An Alternating Sheared AA Pair and Elements of Stability for a Single Sheared Purine-Purine Pair Flanked by Sheared GA Pairs in RNA[†,‡]

Gang Chen,[§] Scott D. Kennedy,[||] Jing Qiao,[§] Thomas R. Krugh,[§] and Douglas H. Turner*[,§,⊥]

Department of Chemistry, University of Rochester, RC Box 270216, Rochester, New York 14627, and Department of Biochemistry and Biophysics and Center for Pediatric Biomedical Research and Department of Pediatrics, School of Medicine and Dentistry, University of Rochester, Rochester, New York 14642

Received November 30, 2005; Revised Manuscript Received March 8, 2006

ABSTRACT: A previous NMR structure of the duplex $\genfrac{}{}{0pt}{}{\mathrm{5'GGU}}{\mathrm{PCCG}}\frac{\mathrm{GGA}}{\mathrm{AAG}}\genfrac{}{}{0pt}{}{\mathrm{GGCU}}{\mathrm{CCG5'}}$ revealed an unusually stable RNA internal loop with three consecutive sheared GA pairs. Here, we report NMR studies of two duplexes, $\genfrac{}{}{0pt}{}{\mathrm{5'GGU}}{\mathrm{PCCA}}\frac{\mathrm{GGA}}{\mathrm{AAG}}\genfrac{}{}{0pt}{}{\mathrm{GGCU}}{\mathrm{CCG5'}}$ (replacing the UG pair with a UA closing pair) and $\genfrac{}{}{0pt}{}{\mathrm{5'GGU}}{\mathrm{PCCG}}\frac{\mathrm{GAA}}{\mathrm{AAG}}\genfrac{}{}{0pt}{}{\mathrm{GGCU}}{\mathrm{CCG5'}}$ (replacing the middle GA pair with an AA pair). An unusually stable loop with three consecutive sheared GA pairs forms in the duplex $\genfrac{}{}{0pt}{}{\mathrm{5'GGU}}{\mathrm{PCCA}}\frac{\mathrm{GGA}}{\mathrm{AAG}}\genfrac{}{}{0pt}{}{\mathrm{GGCU}}{\mathrm{CCG5'}}$. The structure contrasts with that reported for this loop in the crystal structure of the large ribosomal subunit of *Deinococcus radiodurans* [Harms, J., Schluenzen, F., Zarivach, R., Bashan, A., Gat, S., Agmon, I., Bartels, H., Franceschi, F., and Yonath, A. (2001) *Cell 107*, 679−688]. The middle AA pair in the duplex $\genfrac{}{}{0pt}{}{\mathrm{5'GGU}}{\mathrm{PCCG}}\frac{\mathrm{GAA}}{\mathrm{AAG}}\genfrac{}{}{0pt}{}{\mathrm{GGCU}}{\mathrm{CCG5'}}$ rapidly exchanges orientations, resulting in alternative base stacking and pseudosymmetry with exclusively sheared pairs. The $\genfrac{}{}{0pt}{}{\mathrm{U}}{\mathrm{G}}\frac{\mathrm{GAA}}{\mathrm{AAG}}\genfrac{}{}{0pt}{}{\mathrm{G}}{\mathrm{C}}$ internal loop is 2.1 kcal/mol less stable than the $\genfrac{}{}{0pt}{}{\mathrm{U}}{\mathrm{G}}\frac{\mathrm{GGA}}{\mathrm{AAG}}\genfrac{}{}{0pt}{}{\mathrm{G}}{\mathrm{C}}$ internal loop at 37 °C. Structural, energetic, and dynamic consequences upon functional group substitutions within related 3 × 3 and 3 × 6 internal loops are also reported.

Noncanonical pairs within the internal loops of RNA are important elements for folding and function. Understanding the sequence-dependent folding free energy and dynamics of internal loops can facilitate prediction of structure (*1*, *2*), dynamics, and functional significance from sequence.

AA and GA can form isosteric sheared-type (trans Hoogsteen/sugar edge A-A or A-G) noncanonical pairs (Figure 1a) (*3*−*9*). Typically, the AA pair is thermodynamically destabilizing, but the GA pair is stabilizing (*7*, *9*−*13*). Depending on the sequence context, GA often forms a sheared pair, but AA is more flexible (Figure 1). Two A's can potentially switch base pairing orientation in a sheared AA pair (i.e., trans Hoogsteen/sugar edge A1-A2 or A2-A1) without the loss of base−base hydrogen bonding. In a sheared GA pair, the equivalent interchange of bases would result in the loss of the two hydrogen bonds between G and A in a sheared GA pair.

The duplex $\genfrac{}{}{0pt}{}{\mathrm{5'GGU}}{\mathrm{PCCG}}\frac{\mathrm{GGA}}{\mathrm{AAG}}\genfrac{}{}{0pt}{}{\mathrm{GGCU}}{\mathrm{CCG5'}}$ (P[1] is a purine riboside) contains an unusually stable and relatively abundant internal loop, $\frac{\mathrm{GGA}}{\mathrm{AAG}}$ (*9*). The NMR structure of this duplex reveals three consecutive sheared GA pairs (trans Hoogsteen/sugar edge A-G) with separate stacks of three G's (G4, G5, and G14 in the major groove) and three A's (A6, A15, and A16 in the minor groove), which are closed by wobble UG (cis Watson−Crick/Watson−Crick U-G) and Watson−Crick CG pairs (*9*). (Throughout the paper, each top strand is written from 5′ to 3′ in going from left to right. Numbering starts at the left-most (5′) nucleotide of the top strand and ends at the left-most (3′) nucleotide of the bottom strand.)

Helix 68 of the crystal structure of the large ribosomal subunit of *Deinococcus radiodurans* contains a $\genfrac{}{}{0pt}{}{\mathrm{U}}{\mathrm{A}}\frac{\mathbf{G}\mathrm{GA}}{\mathbf{A}\mathrm{AG}}\genfrac{}{}{0pt}{}{\mathrm{G}}{\mathrm{C}}$ loop that has only one sheared GA pair (shown in bold) (*14*). There is less hydrogen bonding, and the base stacking pattern is equivalent to the A6/G5/A16 pattern in the minor groove instead of the A6/A15/A16 pattern found in the NMR structure for the equivalent loop with a UG rather than UA closing pair.

Here, we report NMR and thermodynamic studies of $\genfrac{}{}{0pt}{}{\mathrm{GGU}}{\mathrm{PCCA}}\frac{\mathrm{GGA}}{\mathrm{AAG}}\genfrac{}{}{0pt}{}{\mathrm{GGCU}}{\mathrm{CCG}}$ (A17 duplex) and $\genfrac{}{}{0pt}{}{\mathrm{GGU}}{\mathrm{PCCG}}\frac{\mathrm{GAA}}{\mathrm{AAG}}\genfrac{}{}{0pt}{}{\mathrm{GGCU}}{\mathrm{CCG}}$ (A5 duplex) to determine the effects of replacing a UG closing pair with UA and a middle GA pair with AA, respectively, relative to $\genfrac{}{}{0pt}{}{\mathrm{GGU}}{\mathrm{PCCG}}\frac{\mathrm{GGA}}{\mathrm{AAG}}\genfrac{}{}{0pt}{}{\mathrm{GGCU}}{\mathrm{CCG}}$ (3GA duplex) (Figure 2). NMR restrained molecular dynamics reveals a conformation of three consecutive sheared GA pairs for the loop in $\genfrac{}{}{0pt}{}{\mathrm{GGU}}{\mathrm{PCCA}}\frac{\mathrm{GGA}}{\mathrm{AAG}}\genfrac{}{}{0pt}{}{\mathrm{GGCU}}{\mathrm{CCG}}$. A5 and A15 in $\genfrac{}{}{0pt}{}{\mathrm{GGU}}{\mathrm{PCCG}}\frac{\mathrm{GAA}}{\mathrm{AAG}}\genfrac{}{}{0pt}{}{\mathrm{GGCU}}{\mathrm{CCG}}$ rapidly exchange positions, forming alternative sheared AA pairs (i.e., exchanging between

[†] This work was supported by NIH Grant GM22939 (D.H.T.).

[‡] Protein Data Bank entries 2DD1 (A17 duplex) and 2DD2 and 2DD3 (A5 duplex).

* To whom correspondence should be addressed. Phone: (585) 275-3207. Fax: (585) 276-0205. E-mail: turner@chem.rochester.edu.

[§] Department of Chemistry.

[||] Department of Biochemistry and Biophysics.

[⊥] Center for Pediatric Biomedical Research and Department of Pediatrics.

[1] Abbreviations: a, deoxyadenosine; C_T, total concentration of oligonucleotide strands; D, 2,6-diaminopurine riboside; g, deoxyguanosine; I, inosine; M, 2′-*O*-methyladenosine; P, purine riboside; R, any purine nucleotide; T_M, melting temperature in kelvin; T_m, melting temperature in degrees Celsius.

10.1021/bi0524464 CCC: $33.50
Published on Web 05/12/2006

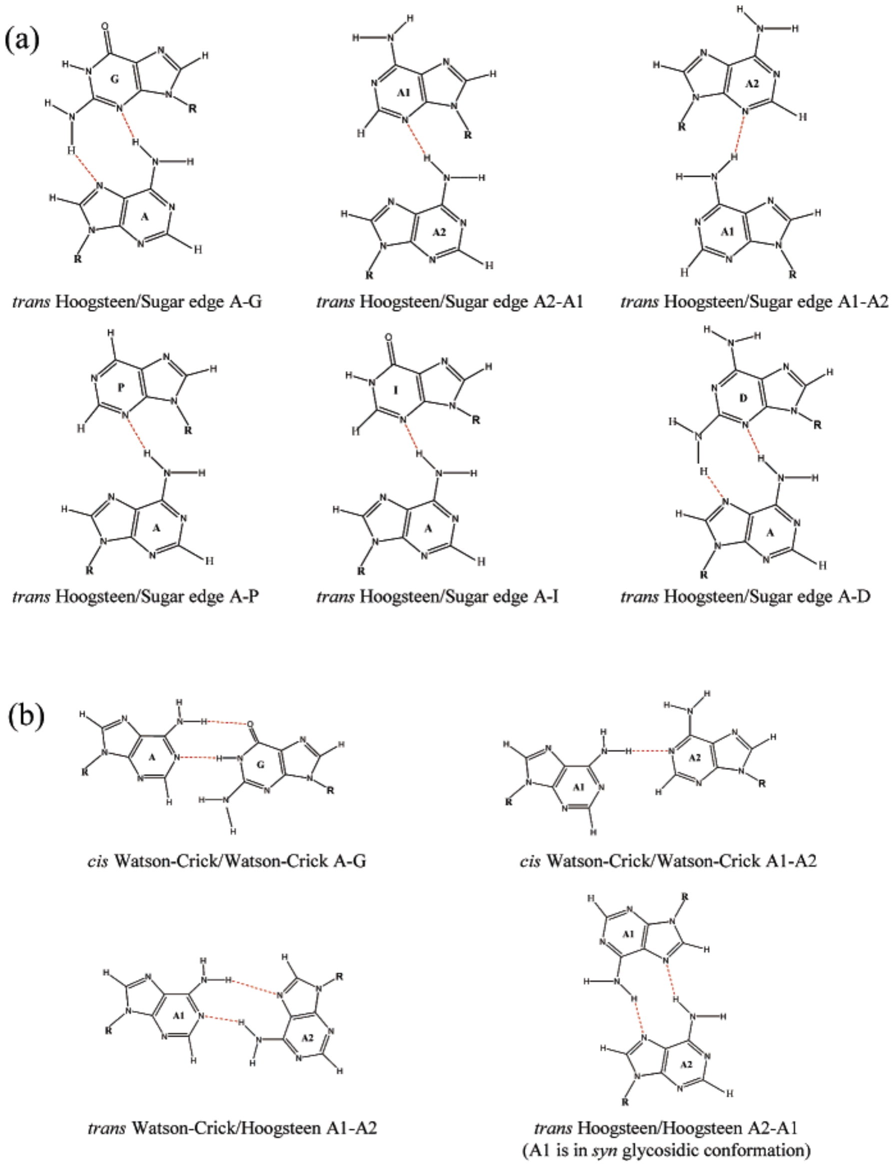

FIGURE 1: Schematic representation of (a) different sheared pairs and (b) a GA pair and various AA pairs mentioned in this paper. The hydrogen bonds between base and backbone are not shown. Note that two conformations with one base−base hydrogen bond are possible for a sheared AA pair because the amino group of either A can form the hydrogen bond. Only one such conformation is possible for the PA and IA pairs because neither P nor I has amino groups.

trans Hoogsteen/sugar edge A15-A5 and trans Hoogsteen/ sugar edge A5-A15) flanked by sheared GA pairs. The exchanging AA pair results in alternative A6/A15/A16 or A6/A5/A16 base stacking in the minor groove. The flexibility of alternative orientations of a middle adenine base edge in the minor groove, i.e., from A15 (N3-C2-N1) (as also observed in $\frac{\text{GGU}\ \underline{\text{GGA}}\ \text{GGCU}}{\text{PCCG}\ \overline{\text{AAG}}\ \text{CCG}}$ and $\frac{\text{GGU}\ \underline{\text{GGA}}\ \text{GGCU}}{\text{PCCA}\ \overline{\text{AAG}}\ \text{CCG}}$) to A5 (N1-C2-N3), might provide switching between different binding partners for dynamic functions.

Functional group substitutions (atomic mutations) have been extensively used for studying elements of molecular recognition in RNA (*10*, *15*−*24*). Here, the structural, energetic, and dynamic consequences of functional group substitutions are explored by studying duplexes of the form $\frac{\text{GGU}\ \underline{\text{GRA}}\ \text{GGCU}}{\text{PCCG}\ \overline{\text{AQG}}\ \text{CCG}}$, where R and Q are various purine nucleotides (Figures 1a and 2). Single predominant conformations form in the a5, P5, I5, and I15 duplexes. Functional group substitutions also facilitate interpretation of NMR data. The thermodynamic effects of functional group substitutions within 3 × 6 internal loops are also reported.

MATERIALS AND METHODS

Oligonucleotide Synthesis and Purification. Oligonucleotides were synthesized using the phosphoramidite method (*25*, *26*) and purified as described previously (*9*, *12*). CPG supports and phosphoramidites were acquired from Proligo, Glen Research, or ChemGenes. 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 and 80 °C and extinction coefficients predicted from those of dinucleotide monophosphates and nucleosides (*27*, *28*) with RNAcalc (http://www.meltwin.com) (*29*). The extinction coefficients were

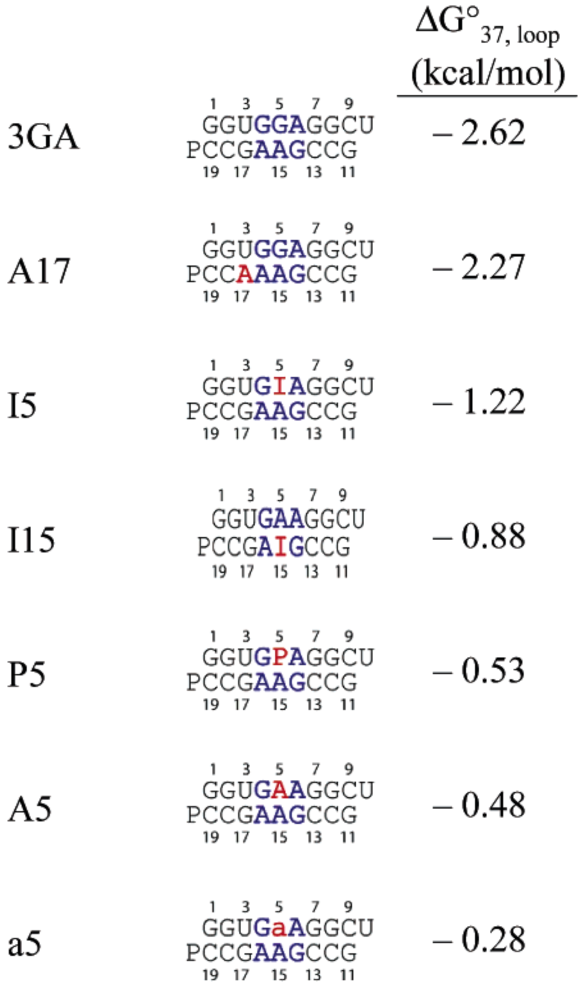

FIGURE 2: Secondary structure, numbering, and abbreviations for the duplexes studied previously (*9*, *13*) and here. The lowercase a represents deoxyadenosine. The value to the right of each duplex is the free energy increment in kilocalories per mole for formation of the internal loop at 37 °C and pH 7 in 1 M NaCl.

estimated by replacing purine riboside, 2,6-diaminopurine riboside, deoxyadenosine, and 2′-*O*-methyladenosine with adenosine and replacing inosine and deoxyguanosine with guanosine. Although extinction coefficients differ with functional group substitutions, individual nucleotides contribute only a small portion of the oligomer extinction and thus do not significantly affect thermodynamic measurements. UV melting buffer included 1.0 M NaCl, 20 mM sodium cacodylate, and 0.5 mM disodium EDTA (pH 7) or 80 mM NaCl, 10 mM sodium phosphate, and 0.5 mM disodium EDTA (pH 7). 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 Peltier temperature controller.

Melting curves were fit to a two-state model with MeltWin (http://www.meltwin.com), assuming linear sloping baselines and temperature-independent $\Delta H°$ and $\Delta S°$ (*29*−*31*). Additionally, the temperature at which half the strands are in a duplex, T_M, at a total strand concentration, C_T, was used to calculate thermodynamic parameters for non-self-complementary duplexes according to (*32*)

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

where R is the gas constant (1.987 cal mol^{-1} K^{-1}). All of the $\Delta H°$ 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 15%, suggesting that the two-state model is a reasonable approximation for these transitions. 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. With minor modification, sample preparation was similar to that previously reported (*7*, *9*). The sample buffer included 80 mM NaCl, 10 mM sodium phosphate, 0.5 mM disodium EDTA at pH 5.1 for H_2O and pD 7.3 for D_2O for $\substack{\text{GGU}\,\underline{\text{GAA}}\,\text{GGCU} \\ \text{PCCG}\,\overline{\text{AAG}}\,\text{CCG}}$ and pH 5.4 for H_2O and pD 6.8 for D_2O for $\substack{\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU} \\ \text{PCCA}\,\overline{\text{AAG}}\,\text{CCG}}$. Exchangeable proton spectra at pH 6.0 for $\substack{\text{GGU}\,\underline{\text{GAA}}\,\text{GGCU} \\ \text{PCCG}\,\overline{\text{AAG}}\,\text{CCG}}$ were very similar to those at pH 5.1. Moreover, chemical shifts and critical loop NOEs involving nonexchangeable protons were essentially the same in water at pH 5.1 and pD 7.3. Total volumes were 300 μL with a 90/10 (v/v) H_2O/D_2O mixture for exchangeable proton spectra and 99.996% D_2O (Cambridge Isotope Laboratories) for nonexchangeable spectra. The total duplex concentrations were ∼2 mM. The total duplex concentrations of other sequences were 0.5−1.2 mM.

NMR Spectroscopy. Unless otherwise noted, all exchangeable and nonexchangeable proton spectra were acquired on a Varian Inova 500 MHz (^{1}H) spectrometer (*33*). One-dimensional imino proton spectra were acquired with an S pulse sequence (*33*) with a sweep width of 12 kHz and temperatures ranging from 0 to 55 °C. SNOESY spectra were recorded with a mixing time of 150 ms at 5 and 30 °C. NOESY spectra of samples in D_2O were acquired at 30 °C with mixing times of 100, 200, and 400 ms. TOCSY spectra were acquired at 30 °C with mixing times of 8, 20, and 40 ms. Natural abundance ^{1}H−^{13}C HMQC spectra for $\substack{\text{GGU}\,\underline{\text{GAA}}\,\text{GGCU} \\ \text{PCCG}\,\overline{\text{AAG}}\,\text{CCG}}$ and $\substack{\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU} \\ \text{PCCA}\,\overline{\text{AAG}}\,\text{CCG}}$ were acquired with a 5000 Hz spectral width for proton and a 15 000 Hz spectral width for carbon. The ^{1}H−^{31}P HETCOR and natural abundance ^{1}H−^{13}C HSQC spectra were acquired on a Varian Inova 600 MHz (^{1}H) spectrometer. The one-dimensional ^{1}H-decoupled ^{31}P spectra (referenced to an external standard of 85% H_3PO_4 at 0 ppm) were acquired on a Bruker Avance 500 MHz (^{1}H) spectrometer at 30 °C. Proton spectra were referenced to H_2O or HDO at a known temperature-dependent chemical shift relative to 3-(trimethylsilyl)tetradeuterosodium propionate (TSP). The Felix (2000) software package (Molecular Simulations Inc.) was used to process two-dimensional spectra.

Restraint Generation. Very similar restraints were generated for $\substack{\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU} \\ \text{PCCA}\,\overline{\text{AAG}}\,\text{CCG}}$ (Table S1 of the Supporting Information) like they were for $\substack{\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU} \\ \text{PCCG}\,\overline{\text{AAG}}\,\text{CCG}}$ (*9*). For $\substack{\text{GGU}\,\underline{\text{GAA}}\,\text{GGCU} \\ \text{PCCG}\,\overline{\text{AAG}}\,\text{CCG}}$ and $\substack{\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU} \\ \text{PCCA}\,\overline{\text{AAG}}\,\text{CCG}}$, 15 hydrogen bond restraints limiting proton and hydrogen bond acceptor distances to 1.8−2.5 Å were applied for the five Watson−Crick GC pairs, but no hydrogen bond restraints were used within the loop and UG or UA pair. Dihedral angles of residues in the Watson−Crick stems and UG or UA pair were loosely restrained: $0 \pm 120°$ for α, $180 \pm 30°$ for β, $60 \pm 30°$ for γ, $85 \pm 30°$ for δ, $-140 \pm 40°$ for ϵ, $0 \pm 120°$ for ζ (ζ was mistakenly given as ξ in ref *9*), and $-170 \pm 40°$ for χ. For loop residues, glycosidic bond dihedral angles, χ's, were loosely restrained ($-120 \pm 90°$) because there was no indication of a syn glycosidic conformation. For the structural modeling of $\begin{smallmatrix}\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU}\\ \text{PCCA}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$, the δ dihedral angle for G5 was restrained to be *C*2′-*endo* with δ (160 ± 30°), and for A6, G14, U10, and P20, the δ dihedral angles were restrained to cover both *C*2′-*endo* and *C*3′-*endo* conformations with δ (122.5 ± 67.5°).

Two sets of distance and dihedral angle restraints (set I, A6/A15/A16; and set II, A6/A5/A16) were run for $\begin{smallmatrix}\text{GGU}\,\underline{\text{GAA}}\,\text{GGCU}\\ \text{PCCG}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$ because NOEs were inconsistent with a single structure. The previous NMR structure of $\begin{smallmatrix}\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU}\\ \text{PCCG}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$ (*9*) facilitated segregation of restraints for structural modeling. For interproton distance restraints that differ for A6/A15/A16 and A6/A5/A16 structural modeling, lower and upper bounds were loosened; all other restraints are the same for sets I and II (Table S2 of the Supporting Information).

Two sets of δ dihedral angle restraints were generated for loop residues (G4, A5, A6, G14, A15, and A16) in $\begin{smallmatrix}\text{GGU}\,\underline{\text{GAA}}\text{GGCU}\\ \text{PCCG}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$. For set I (A6/A15/A16), the δ dihedral angle for A5 was restrained to be *C*2′-*endo* with δ (160 ± 30°), and all other loop residues and the two 3′-dangling residues, U10 and P20, were restrained to cover both *C*2′-*endo* and *C*3′-*endo* conformations with δ (122.5 ± 67.5°). For set II (A6/A5/A16), all the loop residues were restrained to cover both *C*2′-*endo* and *C*3′-*endo* conformations, with the other dihedral angle restraints being the same as those of set I (A6/A15/A16) and $\begin{smallmatrix}\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU}\\ \text{PCCA}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$.

In summary, a total of 222 distance restraints (110 intranucleotide and 112 internucleotide), including hydrogen bond restraints, and 98 dihedral angle restraints were used for the structural modeling of $\begin{smallmatrix}\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU}\\ \text{PCCA}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$ (Table S1 of the Supporting Information). For the structural modeling of $\begin{smallmatrix}\text{GGU}\,\underline{\text{GAA}}\,\text{GGCU}\\ \text{PCCG}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$, a total of 250 distance restraints (128 intranucleotide and 122 internucleotide), including hydrogen bond restraints, and 98 dihedral angle restraints were used for the structural modeling of set I (A6/A15/A16), and a total of 249 distance restraints (128 intranucleotide and 121 internucleotide), including hydrogen bond restraints, and 98 dihedral angle restraints were used for the structural modeling of set II (A6/A5/A16) (Table S2 of the Supporting Information).

Structural Modeling. NMR restrained molecular dynamics and energy minimization were carried out 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). The AMBER 95 force field (*34*) was used with addition of flat-bottom restraint pseudopotentials, with force constants of 25 kcal mol^{-1} Å^{-2} for NOE distance restraints and 50 kcal mol^{-1} 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 (*35*), with a distance-dependent dielectric constant ($\epsilon = 2r$), was used for calculating electrostatic interactions. The progression of the structure simulation was the same as previously reported (*7*, *9*, *17*). Several figures were generated with PyMOL (*36*).

RESULTS

Functional Group Substitutions and Thermodynamics of Molecular Recognition. Measured thermodynamic parameters for several duplexes and internal loops with and without functional group substitutions are listed in Tables 1 and 2, respectively. Most were measured at 1 M NaCl to allow comparison to existing databases, but four were also measured in the 80 mM NaCl buffer used for most NMR experiments. The lower salt concentration makes duplex formation less favorable on average by 3.41 ± 0.15 kcal/mol at 37 °C, which is consistent with a sequence-independent salt effect. Measured thermodynamic parameters for formation of the internal loops (Table 2) are calculated according to the following equation which relies on the nearest neighbor model for predicting duplex stability (*37*):

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

For example

$$\Delta G°_{37\,\begin{smallmatrix}\text{U}\,\underline{\text{GGA}}\,\text{G}\\ \text{G}\,\overline{\text{AMG}}\,\text{C}\end{smallmatrix}} = \Delta G°_{37\,\begin{smallmatrix}\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU}\\ \text{PCCG}\,\overline{\text{AMG}}\,\text{CCG}\end{smallmatrix}} - \Delta G°_{37\,\begin{smallmatrix}\text{GGU}\,\text{GGCU}\\ \text{PCCG}\,\text{CCG}\end{smallmatrix}} + \Delta G°_{37\begin{smallmatrix}\text{U}\,\text{G}\\ \text{G}\,\text{C}\end{smallmatrix}} \quad (2b)$$

where $\Delta G°_{37\begin{smallmatrix}\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU}\\ \text{PCCG}\,\overline{\text{AMG}}\,\text{CCG}\end{smallmatrix}}$ is the measured value of the duplex containing the internal loop (Table 1), $\Delta G°_{37\begin{smallmatrix}\text{GGU}\,\text{GGCU}\\ \text{PCCG}\,\text{CCG}\end{smallmatrix}}$ is the measured value of the duplex without the loop (*9*), and $\Delta G°_{37\,\begin{smallmatrix}\text{U}\,\text{G}\\ \text{G}\,\text{C}\end{smallmatrix}}$ is the free energy increment for the nearest neighbor base stack interaction interrupted by the internal loop (*1*, *31*). $\Delta H°_{\text{loop}}$ and $\Delta S°_{\text{loop}}$ are calculated similarly. All the thermodynamic parameters used in this calculation are derived from T_M^{-1} versus $\ln(C_T/4)$ plots (eq 1). When eq 2a was applied to 2 × 2 nucleotide internal loops of noncanonical pairs flanked by different stems, the values for $\Delta G°_{37,\text{loop}}$ for a given loop sequence differed by an average of 0.40 kcal/mol (*2*). The model should be even better for sequences with identical stems.

Functional Group Substitutions, NMR Assignments, and Structural Features. The base−(H1′/H5) "NOESY walk" regions of the 400 ms NOESY spectra at 30 °C are shown in Figure 3. NMR resonances were assigned essentially as described previously (*7*, *9*, *38*, *39*). Comparison of spectra with those of the duplex $\begin{smallmatrix}\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU}\\ \text{PCCG}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$ (*9*) facilitates NMR assignments of $\begin{smallmatrix}\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU}\\ \text{PCCA}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$ and $\begin{smallmatrix}\text{GGU}\,\underline{\text{GAA}}\,\text{GGCU}\\ \text{PCCG}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$ (see Tables S1−S3 of the Supporting Information for assignments and restraints used for structural modeling).

U3A17 in $\begin{smallmatrix}\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU}\\ \text{PCCA}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$ forms a Watson−Crick pair as indicated by a strong NOE between U3H3 and A17H2. The imino proton, U3H3, is relatively broad and shifted upfield (12.53 ppm) (*13*) relative to the usual range of 13−15 ppm for a Watson−Crick UA pair. A similar upfield shift (11.75 ppm) was observed in the 2 × 2 loop $\begin{smallmatrix}\text{U}\,\underline{\text{GA}}\,\text{A}\\ \text{A}\,\overline{\text{AG}}\,\text{U}\end{smallmatrix}$ (*40*). Three consecutive sheared GA pairs form in $\begin{smallmatrix}\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU}\\ \text{PCCA}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$, as in $\begin{smallmatrix}\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU}\\ \text{PCCG}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$ (Figures 4 and 5a) (*9*). Several medium to strong NOEs, which are similar to NOEs observed for $\begin{smallmatrix}\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU}\\ \text{PCCG}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$, define the loop structure, e.g., G14H2′−G5H1, A6H1′−A15H2, A15H1′−A6H2, A16H1′−A15H2, G7H1′−A6H2, and A17H1′−A16H2 (compare to G17H1′−A16H2 in $\begin{smallmatrix}\text{GGU}\,\underline{\text{GGA}}\,\text{GGCU}\\ \text{PCCG}\,\overline{\text{AAG}}\,\text{CCG}\end{smallmatrix}$) (Figure 3a and Table 3) (*9*). In the loop of $\begin{smallmatrix}\text{U}\,\underline{\text{GGA}}\,\text{G}\\ \text{A}\,\overline{\text{AAG}}\,\text{C}\end{smallmatrix}$, G4 and G5 have *C*3′-*endo* and *C*2′-*endo* sugar puckers, respectively, and G14 is populated in both conformations as evidenced by TOCSY (Figure S1a of the

Table 1: Measured Thermodynamic Parameters for Duplex Formation in 1 M NaCl (pH 7) and in 80 mM NaCl (Listed in Parentheses)

Sequences	T_M^{-1} vs ln(C_T/4) plots (eq 1)				Average of melt curve fits			
	$-\Delta H°$ (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 PCCG**AMG**CCG	108.4±3.5	302.9±10.4	14.47±0.24	61.5	105.1±1.2	293.0±3.6	14.24±0.12	61.6
GGU**GGA**GGCU PCCG**DAG**CCG	104.2±3.1	290.9±9.4	13.96±0.22	60.8	105.4±3.0	294.4±9.0	14.04±0.23	60.8
GGU**GgA**GGCU PCCG**AAG**CCG	103.9±3.1	290.8±9.3	13.68±0.20	59.9	100.1±3.0	279.3±8.9	13.44±0.21	60.0
GGU**GGA**GGCU[b] PCCG**AAG**CCG	94.3±8.2 (82.2±3.4)	261.2±24.5 (233.5±10.7)	13.26±0.57 (9.79±0.10)	60.8 (49.8)	94.5±2.4 (87.8±4.5)	261.9±7.2 (251.1±14.1)	13.27±0.25 (9.94±0.15)	60.8 (49.5)
GGU**GGA**GGCU[b] PCCA**AAG**CCG	92.7±2.2 (84.6±4.3)	258.0±6.6 (243.2±13.6)	12.64±0.13 (9.15±0.13)	58.9 (46.9)	93.2±3.9 (81.9±9.3)	259.7±12.0 (234.9±29.2)	12.67±0.23 (9.09±0.30)	58.9 (47.0)
GGC**GAA**GGCU[b] PCCG**AAG**CCG	81.2±7.0 (56.1±1.9)	223.4±21.2 (152.6±6.1)	11.92±0.47 (8.74±0.05)	59.1 (49.7)	77.8±5.8 (60.7±8.3)	213.1±17.7 (167.2±26.3)	11.76±0.36 (8.88±0.16)	59.3 (49.5)
GGU**GDA**GGCU PCCG**AAG**CCG	93.6±1.9	262.9±5.9	12.09±0.11	56.6	89.4±4.4	250.1±13.6	11.88±0.22	56.7
GGU**GIA**GGCU PCCG**AAG**CCG	90.7±2.5	254.3±7.7	11.86±0.14	56.4	89.6±5.0	250.8±15.3	11.84±0.27	56.5
GGU**GAA**GGCU PCCG**AIG**CCG	90.6±3.4	255.0±10.4	11.52±0.18	55.1	92.0±3.4	259.0±10.2	11.61±0.20	55.1
GGU**GAA**GGCU PCCG**AMG**CCG	92.6±3.8	261.7±11.7	11.47±0.19	54.5	90.3±4.6	254.4±14.1	11.38±0.23	54.6
GGU**GAA**GGCU PCCG**DAG**CCG	89.2±1.8	250.8±5.6	11.42±0.09	55.0	88.5±4.6	248.6±14.1	11.41±0.23	55.1
GGU**GPA**GGCU PCCG**AAG**CCG	88.9±2.1	250.5±6.4	11.17±0.10	54.1	84.3±2.4	236.6±7.5	10.97±0.10	54.2
GGU**GAA**GGCU[b] PCCG**AAG**CCG	84.2±6.1 (75.7±3.3)	235.7±18.6 (219.4±10.4)	11.12±0.32 (7.64±0.05)	54.9 (41.6)	86.5±5.4 (74.9±3.6)	242.8±16.9 (216.8±11.4)	11.23±0.25 (7.64±0.11)	54.8 (41.7)
GGU**GaA**GGCU PCCG**AMG**CCG	88.7±3.4	250.9±10.6	10.93±0.16	53.2	85.5±2.4	240.9±7.3	10.79±0.14	53.3
GGU**GaA**GGCU PCCG**AAG**CCG	86.7±1.5	244.2±4.7	10.92±0.07	53.5	82.9±1.5	232.7±4.8	10.75±0.07	53.6
GGU**GGA**___GGCU[c] PCCG**AAGAAA**CCG	90.8±1.9	259.1±6.0	10.47±0.07	51.1	84.5±4.2	239.4±13.0	10.25±0.19	51.3
GGU**GgA**___GGCU PCCG**AAGAAA**CCG	89.4±2.9	256.1±8.9	9.97±0.09	49.4	81.3±5.2	230.8±16.2	9.74±0.20	49.7
GGU**GIA**GGCU PCCG**AIG**CCG	80.7±1.4	228.9±4.4	9.70±0.05	49.7	78.5±1.9	222.2±5.9	9.61±0.08	49.7
GGU**GIA**___GGCU PCCG**AAGAAA**CCG	80.3±4.1	231.7±13.0	8.47±0.09	44.7	71.8±2.5	204.8±7.8	8.31±0.14	44.9
GGU**GDA**___GGCU PCCG**AAGAAA**CCG	70.0±3.4	199.3±11.0	8.17±0.06	44.4	65.2±2.9	184.0±9.3	8.12±0.13	44.8
GGU**GAA**___GGCU PCCG**AAGAAA**CCG	73.6±4.3	211.1±13.7	8.07±0.08	43.6	67.8±7.1	192.8±22.5	8.01±0.12	43.9
GGU**GaA**___GGCU PCCG**AAGAAA**CCG	72.7±4.8	209.0±15.3	7.91±0.09	43.0	66.1±3.3	187.7±10.4	7.84±0.14	43.3
GGU**GPA**___GGCU PCCG**AAGAAA**CCG	66.1±3.9	188.7±12.3	7.57±0.07	42.0	59.2±2.2	166.6±7.3	7.56±0.14	42.5
GGUGGCU[b] PCCGCCG	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

[a] At 0.1 mM (C_T). [b] Data measured in 1 M NaCl are from ref *9*. [c] From ref *13*.

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

Sequence	$\Delta G°_{37,\,loop}$ (kcal/mol)	$\Delta H°_{loop}$ (kcal/mol)	$\Delta S°_{loop}$ (eu)	$\Delta\Delta G°_{37,\,loop}$ (kcal/mol)
GGU**GGA**GGCU PCCG**AMG**CCG	–3.83±0.59	–38.4±9.9	–111.4±30.3	–1.21[b]
GGU**GGA**GGCU PCCG**DAG**CCG	–3.32±0.58	–34.2±9.7	–99.4±30.0	–0.70[b]
GGU**GgA**GGCU PCCG**AAG**CCG	–3.04±0.57	–33.9±9.7	–99.3±29.9	–0.42[b]
GGU**GGA**GGCU[c] PCCG**AAG**CCG	–2.62±0.78 [–2.39]	–24.3±12.4	–69.7±37.5	–
GGU**GGA**GGCU[c] PCCA**AAG**CCG	–2.27±0.59 [–1.44]	–23.9±9.7	–69.5±29.5	–
GGU**GDA**GGCU PCCG**AAG**CCG	–1.45±0.54	–23.6±9.4	–71.4±29.0	1.17[b], –0.97[d]
GGU**GIA**GGCU PCCG**AAG**CCG	–1.22±0.55	–20.7±9.6	–62.8±29.4	1.40[b], –0.74[d]
GGU**GAA**GGCU PCCG**AIG**CCG	–0.88±0.56	–20.6±9.8	–63.5±30.3	1.74[b], –0.40[d]
GGU**GAA**GGCU PCCG**AMG**CCG	–0.83±0.57	–22.6±10.0	–70.2±30.8	–0.35[d]
GGU**GAA**GGCU PCCG**DAG**CCG	–0.78±0.54	–19.2±9.4	–59.3±29.0	–0.30[d]
GGU**GPA**GGCU PCCG**AAG**CCG	–0.53±0.54	–18.9±9.4	–59.0±29.1	2.09[b], –0.05[d]
GGU**GAA**GGCU[c] PCCG**AAG**CCG	–0.48±0.57 [–0.03]	–14.2±11.1	–44.2±34.0	0.00[d]
GGU**GaA**GGCU PCCG**AMG**CCG	–0.29±0.56	–18.7±9.8	–59.4±30.4	0.19[d]
GGU**GaA**GGCU PCCG**AAG**CCG	–0.28±0.54	–16.7±9.3	–52.7±28.8	0.20[d]
GGU**GGA**GGCU[e] PCCG**AAGAAA**CCG	0.17±0.54 [0.20]	–20.8±9.4	–67.6±29.0	0.00[f]
GGU**GgA**GGCU PCCG**AAGAAA**CCG	0.67±0.54	–19.4±9.6	–64.6±29.8	0.50[f]
GGU**GIA**GGCU PCCG**AIG**CCG	0.94±0.54	–10.7±9.3	–37.4±28.8	–
GGU**GIA**GGCU PCCG**AAGAAA**CCG	2.17±0.54	–10.3±10.1	–40.2±31.2	2.00[f], –0.40[g]
GGU**GDA**GGCU PCCG**AAGAAA**CCG	2.47±0.54	0.0±9.8	–7.8±30.5	2.30[f], –0.10[g]
GGU**GAA**GGCU PCCG**AAGAAA**CCG	2.57±0.54 [2.56]	–3.6±10.2	–19.6±31.6	0.00[g]
GGU**GaA**GGCU PCCG**AAGAAA**CCG	2.73±0.54	–2.7±10.4	–17.5±32.3	2.56[f], 0.16[g]
GGU**GPA**GGCU PCCG**AAGAAA**CCG	3.07±0.54	3.9±10.0	2.8±31.0	2.90[f], 0.50[g]

[a] Experimental errors for $\Delta G°_{37}$, $\Delta H°$, and $\Delta S°$ for the canonical stems are estimated to be 4, 12, and 13.5%, respectively, according to ref *31*. 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 brackets are predicted according to ref *13*. [b] Compared with (GGU GGA GGCU / PCCG AAG CCG) from ref *9*. [c] From ref *9*. [d] Compared with (GGU GAA GGCU / PCCG AAG CCG) from ref *9*. [e] From ref *13*. [f] Compared with (GGU GGA GGCU / PCCG AAGAAA CCG) from ref *13*. [g] Compared with (GGU GAA GGCU / PCCG AAGAAA CCG).

Supporting Information) and NOESY (Figure 3a) spectra. The same sugar puckers are found in (GGU GGA GGCU / PCCG AAG CCG) (*9*).

The similarity of sugar edges (*6*) of adenosine, purine riboside, and inosine (Figure 1) facilitates H2 assignments

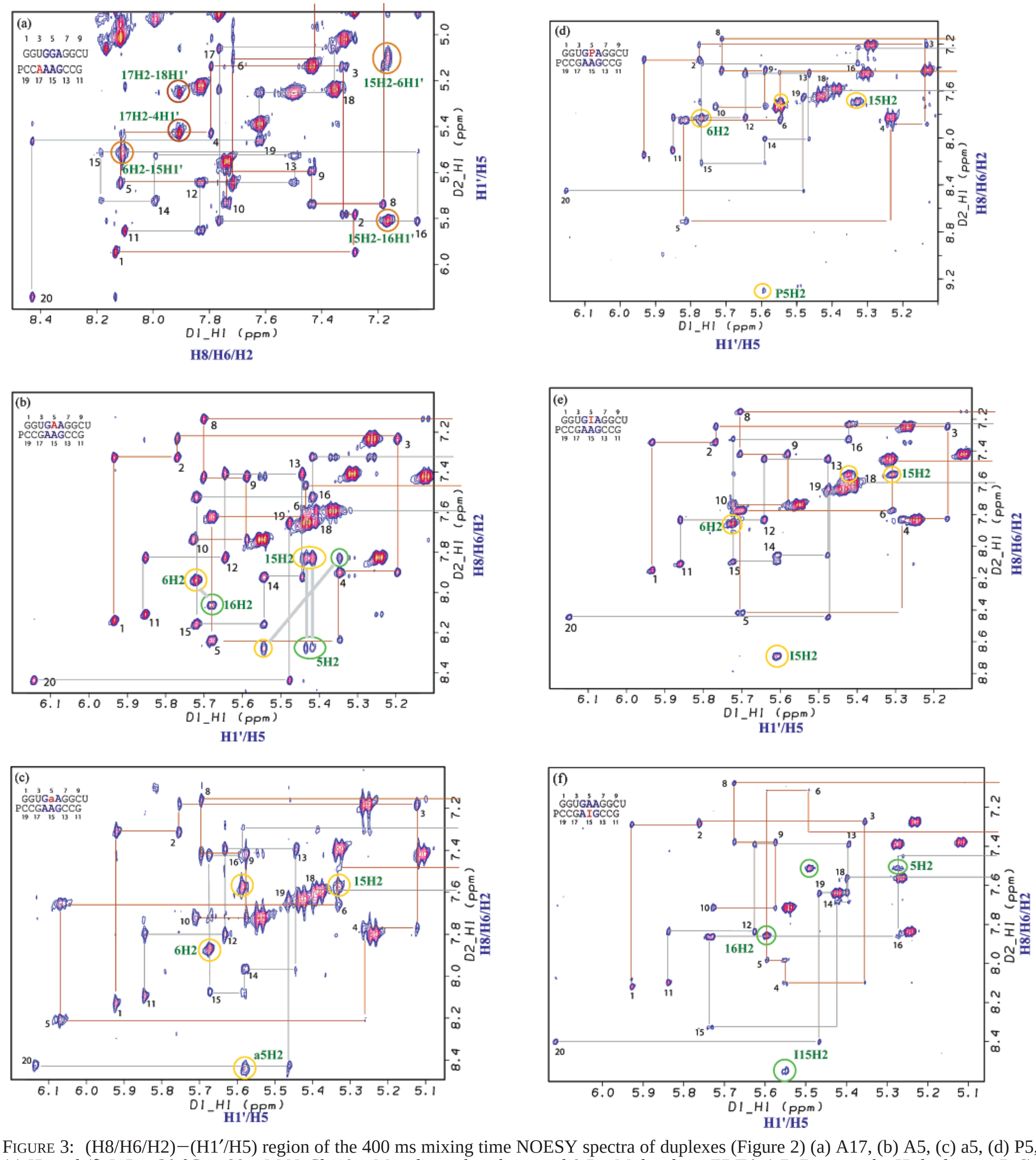

FIGURE 3: (H8/H6/H2)−(H1′/H5) region of the 400 ms mixing time NOESY spectra of duplexes (Figure 2) (a) A17, (b) A5, (c) a5, (d) P5, (e) I5, and (f) I15 at 30 °C in 80 mM NaCl, 10 mM sodium phosphate, and 0.5 mM disodium EDTA (pD 7 except the I5 duplex at pD 6). For the A5 sequence in panel b, yellow and green circles or ovals connected by gray lines identify related cross-peaks of major and minor conformations, respectively. Yellow and green circles or ovals in other spectra identify cross-peaks related to those in circles or ovals of the same color for the A5 duplex.

of middle purine-purine pairs. Also, the NOEs of G14H1′/H2′−A5H2 in $\frac{\text{GGU }\overline{\text{GAA}}\text{ GGCU}}{\text{PCCG }\underline{\text{AAG}}\text{ CCG}}$ (Figure 3b) are similar to those of G14H1′/H2′−G5H1 as present for $\frac{\text{GGU }\overline{\text{GGA}}\text{ GGCU}}{\text{PCCG }\underline{\text{AAG}}\text{ CCG}}$ (*9*) and $\frac{\text{GGU }\overline{\text{GGA}}\text{ GGCU}}{\text{PCCA }\underline{\text{AAG}}\text{ CCG}}$ (Table S1). Both sets of NOEs, G14H1′/H2′−I5H2 and G14H1′/H2′−I5H1 (data not shown), are observed in the I5 duplex, $\frac{\text{GGU }\overline{\text{GIA}}\text{ GGCU}}{\text{PCCG }\underline{\text{AAG}}\text{ CCG}}$ (Figure 3e). The chemical shift of I5H1, 12.1 ppm (Figure S2 of the Supporting Information), is in agreement with the formation of a sheared IA pair (*41*). Downfield chemical shifts for inosine imino protons beyond 14 ppm were observed in face-to-face IA pairs (*4*, *42*).

While comparison of NOESY spectra (Figure 3, Table 3, and Table S2 of the Supporting Information) indicates very similar base pairing and stacking geometries with three sheared-type purine-purine pairs for $\frac{\text{GGU }\overline{\text{GaA}}\text{ GGCU}}{\text{PCCG }\underline{\text{AAG}}\text{ CCG}}$ (a5 duplex), $\frac{\text{GGU }\overline{\text{GPA}}\text{ GGCU}}{\text{PCCG }\underline{\text{AAG}}\text{ CCG}}$ (P5 duplex), and $\frac{\text{GGU }\overline{\text{GIA}}\text{ GGCU}}{\text{PCCG }\underline{\text{AAG}}\text{ CCG}}$ (I5

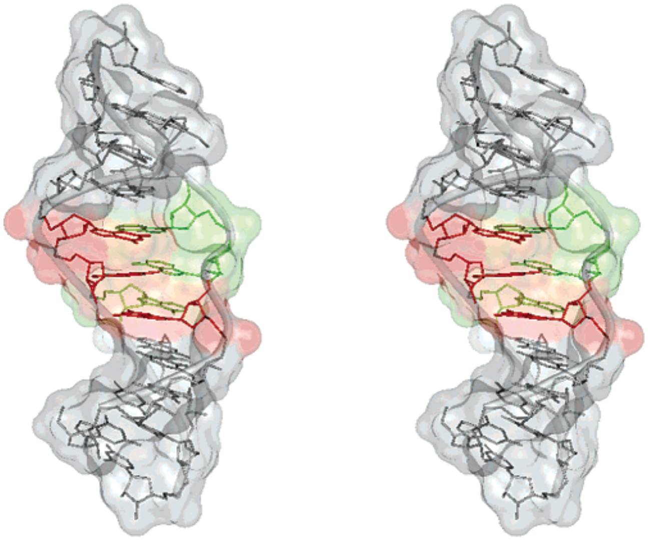

FIGURE 4: Major groove stereoview of $\substack{\text{GGU }\underline{\text{GGA}}\text{ GGCU}\\ \text{PCCA }\overline{\text{AAG}}\text{ CCG}}$, with the A6/A15/A16 stack in the minor groove colored green and the G4/G5/G14 stack in the major groove colored red. Hydrogen and nonbridging oxygen atoms were omitted for clarity.

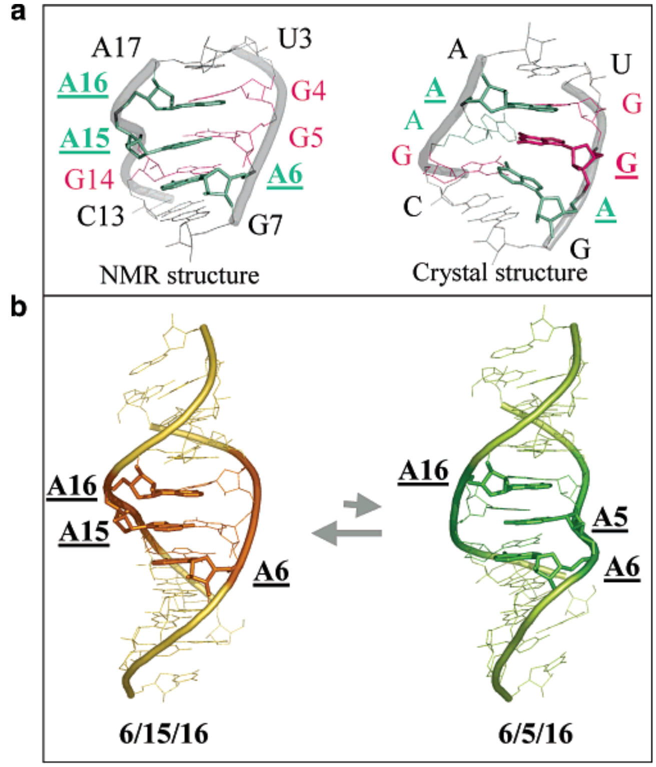

FIGURE 5: Comparison of the minor groove view of the crystal and NMR structures. Residues stacking in the minor groove are shown in sticks and labeled in bold. Hydrogen and nonbridging oxygen atoms were omitted for clarity. (a) NMR and crystal structures of the internal loop $\substack{\text{U }\underline{\text{GGA}}\text{ G}\\ \text{A }\overline{\text{AAG}}\text{ C}}$. Three G's and three A's in the loop are colored red and green, respectively. (b) NMR structures of the duplex $\substack{\text{GGU }\underline{\text{GAA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$ with an alternating middle sheared AA pair.

duplex), the NMR spectra provide evidence for two populations of structures for the middle A5A15 pair in $\substack{\text{GGU }\underline{\text{GAA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$ (A5 duplex) (Figures 5b and 6). For example, in addition to NOEs of R5H2−G14H1′/H2′ (R is any purine), A6H2−A15H1′, A15H2−A6H1′, and A15H2−A16H1′ (denoted with yellow circles and ovals) for the a5, P5, and I5 duplexes, one extra set of NOEs A15H2−G4H1′/H2′, A16H2−A5H1′, A5H2−A16H1′, and A5H2−A6H1′ (denoted with green circles and ovals) are present for $\substack{\text{GGU }\underline{\text{GAA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$ (Figure 3). Moreover, on the basis of TOCSY (Figure S1 of the Supporting Information) and NOESY (Figure 3) spectra, sugar puckers of a5, P5, and I5 are *C2′-endo* as indicated by strong H1′−H2′ couplings (≥ 8 Hz), which corresponds to the *C2′-endo* conformation of the G5 sugar pucker in $\substack{\text{GGU }\underline{\text{GGA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$ (*9*). Sugar puckers for A15 are *C3′-endo* in each of these duplexes with the possible exception of the a5 duplex. For $\substack{\text{GGU }\underline{\text{GAA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$, however, A5 and A15 are populated in both *C2′-endo* and *C3′-endo* conformations with A5 having a greater *C2′-endo* population than A15. Evidently, the middle A5A15 pair is more populated in trans Hoogsteen/sugar edge A15-A5 than in trans Hoogsteen/sugar edge A5-A15 and related conformations. This is in agreement with NOEs (Figure 3b) observed for A5H2−G14H1′/H2′, A6H2−A15H1′, A15H2−A6H1′, and A15H2−A16H1′ (denoted with yellow circles and ovals) being relatively stronger than NOEs A15H2−G4H1′/H2′, A16H2−A5H1′, A5H2−A16H1′, and A5H2−A6H1′ (denoted with green circles and ovals), respectively, in $\substack{\text{GGU }\underline{\text{GAA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$. The presence of a single set of chemical shifts for $\substack{\text{GGU }\underline{\text{GAA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$ indicates that the middle A5A15 pair is alternating rapidly (fast exchange on the NMR time scale), with the sugar edge of either A5 or A15 on the base pairing edge of the other, forming trans Hoogsteen/sugar edge A15-A5 (denoted with yellow) or A5-A15 (denoted with green) and related pairs (Figure 5b).

NMR spectra acquired at 1 M NaCl indicate that the structural and dynamical properties of the A5 loop are similar at 80 mM and 1 M NaCl (Figures 3b and S3, respectively). Chemical shifts show an only modest salt dependence, and the same pattern of NOEs for the major and minor conformations is observed, including the A5H2−A15H8 (major) and A15H2−A5H8 (minor) cross-peaks (data not shown).

Relatively downfield chemical shifts of H2 protons on the base pairing edge (Figure 1) are observed: 9.30 ppm for P5H2 [compared to 8.16 ppm for P20H2 (*9*)], 8.44 ppm for a5H2 (compared to 7.57 ppm for A15H2 in the a5 duplex), and 8.69 ppm for I5H2 in $\substack{\text{GGU }\underline{\text{GPA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$, $\substack{\text{GGU }\underline{\text{GaA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$, and $\substack{\text{GGU }\underline{\text{GIA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$, respectively (Figure 3 and Table 4). Such downfield chemical shifts of H2 on the base pairing edge of sheared purine-purine pairs are expected due to ring current deshielding effects (Figure 1a), as observed previously in 2 × 2 loops: 8.19 ppm for A4H2 in $\substack{\text{GGC }\underline{\text{AA}}\text{ GGCU}\\ \text{UCCG }\overline{\text{AA}}\text{ CCG}}$ (compared to 7.88 ppm for A5H2, which is not on a base pairing edge) and 8.97 ppm for P4H2 in $\substack{\text{GGC }\underline{\text{PA}}\text{ GGCU}\\ \text{UCCG }\overline{\text{AP}}\text{ CCG}}$ (*18*).

On the edge that is not base paired, relatively upfield chemical shifts of H2 protons are observed. A15H2 chemical shifts are 7.69, 7.57, and 7.55 ppm in $\substack{\text{GGU }\underline{\text{GPA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$, $\substack{\text{GGU }\underline{\text{GaA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$, and $\substack{\text{GGU }\underline{\text{GIA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$, respectively (Figure 3 and Table 4). These can be compared to the relatively further upfield chemical shifts of 7.13 and 7.16 ppm for A15H2 protons in $\substack{\text{GGU }\underline{\text{GGA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$ and $\substack{\text{GGU }\underline{\text{GGA}}\text{ GGCU}\\ \text{PCCA }\overline{\text{AAG}}\text{ CCG}}$, respectively, which might reflect stronger base pairing and better stacking of the motif with three consecutive sheared GA pairs, resulting in larger ring current shielding effects from A6 and A16. Intermediate chemical shifts of A5H2 (8.28 ppm) and A15H2 (7.83 ppm) are observed in $\substack{\text{GGU }\underline{\text{GAA}}\text{ GGCU}\\ \text{PCCG }\overline{\text{AAG}}\text{ CCG}}$, which is consistent with a rapidly alternating sheared A5A15 pair (Table 4).

Table 3: Distance Restraints Involving A5 and A15 Residues for the Structural Modeling of $\frac{\text{GGU GAA GGCU}}{\text{PCCG AAG CCG}}$ and Comparison with That of $\frac{\text{GGU GGA GGCU}}{\text{PCCG AAG CCG}}$ (*9*). Lower and Upper Bounds were Calculated from the NMR Derived Distances[a]

Distance Restraints that Differ for the Structural Modeling of A6/A15/A16 and A6/A5/A16

A6/A15/A16						A6/A5/A16					
		distance (Å)						distance (Å)			
atom 1	atom 2	lower	upper	NMR	model[b]	atom 1	atom 2	lower	upper	NMR	model[b]
G4H3′	A5H8	1.80	5.00	3.35	3.50/4.73	G14H3′	A15H8	1.80	5.00	2.79	4.47/2.80
G4H8	A5H8	3.40	6.00	4.86	5.41/6.68	G14H8	A15H8	×[c]	×[c]	×[c]	6.68/5.51
A5H1′	A6H8	1.80	4.12	3.17	3.11/4.98	A15H1′	A16H8	1.80	4.34	3.34	5.20/2.75
G14H1′	A5H2	1.80	5.00		4.62/10.32	G4H1′	A15H2	1.80	5.00		10.22/4.23
G14H2′	A5H2	2.00	5.00		2.21/9.41	G4H2′	A15H2	1.80	5.00		9.98/2.27
A5H2	A15H8	2.00	5.00		2.30/10.41	A5H8	A15H2	1.80	5.00		10.47/2.35
A5H4′	A16H2	2.83	5.26	4.05	4.02/5.38	A15H4′	A6H2	3.00	6.00		5.52/4.44
A6H1′	A15H2	1.80	4.86	3.74	2.55/7.66	A16H1′	A5H2	1.80	5.00		7.11/2.88
A15H1′	A6H2	1.80	4.14	3.19	2.84/6.60	A5H1′	A16H2	1.80	5.00		6.08/2.74
G14H1′	A15H8	2.00	5.50	4.23	4.69/5.99	G4H1′	A5H8	1.80	5.52	4.24	5.73/4.82
A16H1′	A15H2	1.80	4.30	3.31	2.80/7.39	A6H1′	A5H2	1.80	5.00		7.75/2.80
A15H8	A16H8	3.08	5.73	4.41	4.08/6.16	A5H8	A6H8	3.27	6.00	4.67	6.58/4.06
A15H2′	A16H8	1.80	3.99	3.07	3.32/4.64	A5H2′	A6H8	1.80	4.66	3.58	5.07/3.13
A15H3′	A16H8	1.80	5.00		2.48/4.80	A5H3′	A6H8	1.80	4.91	3.77	4.86/2.47

Distance Restraints that Are the Same for A6/A15/A16 and A6/A5/A16, Unless Otherwise Noted

		distance (Å)						distance (Å)			
atom 1	atom 2	lower	upper	NMR	model[b]	atom 1	atom 2	lower	upper	NMR	model[b]
G4H2′	A5H8	1.80	5.00	4.86	3.08/3.10	G14H2′	A15H8	1.80	4.76	3.66	2.83/3.44
A5H1′	A5H2′	1.80	3.94	3.03	2.85/2.45	A15H1′	A15H2′	1.80	3.65	2.80	2.61/2.71
A5H1′	A5H3′	1.80	4.81	3.70	3.70/3.56	A15H1′	A15H3′	1.80	5.00	4.19	3.66/3.69
A5H1′	A5H4′	1.80	4.41	3.39	3.36/3.16	A15H1′	A15H4′	1.80	4.42	3.40	3.32/3.27
A5H1′	A5H8	1.80	4.97	3.82	3.72/3.41	A15H1′	A15H8	1.80	5.04	3.88	3.48/3.68
A5H2′	A5H8	1.80	3.19[d]	2.45	2.10/3.97	A15H2′	A15H8	1.80	5.00	2.79	4.11/2.03
A5H3′	A5H8	1.80	5.00	3.35	4.09/3.42	A15H3′	A15H8	1.80	4.27	3.28	3.42/3.89

Distance Restraints Involving G5 and A15 Residues for the Structural Modeling of $\frac{\text{GGU GGA GGCU}}{\text{PCCG AAG CCG}}$ (*9*)

		distance (Å)						distance (Å)			
atom 1	atom 2	lower	upper	NMR	model[e]	atom 1	atom 2	lower	upper	NMR	model[e]
G4H3′	G5H8	2.41	4.47	3.44	3.87	G14H3′	A15H8[c]	×[c]	×[c]	×[c]	4.50
G4H8	G5H8	2.36	6.00	3.37	5.69	G14H8	A15H8	×[c]	×[c]	×[c]	6.35
A5H1′	A6H8	2.30	4.27	3.29	3.55	A15H1′	A16H8	3.28	6.00	4.69	4.81
G14H1′	A5H2	–[f]	–[f]	–[f]	–[f]	G4H1′	A15H2	×[c]	×[c]	×[c]	8.77
G14H2′	G5H1	2.50	5.84	4.17	4.05	G4H2′	A15H2	×[c]	×[c]	×[c]	8.74
A5H2	A15H8	–[f]	–[f]	–[f]	–[f]	G5H8	A15H2	×[c]	×[c]	×[c]	10.38
G5H4′	A16H2	×[c]	×[c]	×[c]	5.21	A15H4′	A6H2	×[c]	×[c]	×[c]	5.49
A6H1′	A15H2	2.15	3.99	3.07	2.45	A16H1′	A5H2	–[f]	–[f]	–[f]	–[f]
A15H1′	A6H2	2.13	3.95	3.04	2.92	G5H1′	A16H2	×[c]	×[c]	×[c]	6.99
G14H1′	A15H8	3.27	6.00	4.67	4.44	G4H1′	A5H8	×[c]	×[c]	×[c]	6.35
A16H1′	A15H2	2.06	3.83	2.94	2.53	A6H1′	A5H2	–[f]	–[f]	–[f]	–[f]
A15H8	A16H8	×[c]	×[c]	×[c]	4.82	G5H8	A6H8	×[c]	×[c]	×[c]	7.16
A15H2′	A16H8	1.94	3.60	2.77	2.20	G5H2′	A6H8	×[c]	×[c]	×[c]	5.47
A15H3′	A16H8	2.39	4.44	3.42	2.87	G5H3′	A6H8	×[c]	×[c]	×[c]	5.14
G4H2′	G5H8	2.73	5.07	3.90	3.66	G14H2′	A15H8	2.02	3.74	2.88	2.50
G5H1′	G5H2′	2.23	4.14	3.19	2.99	A15H1′	A15H2′	2.30	4.26	3.28	2.73
G5H1′	G5H8	3.29	6.00	4.70	3.91	A15H1′	A15H8	3.15	5.85	4.50	3.74
G5H2′	G5H8	1.43	2.66	2.05	2.25	A15H2′	A16H1′	3.06	5.68	4.37	3.55
G5H3′	G5H8	2.46	4.57	3.52	4.05	A15H3′	A15H8	2.09	3.88	2.99	3.14
G5H1	G4H1	2.34	5.46	3.90	3.52	A15H2	A6H8	3.84	6.00	5.48	4.70
G5H1	G14H1	2.50	7.00		5.41	A15H2	A16H2	3.49	6.00	4.99	4.54

[a] All other distance restraints for both structural modelings are identical and are provided as Supporting Information. [b] Distances measured for the averaged structure of A6/A15/A16 followed by A6/A5/A16. [c] Cross-peaks not observed, which is consistent with the modeled structure. [d] Loosened upper bound to 4.50 Å for the structural modeling of A6/A5/A16. [e] The distances in the columns of models are measured from a representative structure of $\frac{\text{GGU GGA GGCU}}{\text{PCCG AAG CCG}}$ (*9*). [f] Not applicable.

The I15 duplex provides further evidence for two structures of the A5 duplex. Several features in the NMR spectra suggest that the loop conformation in the I15 duplex is most similar to the less populated conformation observed in the A5 duplex. NOEs observed in the I15 duplex that were weak in A5 duplex and not observed at all in a5, P5, or I5 duplexes include 15H2−G4H1′, A16H2−5H1′, 5H2−A16H1′, 5H2−A6H1′, and 15H2−5H8 (Figure 3). NOEs not observed in

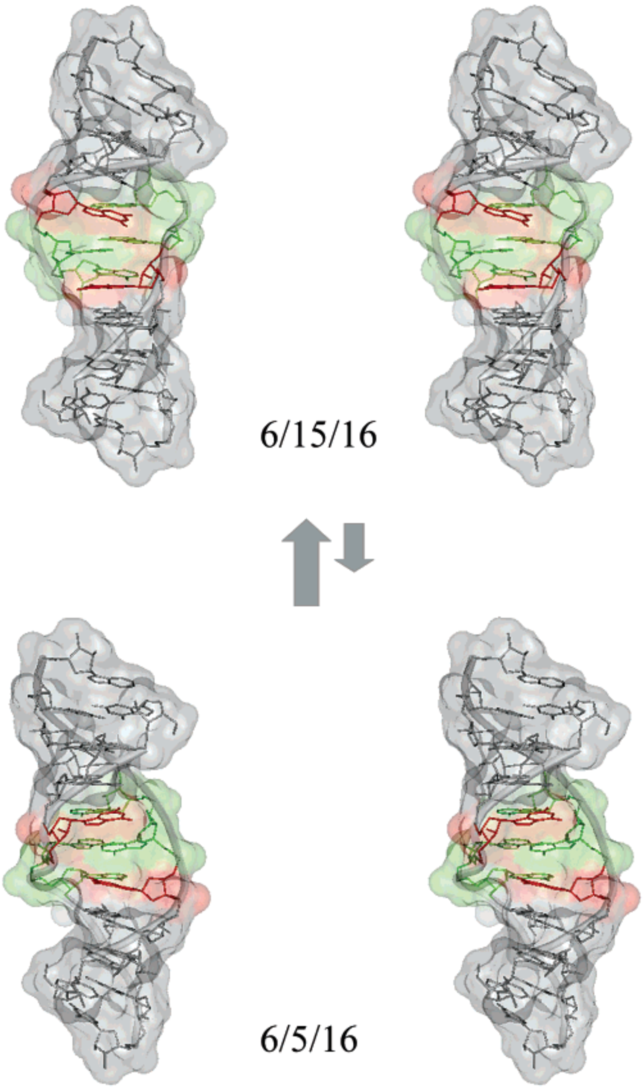

FIGURE 6: Major groove stereoviews of two alternating structures of $\begin{smallmatrix}\text{GGU}\ \underline{\text{GAA}}\ \text{GGCU}\\ \text{PCCG}\ \overline{\text{AAG}}\ \text{CCG}\end{smallmatrix}$ with A6/A15/A16 or A6/A5/A16 stacks in the minor groove. Two G's and four A's in the loop are colored red and green, respectively. Hydrogen and nonbridging oxygen atoms were omitted for clarity. The hydrogen bonding shown for the AA pair in the minor conformation is similar to that shown in Figure 1a, but a variety of hydrogen bonding patterns are seen in the ensemble of structures generated with the restraints from NMR data.

the I15 duplex but observed in $\begin{smallmatrix}\text{GGU}\ \underline{\text{GAA}}\ \text{GGCU}\\ \text{PCCG}\ \overline{\text{AAG}}\ \text{CCG}\end{smallmatrix}$ and all other duplexes include 5H2−G14H1′, A6H2−15H1′, 15H2−A6H1′, and 15H2−A16H1′. The relatively downfield chemical shift of I15H2 (8.55 ppm) and the relatively upfield shift of A5H2 (7.51 ppm) are consistent with those protons being at the A5-I15 base pairing edge and out in the minor groove, respectively, as discussed for the other duplexes (Table 4). A strong scalar coupling (8 Hz), I15H1′−H2′, and a large downfield ^{31}P shift of A16 (2.75 ppm) indicate a *C*2′-*endo* ribose conformation at I15 (*13*, *40*). In contrast, A5H1′−H2′ scalar coupling is much weaker and the ^{31}P shift of A6 is 1.10 ppm. These contrast with a5, P5, and I5 duplexes which show strong 5H1′−H2′ coupling (≥ 8 Hz) and weak 15H1′−H2′ coupling. Moreover, ^{31}P shifts of I5 duplex are 0.86 and 2.58 ppm for A16 and A6, respectively. Additionally, the G4H1′−H2′ and G14H1′−H2′ couplings are moderate and zero, respectively, in the I15 duplex, while they are zero and moderate, respectively, in a5, P5, and I5 duplexes. The moderate couplings probably indicate dynamic interconversion of sugar puckers. The ^{31}P chemical shifts for the A5 duplex are 1.41 and 2.20 ppm for A16 and A6, respectively, which suggest that the A5 duplex is more populated in a conformation similar to a5, P5, and I5 duplexes.

Table 4: Chemical Shifts (Parts per Million) and Full Widths (Hertz) at Half-Height of H2 Peaks of the Central Loop Residues, 5H2 and 15H2, in A5, a5, P5, I5, and I15 Duplexes at 30 °C in 80 mM NaCl[a]

	line width (Hz)			chemical shift (ppm)		
duplex	5H2	15H2	G2H8	5H2	15H2	G2H8
A5	15.0 (1.5)[b]	13.1 (1.0)	4.1 (1.0)	8.28	7.83	7.32
a5	12.0 (2.0)	6.2 (1.5)	3.8 (1.0)	8.44	7.57	7.31
P5	10.0 (1.0)	4.3 (1.0)	3.6 (0.5)	9.30	7.69	7.33
I5	7.1 (1.0)	4.1 (1.0)	4.4 (0.5)	8.69	7.55	7.35
I15	4.4 (1.0)	12.0 (1.5)	4.8 (0.5)	7.51	8.55	7.29

[a] G2H8 from the stem is included for reference. Error limits are listed in parentheses. [b] If the chemical shift of the A5H2 resonance in the A5 duplex differs by 0.87 ppm between major and minor conformations and the fraction of A5 duplexes in the major conformation ranges between 0.6 and 0.9, then a rough calculation (*57*, *58*) assuming an inherent line width of 4 Hz suggests the rate of exchange between the two conformations of the A5 duplex is between 20 000 and 65 000 s^{-1}.

As described above, there is a large chemical shift difference between the H2 resonances for the central purine-purine pairs, and the line widths of these resonances are consistent with two rapidly interconverting structures for the A5 duplex (Table 4). The line widths for A5H2 and A15H2 are both ∼14 Hz for the A5 duplex. In contrast, the H2 resonance not at the base pairing edge in a5, P5, I5, and I15 duplexes has an average line width of 4.75 ± 1.0 Hz, which is approximately the same as the average of the stem resonance G2H8 (4.1 ± 0.5 Hz). This is consistent with a model in which duplexes having a single conformation have a relatively narrow line width, while the A5 duplex resonances are broadened due to switching between two conformations. The base pairing edge H2 resonance in a5, P5, I5, and I15 duplexes has average line width of 10.2 ± 2.3 Hz. This may indicate that this residue is slightly less stable than the pairing partner on the other strand and/or that the chemical shift of the proton in this position is more sensitive to slight structural fluctuations.

Structural Statistics for $\begin{smallmatrix}\text{GGU}\ \underline{\text{GGA}}\ \text{GGCU}\\ \text{PCCA}\ \overline{\text{AAG}}\ \text{CCG}\end{smallmatrix}$ *and* $\begin{smallmatrix}\text{GGU}\ \underline{\text{GAA}}\ \text{GGCU}\\ \text{PCCG}\ \overline{\text{AAG}}\ \text{CCG}\end{smallmatrix}$. A total of 24 of 40 modeled structures (PDB entry 2DD1) were selected for analysis of $\begin{smallmatrix}\text{GGU}\ \underline{\text{GGA}}\ \text{GGCU}\\ \text{PCCA}\ \overline{\text{AAG}}\ \text{CCG}\end{smallmatrix}$. The average root-mean-square deviation of all selected structures from the average structure for all atoms is 0.80 ± 0.15 Å. No distance or dihedral angle restraint violations were greater than 0.2 Å or 2°, respectively. The average of the final energies at 300 K from the force field is −428.0 ± 4.9 kcal/mol.

A total of 27 of 40 modeled structures (PDB entry 2DD2) were selected for analysis of the A6/A15/A16 structure of $\begin{smallmatrix}\text{GGU}\ \underline{\text{GAA}}\ \text{GGCU}\\ \text{PCCG}\ \overline{\text{AAG}}\ \text{CCG}\end{smallmatrix}$. The average root-mean-square deviation of all selected structures from the average structure for all atoms is 0.69 ± 0.21 Å. Two distance restraint violations and one dihedral angle restraint violation were greater than 0.2 Å and 2°, respectively. The average of the final energies at 300 K from the force field is −430.6 ± 5.7 kcal/mol.

A total of 18 of 40 modeled structures (PDB entry 2DD3) were selected for analysis of the A6/A5/A16 structure of $\frac{\text{GGU GAA GGCU}}{\text{PCCG AAG CCG}}$. The average root-mean-square deviation of all selected structures from the average structure for all atoms is 0.97 ± 0.25 Å. One distance restraint violation and no dihedral angle restraint violations were greater than 0.2 Å and 2°, respectively. This minor conformation is less convergent than the other structures due to loosened restraints in the loop region. The average of the final energies at 300 K from the force field is −423.4 ± 6.3 kcal/mol.

Other Known AA Geometries Are Not Consistent with the NMR Data. As illustrated in Figure 1b, several nonsheared AA pairs have been observed in crystal and NMR structures (*6*, *43*−*49*).

A potential A-zipper motif, as seen in a $\frac{\text{C GAA G}}{\text{G AAG C}}$ DNA internal loop and the NMR structure of a $\frac{\text{C UAA G}}{\text{G UA U}}$ RNA tetraloop receptor (*50*, *51*), which would place A5 and A15 in $\frac{\text{GGU GAA GGCU}}{\text{PCCG AAG CCG}}$ stacking on each other, is ruled out because no A5H2−A15H1′ or A15H2−A5H1′ NOE is observed in $\frac{\text{GGU GAA GGCU}}{\text{PCCG AAG CCG}}$.

Both AH2 protons are exposed in the minor groove for a trans Watson−Crick/Hoogsteen A-A pair (Figure 1b) (*6*, *43*). This conformation is ruled out for the middle AA pair in $\frac{\text{U GAA G}}{\text{G AAG C}}$ because it would not give the observed cross-strand G4H1′/H2′−A15H2 and G14H1′/H2′−A5H2 cross-peaks (Figure 3b, Table 3, and Table S2 of the Supporting Information).

The cis Watson−Crick/Watson−Crick A-A conformation (*6*, *44*, *45*), with both AH2 protons exposed in the minor groove, is ruled out for the middle AA in $\frac{\text{U GAA G}}{\text{G AAG C}}$ because it would not give all the observed G14H1′/H2′−5H2, A6H1′−A15H2, A15H1′−A6H2, and A16H1′−A15H2 (denoted with yellow circles and ovals) and G4H1′/H2′−A15H2, A16H1′−A5H2, A5H1′−A16H2, and A6H1′−A5H2 (denoted with green circles and ovals) cross-peaks (Figure 3b).

A trans Hoogsteen/Hoogsteen A-A pair (Figure 1b) (*6*, *46*−*49*) is ruled out for the middle AA in $\frac{\text{U GAA G}}{\text{G AAG C}}$ because there is no indication of a syn glycosidic conformation as evidenced by A5H1′−A5H8 and A15H1′−A15H8 cross-peaks and because the 14H1′/H2′−5H2 and 4H1′/H2′−15H2 cross-peaks seen in $\frac{\text{U GAA G}}{\text{G AAG C}}$ are not expected for a trans Hoogsteen/Hoogsteen A-A pair with two AH2 protons exposed in the minor and major groove, respectively.

DISCUSSION

Understanding relationships among sequence, energetics, structure, dynamics, and function can facilitate rapid extraction of the information encoded in the constantly expanding databases of RNA sequences. The internal loop is a common RNA motif where such relationships are not fully understood (*9*, *12*, *13*, *52*−*54*). A detailed understanding of interactions such as hydrogen bonding and base stacking in internal loops will allow prediction of the contributions of internal loops to RNA folding and function.

Three Consecutive Sheared GA Pairs in $\frac{\text{U GGA G}}{\text{A AAG C}}$. The previous NMR structure of $\frac{\text{GGU GGA GGCU}}{\text{PCCG AAG CCG}}$ reveals three consecutive sheared GA pairs in the unusually stable internal loop (*9*). Formation of three consecutive sheared GA pairs in $\frac{\text{U GGA G}}{\text{A AAG C}}$ (−2.27 kcal/mol) (Figures 4 and 5a), as in $\frac{\text{U GGA G}}{\text{G AAG C}}$ (−2.62 kcal/mol), is consistent with the thermodynamic stabilities (Figure 2 and Table 2) and the occurrences of both loops in helix 41a of small subunit rRNA (*52*). (Throughout the paper, the values in parentheses after the duplex are the measured free energy at 37 °C for loop formation in 1 M NaCl unless otherwise noted.)

In contrast to the NMR structures, helix 68 of the crystal structure of *D. radiodurans* large subunit rRNA contains a $\frac{\text{U }\mathbf{G}\text{GA G}}{\text{A }\mathbf{A}\text{AG C}}$ loop that has only one sheared GA pair (shown in bold) (*14*). The major difference is that the corresponding G5 and A15 bases are shifted, opposite to a sheared GA pair, to the minor and major groove, respectively. This results in the loss of hydrogen bonding and in a base stacking pattern equivalent to an A6/G5/A16 pattern in the minor groove, instead of the A6/A15/A16 stacking pattern found in the NMR structure (Figure 5a). Several critical NOEs define the A6/A15/A16 stacking pattern in the NMR structure with three consecutive sheared GA pairs, e.g., A15H2−A6H1′, A15H1′−A6H2, and A15H2−A16H1′ (Figure 3a and Table S1 of the Supporting Information). The distances between the protons in each pair exceed 5 Å in the crystal structure (PDB entry 1NKW) when hydrogens are added (Table S1 of the Supporting Information). Interestingly, the A6/G5/A16 stacking pattern in the crystal structure (Figure 5a) is similar to the A6/A5/A16 stacking pattern determined for the minor NMR structure of $\frac{\text{GGU GAA GGCU}}{\text{PCCG AAG CCG}}$ (Figure 5b), although fewer hydrogen bonds are formed in the crystal structure (*14*).

There are several differences between the environments of the $\frac{\text{U GGA G}}{\text{A AAG C}}$ loop in the crystal and in NMR buffer. The crystals were grown from ribosomal subunits in 10 mM $MgCl_2$, 60 mM NH_4Cl, 5 mM KCl, and 10 mM HEPES (pH 7.8) (*14*). The NMR buffer consists of 80 mM NaCl and 10 mM sodium phosphate (pD 6.8). It would be surprising, however, if Mg^{2+} shifted the local structure. The thermodynamics of the 3GA duplex (Figure 2) were essentially the same in 1 M NaCl and in 10 mM $MgCl_2$ and 150 mM KCl (*9*). It is quite possible, however, that other interactions in the ribosomal subunit crystal are strong enough to break hydrogen bonds and rearrange stacking. There is no tertiary interaction with or protein binding to the loop $\frac{\text{U }\mathbf{G}\text{GA G}}{\text{A }\mathbf{A}\text{AG C}}$ in the crystal, but the loop, $\frac{\text{G UCAAG U}}{\text{C GAAGU A}}$, which is directly 5′ to the UA closing pair of $\frac{\text{U }\mathbf{G}\text{GA G}}{\text{A }\mathbf{A}\text{AG C}}$, has tertiary interactions with helix 75 via consecutive A−minor interactions. Similar A−minor tertiary interactions are observed in the crystal structures of the large ribosomal subunits of *Haloarcula marismortui* and *Escherichia coli* between helix 75 and helix 68 (*55*, *56*). While long-range effects may affect local structure, it may also be difficult to determine such fine details in a large crystal refined to 3.1 Å.

Tandem sheared GA pairs closed by UA Watson−Crick pairs have been reported in an NMR structure of $\frac{\text{GGGCU GA AGCCU}}{\text{UCCGA AG UCGGG}}$ (*40*). The sugars of G's in $\frac{\text{U GA A}}{\text{A AG U}}$ are in a *C*2′-*endo* conformation. In the loop, $\frac{\text{U GGA G}}{\text{A AAG C}}$, of the A17 duplex, G4 and G5 have *C*3′-*endo* and *C*2′-*endo* sugar puckers, respectively, and G14 is populated in both conformations as evidenced by the TOCSY spectrum (Figure S1a of the Supporting Information). Evidently, a *C*2′-*endo* sugar pucker for G is not required for formation of a sheared GA pair in a $\frac{\text{U G}}{\text{A A}}$ motif.

Two Alternating Structures for $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GAA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$. The sheared AA pair in $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GAA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ is rapidly exchanging between alternative conformations (Figures 5b and 6). This exchange is consistent with an intrinsically flexible AA pair with fewer hydrogen bonds than a GA pair. The pseudosymmetry of the dynamic AA pair allows an estimate of the lower limit for the exchange rate. The a5 duplex has only one conformation, and it is the same as the major conformation of the A5 duplex. The chemical shifts for a5H2 and A15H2 in this conformation are 8.44 and 7.57 ppm, respectively. With the assumption that the minor conformation of the A5 duplex would result in an A5H2 chemical shift of 7.57 ppm, the lower limit for the exchange rate is estimated to be $0.87 \times 500 = 435\ \text{s}^{-1}$. A calculation based on line widths (*57*, *58*) of H2 resonances suggests an even faster exchange rate [between 20 000 and 65 000 s^{-1} (Table 4)] for estimates of the major conformation ranging from 90 to 60% of the population. Fast exchange has also been detected between syn and anti G's in a single GG pair in an RNA duplex (*17*). A recent theoretical study has provided insight into possible mechanisms for such rapid exchange in the absence of duplex dissociation (*59*).

In contrast to $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GAA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$, duplexes a5, P5, I5, and I15 all have a single predominant conformation. As indicated by NOE patterns (Figure 3), there is little structural change for a5, P5, and I5 duplexes relative to $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GGA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (*9*) and to the major conformation of the A5 duplex. NMR data for the I15 duplex suggest that its structure differs from that of a5, P5, and I5 duplexes but resembles the minor structure of the A5 duplex.

The duplex with a deoxyadenosine (a5), $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GaA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−0.28 kcal/mol), has a single structure with a trans Hoogsteen/sugar edge A15-a5 pair (Figure 1). This is consistent with the fact that G5 is predominantly in a *C*2′-*endo* sugar pucker conformation with a trans Hoogsteen/sugar edge A15-G5 pair in $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GGA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−2.62 kcal/mol) (*9*). Presumably, the a5 duplex has a single structure because the deoxy sugar favors *C*2′-*endo* sugar pucker which thus favors a single conformation similar to that of $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GGA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (*9*). This is also consistent with the observation that the deoxy g5 substitution in $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GgA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−3.04 kcal/mol) enhances loop stability by 0.42 kcal/mol despite the loss of hydrogen bonds from G5 (2′-hydroxyl) to G14 (imino/amino). The opposite change in thermodynamic stability is observed for $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GgA}}{\text{AAGAAA}}{}^{\text{GGCU}}_{\text{CCG}}$ (0.67 kcal/mol) compared with $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GGA}}{\text{AAGAAA}}{}^{\text{GGCU}}_{\text{CCG}}$ (0.17 kcal/mol). Perhaps the greater flexibility of a 3×6 loop negates the necessity of a *C*2′-*endo* sugar.

Previous NMR studies showed no orientation exchange for the tandem sheared AA pairs of $^{\text{C}}_{\text{G}}\frac{\text{AA}}{\text{AA}}{}^{\text{G}}_{\text{C}}$ (*7*), presumably because switching the orientation would result in making the backbone too narrow for the adjacent Watson−Crick pair (*60*, *61*). Evidently, switching the side where backbone narrowing occurs is not a problem when an AA pair is flanked by sheared GA pairs.

It has been pointed out that $\frac{\text{GA}}{\text{AA}}$ might be a potential groove binding and/or intercalation site (*43*). The alternating sheared AA pair in $^{\text{U}}_{\text{G}}\frac{\text{GAA}}{\text{AAG}}{}^{\text{G}}_{\text{C}}$ could potentially serve as a switch between different binding partners for dynamic functions because the smooth N1-C2-N3 edge of either A5 or A15 is presented differently in the minor groove in alternative orientations (Figure 5b).

Energetics of Molecular Recognition. The consistency of structures for the 3GA, I5, P5, and a5 duplexes (Figures 2 and 3) provides models for studying the interactions determining the energetics of a 3×3 loop with three sheared pairs. The P5 duplex $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GPA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−0.53 kcal/mol) is thermodynamically similar to the a5 duplex $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GaA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−0.28 kcal/mol) and to $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GAA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−0.48 kcal/mol) (Table 2). This is in agreement with the formation of a sheared PA pair (trans Hoogsteen/sugar edge A15-P5) without the loss of hydrogen bonds compared with a sheared AA pair (Figure 1).

The I5 duplex $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GIA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−1.22 kcal/mol) is 1.40 kcal/mol less stable than $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GGA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−2.62 kcal/mol). Similar destabilizations of 1.74 and 2.00 kcal/mol are observed with the G to I substitutions in $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GAA}}{\text{AIG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−0.88 kcal/mol) and $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GIA}}{\text{AAGAAA}}{}^{\text{GGCU}}_{\text{CCG}}$ (2.17 kcal/mol), respectively (Table 2). This is presumably primarily due to the loss of hydrogen bonds of G5 amino to A15N7 and to the A15 nonbridging oxygen in the IA pair compared to GA (Figure 1). The free energy of ∼1.5 kcal/mol attributed to two hydrogen bonds is a lower limit because subtle rearrangement of three-dimensional structure is expected upon inosine substitution, and this can strengthen the remaining hydrogen bonds (*62*).

Interestingly, $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GIA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−1.22 kcal/mol) is more stable than $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GPA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−0.53 kcal/mol), $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GAA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−0.48 kcal/mol), and $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GaA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−0.28 kcal/mol) (Table 2 and Figure 2), even though the number of base−base hydrogen bonds is expected to be the same (Figure 1). A water-mediated hydrogen bond between the G imino proton and the nonbridging oxygen of A was predicted (*15*) and observed in a crystal structure of a GNRA tetraloop (*63*). A similar water-mediated hydrogen bond might exist between I5H1 and an A15 nonbridging oxygen in the I5 duplex $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GIA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$. This water-mediated hydrogen bond might explain the extra stability of the I5 duplex relative to those of the P5, A5, and a5 duplexes (Figure 2).

The D5 (2,6-diaminopurine) duplex, $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GDA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−1.45 kcal/mol), is ∼1.2 kcal/mol less stable than $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GGA}}{\text{AAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−2.62 kcal/mol) (Table 2 and Figure 2), even though no base−base hydrogen bonds are lost (Figure 1). The destabilizing effect upon substitution of D5 for G5 may also be due to loss of the proposed water-mediated hydrogen bond between G5H1 and an A15 nonbridging oxygen (*15*, *63*). It is also possible that the 2-amino group of G is a better hydrogen bond donor than that of D (2,6-diaminopurine) because of relatively larger positive partial charges on the G amino hydrogens (*64*). Greater destabilization of 2.3 kcal/mol is observed for $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GDA}}{\text{AAGAAA}}{}^{\text{GGCU}}_{\text{CCG}}$ (2.47 kcal/mol) compared with $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GGA}}{\text{AAGAAA}}{}^{\text{GGCU}}_{\text{CCG}}$ (0.17 kcal/mol). Perhaps the greater flexibility of the size-asymmetric loop allows binding of a water molecule in a GA pair to be more favorable. In contrast, substitution of D for A16 to give $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GGA}}{\text{DAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−3.32 kcal/mol) and $^{\text{GGU}}_{\text{PCCG}}\frac{\text{GAA}}{\text{DAG}}{}^{\text{GGCU}}_{\text{CCG}}$ (−0.78 kcal/mol) stabilizes the loop by −0.70 and −0.30 kcal/mol, respectively, even though the 6-amino group on D (2,6-diaminopurine) has essentially the same partial charges as A (*64*). The extra amino group on D possibly allows better stacking and/or extra hydrogen bonding to the backbone.

A 2′-*O*-methyl substitution favors the *C*3′-*endo* sugar conformation (*65*−*67*). The 2′-*O*-methyl A15 substitution in ${}^{\mathrm{GGU}}_{\mathrm{PCCG}}\frac{\mathrm{GGA}}{\mathrm{AMG}}{}^{\mathrm{GGCU}}_{\mathrm{CCG}}$ (−3.83 kcal/mol) and ${}^{\mathrm{GGU}}_{\mathrm{PCCG}}\frac{\mathrm{GAA}}{\mathrm{AMG}}{}^{\mathrm{GGCU}}_{\mathrm{CCG}}$ (−0.83 kcal/mol) stabilizes the duplexes by −1.21 and −0.35 kcal/mol relative to ${}^{\mathrm{GGU}}_{\mathrm{PCCG}}\frac{\mathrm{GGA}}{\mathrm{AAG}}{}^{\mathrm{GGCU}}_{\mathrm{CCG}}$ (−2.62 kcal/mol) and ${}^{\mathrm{GGU}}_{\mathrm{PCCG}}\frac{\mathrm{GAA}}{\mathrm{AAG}}{}^{\mathrm{GGCU}}_{\mathrm{CCG}}$ (−0.48 kcal/mol), respectively (Table 2). Part of the reason for the smaller effect in the ${}^{\mathrm{U}}_{\mathrm{G}}\frac{\mathrm{GAA}}{\mathrm{AAG}}{}^{\mathrm{G}}_{\mathrm{C}}$ loop may be that the 2′-*O*-methyl substitution limits the loop to a single conformation. If the two conformations of the natural loop have equal concentrations, then the dynamics would favor loop formation [$\Delta G^\circ_{37} = -T\Delta S = -310(1.987)\ln 2 = -0.4$ kcal/mol].

$\frac{GA}{AA}$ *Motif within Other Internal Loops*. Adjacent sheared GA and AA pairs are also found in the ${}^{\mathrm{C}}_{\mathrm{G}}\frac{\mathrm{G\mathbf{A}C}}{\mathrm{AAC}}{}^{\mathrm{C}}_{\mathrm{G}}$ and ${}^{\mathrm{C}}_{\mathrm{G}}\frac{\mathrm{G\mathbf{A}C}}{\mathrm{AAC}}{}^{\mathrm{G}}_{\mathrm{C}}$ loops in helix 89 of the large ribosomal subunits of *H. marismortui* (*45*) and *E. coli* (*56*), respectively. In this case, the bold A pairs with its Hoogsteen edge and the CC pair is in a cis Watson−Crick bifurcated conformation (*45*, *56*) according to Leontis−Westhof nomenclature (*6*). The same loop sequence occurs in helix 41a of *E. coli* 16S rRNA with a similar three-dimensional conformation (*56*).

Other $\frac{\mathrm{GA}}{\mathrm{AA}}$ motifs have nonsheared pairs. Only face-to-face (cis Watson−Crick/Watson−Crick pair) purine-purine pairs (Figure 1) form in the loop ${}^{\mathrm{G}}_{\mathrm{U}}\frac{\mathrm{GAA}}{\mathrm{AAG}}{}^{\mathrm{C}}_{\mathrm{G}}$ in helix 23 of the crystal structure of *Thermus thermophilus* 16S rRNA (*44*), which is consistent with previous NMR studies which showed that imino GA (face-to-face, cis Watson−Crick/Watson−Crick A-G pair) is favored in the ${}^{\mathrm{G}}_{\mathrm{C}}\frac{\mathrm{GA}}{\mathrm{AG}}{}^{\mathrm{C}}_{\mathrm{G}}$ motif (*68*). The crystal structure of a symmetric 4 × 4 loop, ${}^{\mathrm{C}}_{\mathrm{G}}\frac{\mathrm{G\mathbf{A}AA}}{\mathrm{AA\mathbf{A}G}}{}^{\mathrm{G}}_{\mathrm{C}}$, shows little base overlap for the $\frac{\mathrm{GA}}{\mathrm{AA}}$ nearest neighbors with sheared GA (trans Hoogsteen/sugar edge A-G) pairs and trans Watson−Crick/Hoogsteen A-**A** pairs (*43*). [In the trans Watson−Crick/Hoogsteen A-**A** pairs, the A paired with Hoogsteen edge is bold (Figure 1).] Evidently, $\frac{\mathrm{GA}}{\mathrm{AA}}$ is an intrinsically flexible structure within size-symmetric internal loops.

The free energy increment for ${}^{\mathrm{C}}_{\mathrm{G}}\frac{\mathrm{G\mathbf{A}AA}}{\mathrm{AA\mathbf{A}G}}{}^{\mathrm{G}}_{\mathrm{C}}$ at 37 °C in 1 M NaCl is 0.96 kcal/mol as calculated from the measurement of the duplex ${}^{\mathrm{CGC}}_{\mathrm{GCG}}\frac{\mathrm{GAAA}}{\mathrm{AAAG}}{}^{\mathrm{GGC}}_{\mathrm{CCG}}$ (*13*). In contrast, internal loops with consecutive GA pairs are very stable, e.g., ${}^{\mathrm{U}}_{\mathrm{G}}\frac{\mathrm{GGAA}}{\mathrm{AAGG}}{}^{\mathrm{G}}_{\mathrm{C}}$ (−4.27 kcal/mol) (*13*), which is consistent with extensive stacking and hydrogen bonding as observed in the crystal structure of the loop ${}^{\mathrm{C}}_{\mathrm{G}}\frac{\mathrm{GGAA}}{\mathrm{AAGG}}{}^{\mathrm{G}}_{\mathrm{C}}$ with four sheared GA pairs (*69*). Previous thermodynamic studies showed that the destabilizing 2 × 2 loop ${}^{\mathrm{C}}_{\mathrm{G}}\frac{\mathrm{AA}}{\mathrm{AA}}{}^{\mathrm{G}}_{\mathrm{C}}$ (1.2 kcal/mol) with two sheared AA pairs has base pairing and stacking geometries similar to those of but fewer hydrogen bonds than ${}^{\mathrm{C}}_{\mathrm{G}}\frac{\mathrm{GA}}{\mathrm{AG}}{}^{\mathrm{G}}_{\mathrm{C}}$ (−0.7 kcal/mol) with two sheared GA pairs (*2*, *5*, *7*, *8*). Evidently, the thermodynamic and structural effects of replacing a GA pair with an AA pair are context-dependent.

A recently proposed "reverse kink-turn" motif involves a size-asymmetric 2 × 5 internal loop, ${}^{\mathrm{G}}_{\mathrm{C}}\frac{\mathrm{G\mathbf{A}}}{\mathrm{AAACA}}{}^{\mathrm{G}}_{\mathrm{C}}$, with a sheared GA followed by a symmetric AA pair (trans Hoogsteen/Hoogsteen A-A pair) (Figure 1b) (*49*). Such a conformation is also observed in some loop E motifs, $\frac{\mathrm{G\mathbf{A}\ GUA}}{\mathrm{AA\quad AR}}$ (*45*, *47*, *48*). The glycosidic bond of the A (in bold) 3′ to the G of the sheared GA pair is in a syn conformation. Thus, it might facilitate the packing between two stems via major grooves, with the smooth N1-C2-N3 edge of the bold A flipped to the major groove. Different detailed structures of an AA pair adjacent to a sheared GA pair in $\frac{\mathrm{G\mathbf{A}}}{\mathrm{AA}}$ nearest neighbors with the bold A paired on its Hoogsteen side are also observed in kink-turn motifs (*70*) within internal loops, such as kt-11 (trans Watson−Crick/Hoogsteen A-**A** pair), and multibranch loops such as kt 94/99 (trans Hoogsteen/Sugar edge **A**-A pair), kt 4/5 (trans Hoogsteen/Hoogsteen **A**-**<u>A</u>**, with the **<u>A</u>** in a syn glycosidic conformation). These kink turns facilitate local and long-range tertiary interactions (*70*). In these cases, the A 3′ to the G of a sheared GA pair prefers to base pair with its Hoogsteen edge. Evidently, the $\frac{\mathrm{G\mathbf{A}}}{\mathrm{AA}}$ nearest neighbor is intrinsically flexible compared with the motif of consecutive GA pairs (*13*) in both size-symmetric and -asymmetric internal loops.

ACKNOWLEDGMENT

We thank Mr. B. Tolbert for help with PyMOL and critical reading of the manuscript and Prof. R. Kierzek and Dr. E. Kierzek for discussions on synthesis of oligonucleotides with modified nucleotides.

SUPPORTING INFORMATION AVAILABLE

Tables listing chemical shift assignments, tables of NMR distance restraints, and figures of TOCSY, one-dimensional proton spectra (9−14.5 ppm) and a two-dimensional NOESY spectrum at 1 M NaCl. This material is available free of charge via the Internet at http://pubs.acs.org.

NOTE ADDED IN PROOF

A recent crystal structure of a ribonuclease P RNA reveals an internal loop, ${}^{\mathrm{C}}_{\mathrm{G}}\frac{\mathrm{GGA}}{\mathrm{AAG}}{}^{\mathrm{G}}_{\mathrm{C}}$, with three consecutive sheared GA pairs in P15.1 and P19 (*71*).

REFERENCES

1. 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.
2. 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.
3. 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.
4. Li, Y., Zon, G., and Wilson, W. D. (1991) NMR and molecular modeling evidence for a GA mismatch base pair in a purine rich DNA duplex, *Proc. Natl. Acad. Sci. U.S.A. 88*, 26−30.
5. SantaLucia, J., Jr., and Turner, D. H. (1993) Structure of $(\mathrm{rGGC\underline{GA}GCC})_2$ in solution from NMR and restrained molecular dynamics, *Biochemistry 32*, 12612−12623.
6. 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.
7. Znosko, B. M., Burkard, M. E., Schroeder, S. J., Krugh, T. R., and Turner, D. H. (2002) Sheared $\mathrm{A_{anti}{\cdot}A_{anti}}$ base pairs in a destabilizing 2 × 2 internal loop: The NMR structure of $5'(\mathrm{rGGC\underline{AA}GCCU})_2$, *Biochemistry 41*, 14969−14977.
8. 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. D60*, 829−835.

9. Chen, G., Znosko, B. M., Kennedy, S. D., Krugh, T. R., and Turner, D. H. (2005) Solution structure of an RNA internal loop with three consecutive sheared GA pairs, *Biochemistry 44*, 2845−2856.
10. SantaLucia, J., Kierzek, R., and Turner, D. H. (1991) Functional group substitutions as probes of hydrogen bonding between GA mismatches in RNA internal loops, *J. Am. Chem. Soc. 113*, 4313−4322.
11. 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.
12. Chen, G., Znosko, B. M., Jiao, X. Q., and Turner, D. H. (2004) Factors affecting thermodynamic stabilities of RNA 3 × 3 internal loops, *Biochemistry 43*, 12865−12876.
13. Chen, G., and Turner, D. H. (2006) Consecutive GA pairs stabilize medium size RNA internal loops, *Biochemistry 45*, 4025−4043.
14. 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.
15. SantaLucia, J., Kierzek, R., and Turner, D. H. (1992) Context dependence of hydrogen bond free energy revealed by substitutions in an RNA hairpin, *Science 256*, 217−219.
16. 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.
17. Burkard, M. E., and Turner, D. H. (2000) NMR structures of r(GCA$\underline{G}$GC$\underline{G}$UGC)$_2$ and determinants of stability for single guanosine-guanosine base pairs, *Biochemistry 39*, 11748−11762.
18. 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′(rGGC$\underline{AA}$GCCU)$_2$, *Biochemistry 41*, 14978−14987.
19. 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.
20. Broda, M., Kierzek, E., Gdaniec, Z., Kulinski, T., and Kierzek, R. (2005) Thermodynamic stability of RNA structures formed by CNG trinucleotide repeats. Implication for prediction of RNA structure, *Biochemistry 44*, 10873−10882.
21. Moody, E. M., Feerrar, J. C., and Bevilacqua, P. C. (2004) Evidence that folding of an RNA tetraloop hairpin is less cooperative than its DNA counterpart, *Biochemistry 43*, 7992−7998.
22. 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.
23. Szewczak, L. B. W., DeGregorio, S. J., Strobel, S. A., and Steitz, J. A. (2002) Exclusive interaction of the 15.5 kD protein with the terminal box C/D motif of a methylation guide snoRNP, *Chem. Biol. 9*, 1095−1107.
24. Blount, K. E., and Uhlenbeck, O. C. (2005) The structure−function dilemma of the hammerhead ribozyme, *Annu. Rev. Biophys. Biomol. Struct. 34*, 415−440.
25. 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.
26. 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.
27. Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., pp 589−595, CRC Press, Cleveland, OH.
28. Richards, E. G. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., pp 596−603, CRC Press, Cleveland, OH.
29. McDowell, J. A., and Turner, D. H. (1996) Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: Solution structure of (rGAG$\underline{GU}$CUC)$_2$ by two-dimensional NMR and simulated annealing, *Biochemistry 35*, 14077−14089.
30. 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.
31. 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.
32. 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.
33. Lukavsky, P. J., and Puglisi, J. D. (2001) RNAPack: An integrated NMR approach to RNA structure determination, *Methods 25*, 316−332.
34. 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.
35. Ding, H. Q., Karasawa, N., and Goddard, W. A. (1992) Atomic level simulations on a million particles: The cell multipole method for coulomb and London nonbond interactions, *J. Chem. Phys. 97*, 4309−4315.
36. DeLano, W. L. (2002) *The PyMOL User's Manual*, DeLano Scientific, San Carlos, CA.
37. 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.
38. Varani, G., and Tinoco, I. (1991) RNA structure and NMR spectroscopy, *Q. Rev. Biophys. 24*, 479−532.
39. Varani, G., Aboulela, F., and Allain, F. H. T. (1996) NMR investigation of RNA structure, *Prog. Nucl. Magn. Reson. Spectrosc. 29*, 51−127.
40. 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.
41. Uesugi, S., Oda, Y., Ikehara, M., Kawase, Y., and Ohtsuka, E. (1987) Identification of I-A mismatch base pairing structure in DNA, *J. Biol. Chem. 262*, 6965−6968.
42. SantaLucia, J., Jr. (1991) The Role of Hydrogen Bonding in the Thermodynamics and Structure of Mismatches in RNA Oligonucleotides, Ph.D. Thesis, University of Rochester, Rochester, NY.
43. Baeyens, K. J., DeBondt, H. L., Pardi, A., and Holbrook, S. R. (1996) A curved RNA helix incorporating an internal loop with GA and AA non-Watson−Crick base pairing, *Proc. Natl. Acad. Sci. U.S.A. 93*, 12851−12855.
44. 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.
45. 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 angstrom resolution, *Science 289*, 905−920.
46. Correll, C. C., Munishkin, A., Chan, Y. L., Ren, Z., Wool, I. G., and Steitz, T. A. (1998) Crystal structure of the ribosomal RNA domain essential for binding elongation factors, *Proc. Natl. Acad. Sci. U.S.A. 95*, 13436−13441.
47. Lukavsky, P. J., Otto, G. A., Lancaster, A. M., Sarnow, P., and Puglisi, J. D. (2000) Structures of two RNA domains essential for hepatitis C virus internal ribosome entry site function, *Nat. Struct. Biol. 7*, 1105−1110.
48. Krasilnikov, A. S., Yang, X., Pan, T., and Mondragon, A. (2003) Crystal structure of the specificity domain of ribonuclease P, *Nature 421*, 760−764.
49. Strobel, S. A., Adams, P. L., Stahley, M. R., and Wang, J. M. (2004) RNA kink turns to the left and to the right, *RNA 10*, 1852−1854.
50. Chou, S. H., Zhu, L., and Reid, B. R. (1997) Sheared purine·purine pairing in biology, *J. Mol. Biol. 267*, 1055−1067.
51. Butcher, S. E., Dieckmann, T., and Feigon, J. (1997) Solution structure of a GAAA tetraloop receptor RNA, *EMBO J. 16*, 7490−7499.

52. Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Muller, K. M., Pande, N., Shang, Z., Yu, N., and Gutell, R. R. (2002) The Comparative RNA Web (CRW) Site: An online database of comparative sequence and structure information for ribosomal, intron, and other RNAs, *BMC Bioinf. 3*, 2.
53. Mathews, D. H., Schroeder, S. J., Turner, D. H., and Zuker, M. (2005) in *The RNA World* (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) 3rd ed., pp 631−657, Cold Spring Harbor Laboratory Press, Woodbury, NY.
54. Schroeder, S. J., Burkard, M. E., and Turner, D. H. (1999) The energetics of small internal loops in RNA, *Biopolymers 52*, 157−167.
55. 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.
56. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. D. (2005) Structures of the bacterial ribosome at 3.5 angstrom resolution, *Science 310*, 827−834.
57. Gutowsky, H. S., and Holm, C. H. (1956) Rate processes and nuclear magnetic resonance spectra 2. Hindered internal rotation of amides, *J. Chem. Phys. 25*, 1228−1234.
58. Legault, P., and Pardi, A. (1997) Unusual dynamics and pK_a shift at the active site of a lead-dependent ribozyme, *J. Am. Chem. Soc. 119*, 6621−6628.
59. Mathews, D. H., and Case, D. A. (2006) Nudged elastic band calculation of minimal energy paths for the conformational change of a GG noncanonical pair, *J. Mol. Biol. 357*, 1683−1693.
60. 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.
61. Wu, M., SantaLucia, J., Jr., and Turner, D. H. (1997) Solution structure of (rGGC$\underline{\text{AG}}$GCC)$_2$ by two-dimensional NMR and the iterative relaxation matrix approach, *Biochemistry 36*, 4449−4460.
62. 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.
63. Batey, R. T., and Doudna, J. A. (2002) Structural and energetic analysis of metal ions essential to SRP signal recognition domain assembly, *Biochemistry 41*, 11703−11710.
64. Ornstein, R. L., and Fresco, J. R. (1983) Correlation of crystallographically determined and computationally predicted hydrogen bonded pairing configurations of nucleic acid bases, *Proc. Natl. Acad. Sci. U.S.A. 80*, 5171−5175.
65. Adamiak, D. A., Milecki, J., Popenda, M., Adamiak, R. W., Dauter, Z., and Rypniewski, W. R. (1997) Crystal structure of 2′-O-Me-(CGCGCG)$_2$, an RNA duplex at 1.30 angstrom resolution. Hydration pattern of 2′-O-methylated RNA, *Nucleic Acids Res. 25*, 4599−4607.
66. Popenda, M., Biala, E., Milecki, J., and Adamiak, R. W. (1997) Solution structure of RNA duplexes containing alternating CG base pairs: NMR study of r(CGCGCG)$_2$ and 2′-O-Me(CGCGCG)$_2$ under low salt conditions, *Nucleic Acids Res. 25*, 4589−4598.
67. Kierzek, E., Ciesielska, A., Pasternak, K., Mathews, D. H., Turner, D. H., and Kierzek, R. (2005) The influence of locked nucleic acid residues on the thermodynamic properties of 2′-O-methyl RNA/RNA heteroduplexes, *Nucleic Acids Res. 33*, 5082−5093.
68. 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.
69. 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.
70. Lescoute, A., Leontis, N. B., Massire, C., and Westhof, E. (2005) Recurrent structural RNA motifs, isostericity matrices and sequence alignments, *Nucleic Acids Res. 33*, 2395−2409.
71. Kazantsev, A. V., Krivenko, A. A., Harrington, D. J., Holbrook, S. R., Adams, P. D., and Pace, N. R. (2005) Crystal structure of a bacterial ribonuclease P RNA, *Proc. Natl. Acad. Sci. U.S.A. 102*, 13392−13397.

BI0524464