and Borer, P. N. (2003) Stem of SL1 RNA in HIV-1: Structure and nucleocapsid protein binding for a 1 × 3 internal loop, *Biochemistry 42*, 5259−5269.
41. Westhof, E., Dumas, P., and Moras, D. (1988) Restrained refinement of 2 crystalline forms of yeast aspartic acid and phenylalanine transfer-RNA crystals, *Acta Crystallogr., Sect. A 44*, 112−123.
42. Szewczak, A. A., Moore, P. B., Chan, Y. L., and Wool, I. G. (1993) The conformation of the sarcin/ricin loop from 28S ribosomal RNA, *Proc. Natl. Acad. Sci. U.S.A. 90*, 9581−9585.
43. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K. H., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) Crystal structure of a group I ribozyme domain: Principles of RNA packing, *Science 273*, 1678−1685.
44. 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.
45. SantaLucia, J., Jr., and Turner, D. H. (1993) Structure of $(rGGCGAGCC)_2$ in solution from NMR and restrained molecular dynamics, *Biochemistry 32*, 12612−12623.
46. 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.
47. Hoffmann, B., Mitchell, G. T., Gendron, P., Major, F., Andersen, A. A., Collins, R. A., and Legault, P. (2003) NMR structure of the active conformation of the Varkud satellite ribozyme cleavage site, *Proc. Natl. Acad. Sci. U.S.A. 100*, 7003−7008.
48. 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.
49. Jucker, F. M., Heus, H. A., Yip, P. F., Moors, E. H. M., and Pardi, A. (1996) A network of heterogeneous hydrogen bonds in GNRA tetraloops, *J. Mol. Biol. 264*, 968−980.
50. Michiels, P. J. A., Schouten, C. H. J., 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.
51. Flinders, J., and Dieckmann, T. (2001) A pH controlled conformational switch in the cleavage site of the VS ribozyme substrate RNA, *J. Mol. Biol. 308*, 665−679.
52. Case, D. A. (1995) Calibration of ring-current effects in proteins and nucleic acids, *J. Biomol. NMR 6*, 341−346.
53. Nowakowski, J., Miller, J. L., Kollman, P. A., and Tinoco, I. (1996) Time evolution of NMR proton chemical shifts of an RNA hairpin during a molecular dynamics simulation, *J. Am. Chem. Soc. 118*, 12812−12820.
54. Jeffrey, G. A. (1997) *An Introduction to Hydrogen Bonding*, p 12, Oxford University Press, Inc., New York.
55. Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit, *Nature 407*, 327−339.
56. 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 289*, 905−920.
57. Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. (2001) High-resolution structure of the large ribosomal subunit from a mesophilic eubacterium, *Cell 107*, 679−688.
58. 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.
59. 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.
60. Gautheret, D., Konings, D., and Gutell, R. R. (1995) GU base pairing motifs in ribosomal RNA, *RNA 1*, 807−814.
61. Adams, P. L., Stahley, M. R., Kosek, A. B., Wang, J. M., and Strobel, S. A. (2004) Crystal structure of a self-splicing group I intron with both exons, *Nature 430*, 45−50.
62. Jovine, L., Hainzl, T., Oubridge, C., Scott, W. G., Li, J., Sixma, T. K., Wonacott, A., Skarzynski, T., and Nagai, K. (2000) Crystal structure of the Ffh and EF-G binding sites in the conserved domain IV of *Escherichia coli* 4.5S RNA, *Struct. Folding Des. 8*, 527−540.
63. 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.
64. Jang, S. B., Baeyens, K., Jeong, M. S., SantaLucia, J., Jr., Turner, D. H., and Holbrook, S. R. (2004) Structures of two RNA octamers containing tandem G·A base pairs, *Acta Crystallogr., Sect. D 60*, 829−835.
65. Kollman, P. (1993) Free energy calculations: Applications to chemical and biochemical phenomena, *Chem. Rev. 93*, 2395−2417.
66. Sarzynska, J., Nilsson, L., and Kulinski, T. (2003) Effects of base substitutions in an RNA hairpin from molecular dynamics and free energy simulations, *Biophys. J. 85*, 3445−3459.
67. Gautheret, D. F., 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.
68. Leonard, G. A., McAuleyhecht, K. E., Ebel, S., Lough, D. M., Brown, T., and Hunter, W. N. (1994) Crystal and molecular structure of r(CGCGAAUUAGCG): An RNA duplex containing two G(anti)·A(anti) base pairs, *Structure 2*, 483−494.
69. Lukavsky, P. J., Kim, I., Otto, G. A., and Puglisi, J. D. (2003) Structure of HCV IRES domain II determined by NMR, *Nat. Struct. Biol. 10*, 1033−1038.
70. Pan, B. C., Mitra, S. N., and Sundaralingam, M. (1999) Crystal structure of an RNA 16-mer duplex $r(GCAGAGUUAAAU-CUGC)_2$ with nonadjacent $G(syn)\cdot A^+(anti)$ mispairs, *Biochemistry 38*, 2826−2831.
71. Klein, D. J., Schmeing, T. M., Moore, P. B., and Steitz, T. A. (2001) The kink-turn: A new RNA secondary structure motif, *EMBO J. 20*, 4214−4221.
72. Winkler, W. C., Grundy, F. J., Murphy, B. A., and Henkin, T. M. (2001) The GA motif: An RNA element common to bacterial antitermination systems, rRNA, and eukaryotic RNAs, *RNA 7*, 1165−1172.
73. Goody, T. A., Melcher, S. E., Norman, D. G., and Lilley, D. M. J. (2004) The kink-turn motif in RNA is dimorphic, and metal ion-dependent, *RNA 10*, 254−264.
74. 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.
75. 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.
76. Doherty, E. A., Batey, R. T., Masquida, B., and Doudna, J. A. (2001) A universal mode of helix packing in RNA, *Nat. Struct. Biol. 8*, 339−343.
77. Nissen, P., Ippolito, J. A., Ban, N., Moore, P. B., and Steitz, T. A. (2001) RNA tertiary interactions in the large ribosomal subunit: The A-minor motif, *Proc. Natl. Acad. Sci. U.S.A. 98*, 4899−4903.
78. Batey, R. T., Rambo, R. P., Lucast, L., Rha, B., and Doudna, J. A. (2000) Crystal structure of the ribonucleoprotein core of the signal recognition particle, *Science 287*, 1232−1239.
79. Disney, M. D., Haidaris, C. G., and Turner, D. H. (2001) Recognition elements for 5′ exon substrate binding to the *Candida albicans* group I intron, *Biochemistry 40*, 6507−6519.
80. Disney, M. D., and Turner, D. H. (2002) Molecular recognition by the *Candida albicans* group I intron: Tertiary interactions with an imino G·A pair facilitate binding of the 5′ exon and lower the K_M for guanosine, *Biochemistry 41*, 8113−8119.
81. Pyle, A. M., Moran, S., Strobel, S. A., Chapman, T., Turner, D. H., and Cech, T. R. (1994) Replacement of the conserved GU with a GC pair at the cleavage site of the *Tetrahymena* ribozyme decreases binding, reactivity, and fidelity, *Biochemistry 33*, 13856−13863.
82. Misra, V. K., and Draper, D. E. (2001) A thermodynamic framework for Mg^{2+} binding to RNA, *Proc. Natl. Acad. Sci. U.S.A. 98*, 12456−12461.
83. Serra, M. J., Baird, J. D., Dale, T., Fey, B. L., Retatagos, K., and Westhof, E. (2002) Effects of magnesium ions on the stabilization of RNA oligomers of defined structures, *RNA 8*, 307−323.
84. Pyle, A. M. (2002) Metal ions in the structure and function of RNA, *J. Biol. Inorg. Chem. 7*, 679−690.

85. Auffinger, P., Bielecki, L., and Westhof, E. (2004) Symmetric K^+ and Mg^{2+} ion-binding sites in the 5 S rRNA loop E inferred from molecular dynamics simulations, *J. Mol. Biol. 335*, 555−571.
86. Davis, B., Afshar, M., Varani, G., Murchie, A. I. H., Karn, J., Lentzen, G., Drysdale, M., Bower, J., Potter, A. J., Starkey, I. D., Swarbrick, T., and Aboul-ela, F. (2004) Rational design of inhibitors of HIV-1 TAR RNA through the stabilization of electrostatic "hot spots", *J. Mol. Biol. 336*, 343−356.
87. Yoshizawa, S., Fourmy, D., Eason, R. G., and Puglisi, J. D. (2002) Sequence-specific recognition of the major groove of RNA by deoxystreptamine, *Biochemistry 41*, 6263−6270.

BI048079Y