Stacking Effects on Local Structure in RNA: Changes in the Structure of Tandem GA Pairs when Flanking GC Pairs Are Replaced by isoG−isoC Pairs†

Gang Chen,‡ Ryszard Kierzek,§ Ilyas Yildirim,|| Thomas R. Krugh,‡ Douglas H. Turner,*,‡,# and Scott D. Kennedy⊥

Departments of Chemistry and Physics, University of Rochester, Rochester, New York 14627, Center for Pediatric Biomedical Research and Departments of Pediatrics and Biochemistry and Biophysics, School of Medicine and Dentistry, University of Rochester, Rochester, New York 14642, and Institute of Bioorganic Chemistry, Polish Academy of Sciences, 60-714 Poznan, Noskowskiego 12/14, Poland

Received: December 19, 2006; In Final Form: February 19, 2007

The Watson−Crick-like isoG−isoC (iGiC) pair, with the amino and carbonyl groups transposed relative to the Watson−Crick GC pair, provides an expanded alphabet for understanding interactions that shape nucleic acid structure. Here, thermodynamic stabilities of tandem GA pairs flanked by iGiC pairs are reported along with the NMR structures of the RNA self-complementary duplexes (GCiGGAiCGCA)$_2$ and (GGiCGAiGCCA)$_2$. A sheared GA pairing forms in (GCiGGAiCGCA)$_2$, and an imino GA pairing forms in (GGiCGAiGCCA)$_2$. The structures contrast with the formation of tandem imino and sheared GA pairs flanked by GC pairs in the RNA self-complementary duplexes (GCGGACGC)$_2$ and (GGCGAGCC)$_2$, respectively. In both iGiC duplexes, Watson−Crick-like hydrogen bonds are formed between iG and iC, and iGiC substitutions result in less favorable loop stability. The results provide benchmarks for testing computations of molecular interactions that shape RNA three-dimensional structure.

Introduction

Understanding the sequence-dependent interactions and thermodynamics of small RNA motifs can facilitate prediction of RNA secondary[1,2] and three-dimensional structure and, thus, functional significance. The stability[1−7] and three-dimensional structure[7−15] of RNA tandem GA pairs are dependent on sequence context. Tandem sheared GA (trans Hoogsteen/Sugar edge A−G) pairs (Figure 1) form with cross-strand base stacking in the symmetric contexts (CGAG)$_2$ with a loop free energy of −0.7 kcal/mol[1,3,8] and (UGAA)$_2$ with a loop free energy of 0.7 kcal/mol.[1,4,11] Tandem imino GA (cis Watson−Crick/Watson−Crick A−G) pairs (Figure 1) form in the symmetric context (GGAC)$_2$ with a loop free energy of −2.5 kcal/mol.[1,4,9] These results suggest that base stacking interactions (e.g., Coulombic and overlap) determine the hydrogen-bonding patterns of RNA tandem GA pairs in these internal loops.[7,9,16] Different shapes of GA pairs, in turn, provide different molecular features, e.g., van der Waals and electrostatic surface and hydrogen bonding donors and acceptors, for higher-order RNA folding and molecular recognition.[13−15,17−19]

The Watson−Crick-like isoG−isoC (iGiC) pair, with the amino and carbonyl groups transposed relative to the Watson−Crick GC pair, provides an expanded alphabet[20−26] for further understanding interactions that shape nucleic acid structure (Figure 1).[7,25−32] Here, thermodynamic studies of tandem GA pairs flanked by iGiC pairs are reported, along with NMR structures of the RNA duplexes (GCiGGAiCGCA)$_2$ and (GGiCGAiGCCA)$_2$. NMR restrained molecular dynamics and energy minimization of (GCiGGAiCGCA)$_2$ reveal a conformation of tandem sheared GA pairs closed by Watson−Crick-like iGiC pairs. This contrasts with the NMR structure of tandem imino GA pairs in (GCGGACGC)$_2$.[9] The loop free energy of (iGGAiC)$_2$ with sheared GA pairs is −0.7 kcal/mol, on average, at 37 °C in 1 M NaCl, which is less favorable than the −2.5 kcal/mol for (GGAC)$_2$ with imino GA pairs. NMR restrained modeling of (GGiCGAiGCCA)$_2$ reveals a predominant conformation of tandem imino GA pairs closed by Watson−Crick-like iGiC pairs, although at least one other conformation is also sampled. This contrasts with the NMR structure of tandem sheared GA pairs in (GGCGAGCC)$_2$.[8] The loop free energy of (iCGAiG)$_2$ with imino GA pairs is +0.7 kcal/mol, on average, at 37 °C in 1 M NaCl, which is less favorable than the −0.7 kcal/mol found for (CGAG)$_2$ with sheared GA pairs. Thus in both cases, the iGiC substitutions produce less favorable stability and change the shape of the internal loop. Presumably, stability would be even less favorable if the shape were maintained. The results provide benchmarks for testing methods that predict RNA stability and structure.

Experimental Methods

Oligoribonucleotide Synthesis and Purification. Oligonucleotides were synthesized using the phosphoramidite method[33,34] and purified as described previously.[14,35] Phosphoramidites for iG and iC were prepared as previously described.[27] CPG supports and phosphoramidites were acquired from Proligo. The mass of all oligonucleotides was verified by ESI-MS using a Hewlett-Packard 1100 LC/MS Chemstation. Purities were checked by reversed-phase HPLC, and all were greater than 95%.

† Part of the special issue "Norman Sutin Festschrift".

* To whom correspondence should be addressed. Phone: (585) 275-3207. Fax: (585) 276-0205. E-mail: turner@chem.rochester.edu.
‡ Department of Chemistry, University of Rochester.
§ Polish Academy of Sciences.
|| Department of Physics, University of Rochester.
⊥ Department of Biochemistry and Biophysics, University of Rochester.
Center for Pediatric Biomedical Research and Department of Pediatrics, University of Rochester.

10.1021/jp068732m CCC: $37.00
Published on Web 04/06/2007

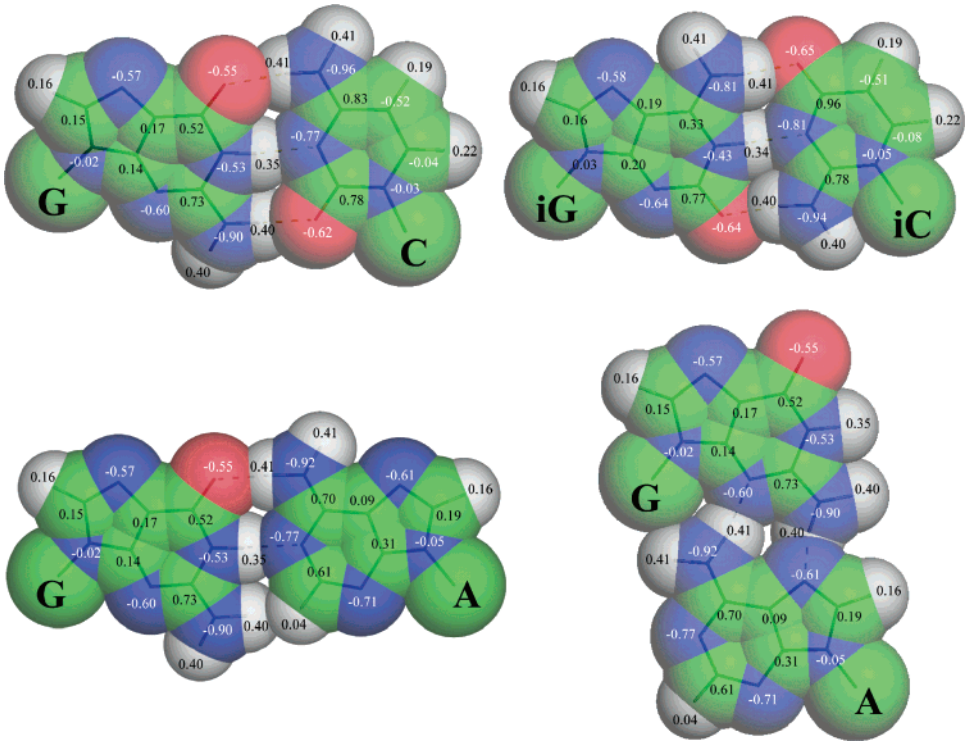

Figure 1. Schematic space-filling representation of Watson−Crick GC (top left), iGiC (top right), imino GA (bottom left), and sheared GA (bottom right) pairs. Partial charges are labeled on each atom of the RNA bases, with negative and positive charges shown in white and black, respectively.

UV Melting Experiments and Thermodynamics. Concentrations of single-stranded oligoribonucleotides were calculated from the absorbance at 280 nm at 80 °C, and extinction coefficients predicted from those of dinucleotide monophosphates and nucleosides[36,37] with the RNAcalc program (http://www.meltwin.com).[38] The extinction coefficients were estimated by replacing iG with G and iC with C. This approximation is within experimental error of measured extinction coefficients.[27] The buffer for UV melting was 1 M NaCl, 20 mM sodium cacodylate, and 0.5 mM disodium EDTA at 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 with water flow.

Melting curves were fit to a two-state model with the MeltWin program (http://www.meltwin.com), assuming linear sloping baselines and temperature-independent $\Delta H°$ and $\Delta S°$.[38−40] Additionally, the temperature in kelvin at which half the strands are in duplex, T_M, at total strand concentration, C_T, was used to calculate thermodynamic parameters for self-complementary duplexes according to[41]

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

Here, R is the gas constant, 1.987 cal/mol·K. All of the $\Delta H°$ values from T_M^{-1} versus $\ln(C_T)$ plots and from the average of the fits of melting curves to two-state transitions agree within 12%, suggesting that the two-state model is a reasonable approximation for these transitions. All the transitions have T_M's that are dependent on C_T, which indicates duplex formation instead of hairpin. 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, RNA sample preparation was similar to that previously reported.[14,42] The buffer was 80 mM NaCl, 10 mM sodium phosphate, 0.5 mM Na_2EDTA, pH 7.1 for RNA samples in H_2O. Total volumes were 300 μL with 9:1 (v:v) H_2O/D_2O for exchangeable proton spectra and 99.996% D_2O (Cambridge Isotope Laboratories) for nonexchangeable spectra. The sample was exchanged to D_2O by lyophilization three times from 99.96% D_2O and finally dissolved in 300 μL of 99.996% D_2O. The total RNA duplex concentrations were ~0.7 mM.

NMR Spectroscopy. Exchangeable and nonexchangeable proton spectra were acquired on Varian Inova 500 and 600 MHz (^{1}H) NMR spectrometers.[43] One-dimensional imino proton spectra were acquired in 9:1 (v:v) H_2O/D_2O with an S-shaped excitation pulse[43] at temperatures ranging from −5 to 40 °C. Two-dimensional SNOESY spectra were recorded with 100 and 150 ms mixing times at various temperatures between −5 and 30 °C. NOESY spectra of samples in D_2O were acquired at 30 °C for $(GCiG\underline{GA}iCGCA)_2$, and at 5 and 42 °C for $(GGiC\underline{GA}iGCCA)_2$ with 100, 200, and 400 ms mixing times. TOCSY spectra were acquired at 5, 30, and 42 °C with 8, 20, 40, and 100 ms mixing times. Exchangeable proton TOCSY spectra of $(GCiG\underline{GA}iCGCA)_2$ were acquired at −5 and 0 °C. The ^{1}H−^{31}P HETCOR and natural abundance ^{1}H−^{13}C HMQC spectra were acquired at 30 °C for $(GCiG\underline{GA}iCGCA)_2$ and at 5 and 42 °C for $(GGiC\underline{GA}iGCCA)_2$. The 1D ^{1}H-decoupled ^{31}P spectra of $(GCiG\underline{GA}iCGCA)_2$ were acquired on a Bruker Avance 500 MHz (^{1}H) NMR spectrometer at 30 °C and referenced to an external standard of 85% H_3PO_4 at 0 ppm. 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.), NMRPipe, and Sparky were used to process and analyze 2D spectra.

NMR Restraint Generation. A total of 12 hydrogen bond distance restraints limiting hydrogen-bond donor and acceptor distances were applied for the four GC pairs (GH1−CN3, 1.8−2.5 Å; GO6−CN4, 1.8−3.5 Å; and GN2−CO2, 1.8−3.5 Å), but no hydrogen bond restraints were used for the loop residues,

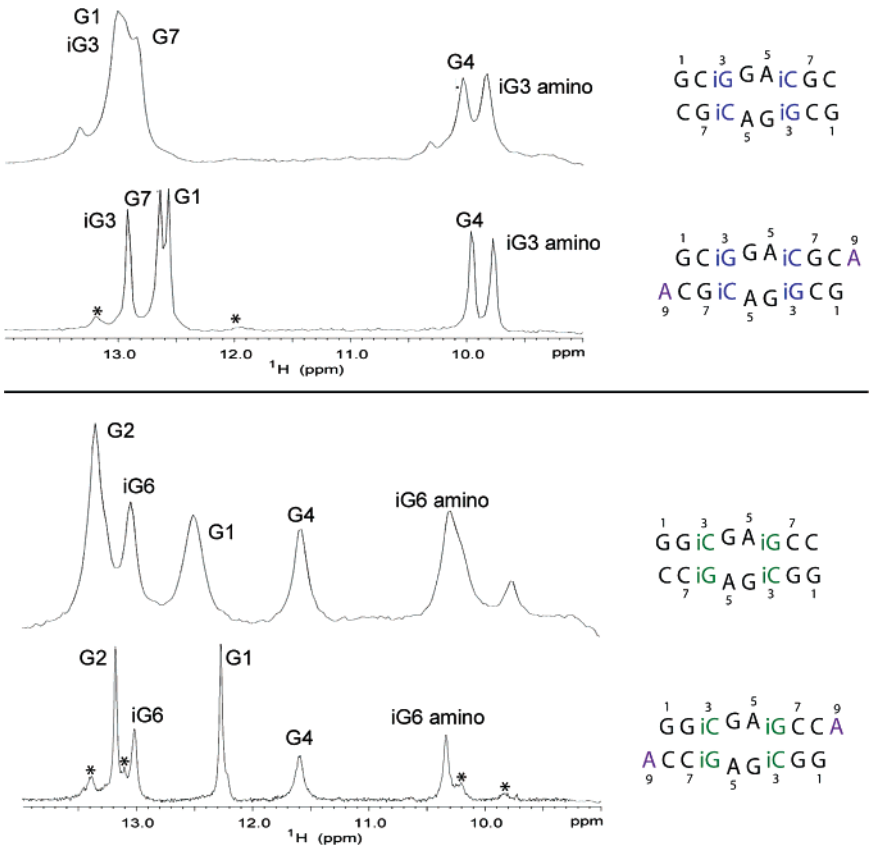

Figure 2. One-dimensional exchangeable proton NMR spectra of $(GCiG\underline{GA}iCGC)_2$ and $(GCiG\underline{GA}iCGCA)_2$ at −5 °C and $(GGiC\underline{GA}iGCC)_2$ and $(GGiC\underline{GA}iGCCA)_2$ at 5 °C. Imino proton peaks are marked with the residue name. Peaks due to a minor conformation are marked with an asterisk.

including iGiC pairs. Dihedral angles of the backbones were loosely restrained on the basis of ^{1}H−^{31}P HETCOR, DQF-COSY, and TOCSY spectra:[44,45] α (0 ± 120°) (with tandem GA pairs not restrained), β (180 ± 30°) (tandem GA pairs not restrained), γ (60 ± 30°), δ (85 ± 30°, the tandem GA pairs and 3′-end dangling A restrained to cover both C2′-endo and C3′-endo conformations (122.5 ± 67.5°)), ϵ (−140 ± 40°; the G of the tandem GA pairs restrained to cover both trans and gauche$^-$ (−135 ± 60°); and the A of the tandem GA pairs not restrained), ζ (0 ± 120°, the tandem GA pairs not restrained), and χ (−170 ± 40°, the tandem GA pairs restrained to be −120 ± 90°). No base planarity restraints were applied.

In summary, a total of 196 distance restraints (104 intranucleotide, 92 internucleotide, 24 cross-strand), including hydrogen bond restraints, and 104 dihedral angle restraints were used for the structure modeling of the RNA duplex $(GCiG\underline{GA}iCGCA)_2$ (Supporting Information Table S1). A total of 186 distance restraints (94 intranucleotide, 92 internucleotide, 28 cross-strand) and 104 dihedral angle restraints were used for the structure modeling of $(GGiC\underline{GA}iGCCA)_2$ (Supporting Information Table S3). No symmetry constraint was imposed, although the duplexes are self-complementary.

Structural Modeling. NMR restrained molecular dynamics and energy minimization were performed in vacuum without counterion and solvent with the Discover 98 package on a Silicon Graphics computer. The starting structures of $(GCiG\underline{GA}iCGCA)_2$ and $(GGiC\underline{GA}iGCCA)_2$ were generated by modifying A-form-like RNA duplexes in the Biopolymer module of Insight II (2000). The partial charges for atoms of iG and iC were calculated as previously described.[7,46,68] Other force field parameters for iG and iC were those previously described.[27] The modified AMBER 95 force field[46] was used with the 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,[47] with distance-dependent dielectric constant ($\epsilon = 2r$), was used for calculating electrostatic interactions. The progression of the structure simulation was the same as previously reported.[14,29,42] Figures 1 and 5 were generated with the PyMOL program.[48] The NMR structures are deposited in the Protein Data Bank with entries 2O81 and 2O83.

Results

Thermodynamics. Measured thermodynamic parameters for duplex formation at 1 M NaCl are listed in Table 1. Thermodynamic parameters for formation of the internal loops (Table 2) were calculated from measured parameters of duplexes according to the following equation:[49]

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

For example,

$$\Delta G^\circ_{37\ \substack{iG\ \underline{GA}\ iC\\ iC\ \underline{AG}\ iG}} = \Delta G^\circ_{37\ \substack{GCiG\ \underline{GA}\ iCGC\\ CGiC\ \underline{AG}\ iGCG}} - \Delta G^\circ_{37\ \substack{GCiGiCGC\\ CGiCiGCG}} + \Delta G^\circ_{37\ \substack{iGiC\\ iCiG}} \quad (2b)$$

Here, $\Delta G^\circ_{37\ \substack{GCiG\ \underline{GA}\ iCGC\\ CGiC\ \underline{AG}\ iGCG}}$ is the measured value of the duplex containing the internal loop (Table 1), $\Delta G^\circ_{37\ \substack{GCiGiCGC\\ CGiCiGCG}}$ is the measured value of the duplex without the loop,[27] and $\Delta G^\circ_{37\ \substack{iGiC\\ iCiG}}$ is the free energy increment for the nearest neighbor base stack interaction interrupted by the internal loop.[27] $\Delta H^\circ_{\text{loop}}$ and $\Delta S^\circ_{\text{loop}}$

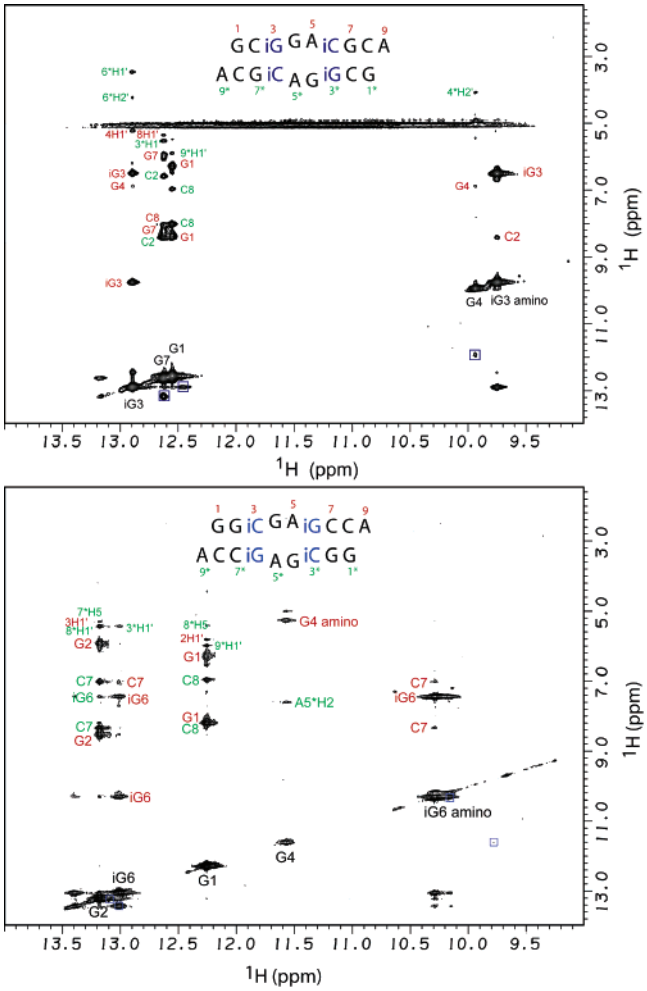

Figure 3. SNOESY spectra of (GCiGGAiCGCA)$_2$ at −5 °C (top) and (GGiCGAiGCCA)$_2$ at +5 °C (bottom) with a mixing time of 150 ms in 80 mM NaCl, 10 mM sodium phosphate, 0.5 mM Na$_2$EDTA, pH 7.1. Intrastrand cross-peaks are labeled in red and cross-strand cross-peaks are labeled in green. Cross-peaks to atoms that are not explicitly identified are to amino protons of the corresponding base. Cross-peaks due to exchange with a minor conformation are enclosed in blue boxes.

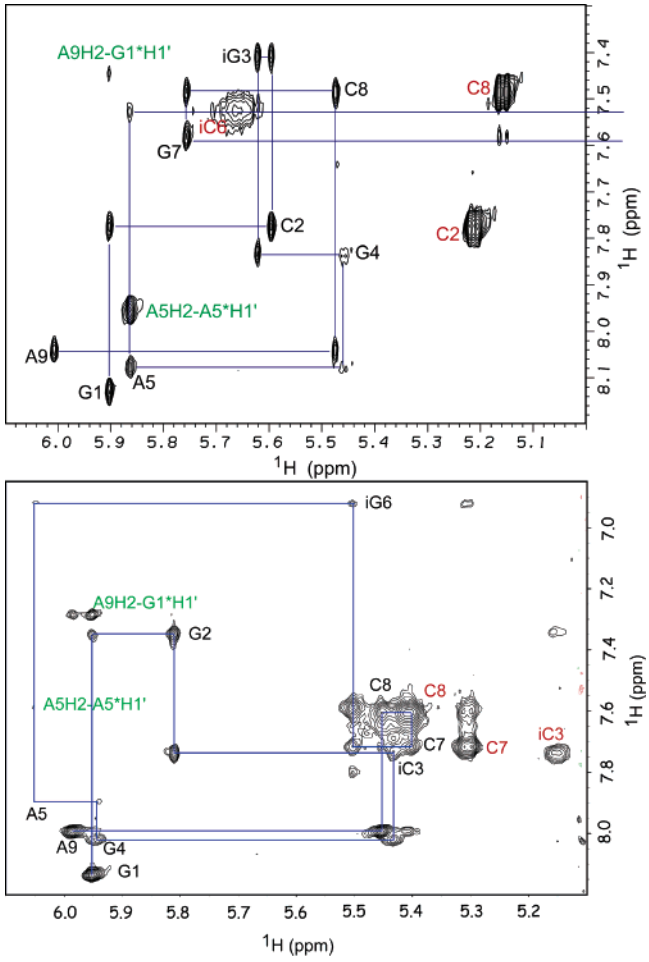

Figure 4. (H1′/H5)−(H8/H6/H2) region of the 400 ms mixing time NOESY spectra of (GCiGGAiCGCA)$_2$ at 30 °C (top) and (GGiCGAiGCCA)$_2$ at 5 °C (bottom) in 80 mM NaCl, 10 mM sodium phosphate, and 0.5 mM Na$_2$EDTA, pH 7.1. Cross-strand cross-peaks are labeled in green. The H5−H6 cross-peaks of C and iC residues are labeled in red.

are calculated similarly. All the thermodynamic parameters used in this calculation are derived from T_M^{-1} versus ln(C_T) plots (eq 1).

The stability increment of a single 3′-dangling A in the motif of $\begin{smallmatrix}CA\\G\end{smallmatrix}$ at the end of a Watson−Crick helix is −1.7 kcal/mol at 37 °C in 1 M NaCl,[50,51] which was applied twice for (GCiGiCGCA)$_2$ to calculate the free energy for formation of the internal loop (iGGAiC)$_2$ in (GCiGGAiCGCA)$_2$.

NMR Assignments and Structural Features. The 1D ^{31}P and ^{1}H−^{31}P HETCOR spectra reveal that all eight phosphorus resonances are within 1 ppm for (GCiGGAiCGCA)$_2$ and (GGiCGAiGCCA)$_2$, indicating no structure with severe backbone distortion. Thus, the sequences form self-complementary RNA duplexes, which is consistent with the dependence of UV melting temperatures on RNA total strand concentration.

Almost all proton and phosphorus NMR resonances were unambiguously assigned essentially as described previously[14,42,44,45] using TOCSY, DQF-COSY, ^{1}H−^{31}P HETCOR, NOESY, and natural abundance ^{1}H−^{13}C HMQC spectra. Because the duplexes are self-complementary, each resonance is due to two residues that have identical chemical shifts. Exchangeable proton 1D NMR spectra are shown in Figure 2 and were assigned by analysis of 2D SNOESY spectra[43,52−54] (Figure 3). Proton assignments are listed in Tables S2 and S4 of the Supporting Information. Formation of Watson−Crick-like iGiC pairs is consistent with the resonance of the iG imino protons in both duplexes near 13 ppm and a pair of resonances near 7 and 10 ppm for iG amino protons. When the iG amino is not hydrogen-bonded in a DNA iGT wobble pair, only a single resonance of iG amino protons at 7.35 ppm is observed.[31] The iC amino protons were not clearly resolved in either (GCiGGAiCGCA)$_2$ or (GGiCGAiGCCA)$_2$, although a broad resonance in (GCiGGAiCGCA)$_2$ at ∼8.75 ppm shows a cross-peak to iC6H1′. Both duplexes exhibit an NOE contact between iCH1′ and iGH1, presumably due to spin diffusion through iC amino protons in the iGiC Watson−Crick-like pair. Similar observations of amino protons of iG and iC in RNA and DNA were reported previously.[27,32] In (GCiGGAiCGCA)$_2$, the cross-peaks of iG3 amino-G4H1, iG3H1-G4 amino, iG3H1-G4H1′, iG3H8-G4H8, and typical A-form NOE contacts between C2 and iG3 (Supporting Information Table S1) further indicate formation of Watson−Crick-like iGiC base pairs. In (GGiCGAiGCCA)$_2$, Watson−Crick-like iGiC base pairs are also indicated by cross-peaks of iG6 amino-C7 amino, iG6H1-G2*H1, and typical A-form NOE contacts between G2 and iC3 (Supporting Information Table S3). Despite the evidence indicating Watson−Crick-like iGiC pairs, no hydrogen-bonding restraints were used for the iGiC pairs in the structural modeling.

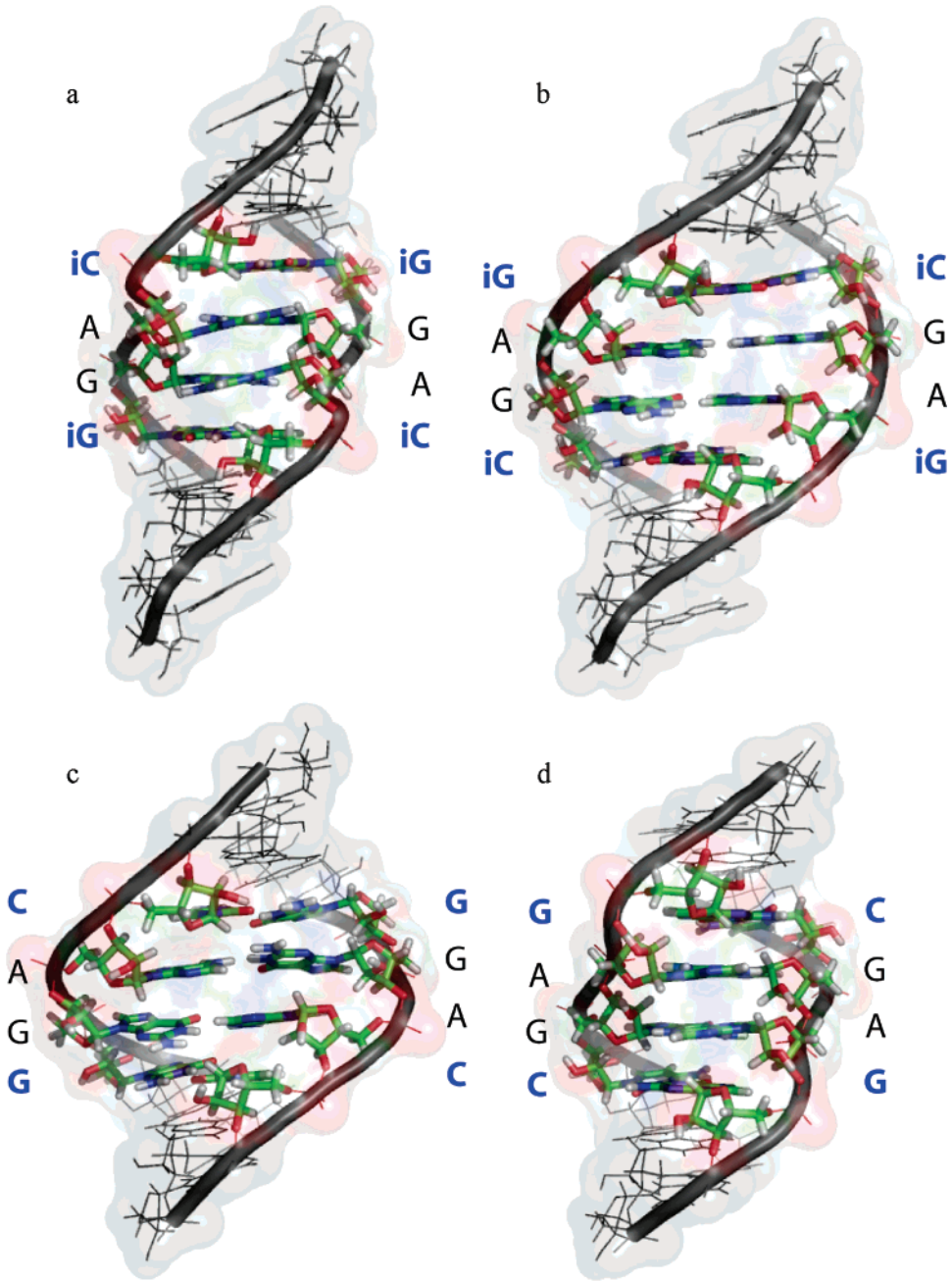

Figure 5. Minor groove view of average structures of the ensemble of modeled structures of (a) (GCiGGAiCGCA)$_2$ and (b) (GGiCGAiGCCA)$_2$, with tandem GA pairs and closing iG−iC pairs shown in sticks and colored according to atoms. For comparison with the case of tandem GA pairs and closing GC pairs, the structures of (c) (GCGGACGC)$_2$ (imino GA pair)[9] and (d) (GGCGAGCC)$_2$ (sheared GA pair)[8] are also shown.

The base−(H1′/H5) "NOESY walk" regions of the 400 ms NOESY spectra in D_2O are shown in Figure 4. The cross-strand cross-peak A9H2-G1*H1′ (asterisk represents cross-strand nucleotide) observed in both duplexes (with NMR distance of ∼4.5 Å) is consistent with the stacking of A9 on the terminal base pair (Figure 5) and thermodynamic stabilization of −1.5 kcal/mol per 3′ dangling A of (GCiGGAiCGCA)$_2$ at 37 °C in 1 M NaCl (Table 1). The favorable stacking of $\begin{smallmatrix}\mathrm{CA}\\ \mathrm{G}\end{smallmatrix}$ is also largely maintained in three-dimensional structures of large RNA.[72]

The 2′-hydroxyl protons were assigned in (GCiGGAiCGCA)$_2$ by combined analysis of exchangeable proton TOCSY and SNOESY spectra because they usually show scalar coupling with their own H2′ and NOE contact with their own H2′ and H1′. Most 2′-hydroxyl protons resonate between 6.5 and 7.5 ppm, which is similar to published results.[55,56] The G4 2′-hydroxyl proton resonates at 5.83 ppm and shows a strong TOCSY cross-peak with G4H2′, which might be enhanced due to slow water exchange of the G4 2′-hydroxyl proton caused by a cross-strand hydrogen bond to the A5 amino (Table 3).[55] No distance or dihedral angle restraints involving 2′-OH were generated for structural modeling.

The G4 in both (GCiGGAiCGCA)$_2$ and (GGiCGAiGCCA)$_2$ shows weak (2−5 Hz) H1′−H2′ coupling, as indicated by the G4H1′−H2′ TOCSY cross-peak that is comparable to the A9H1′−H2′ and G1H1′−H2′ cross-peaks (Figure S4), suggesting that the G4 ribose ring samples both C2′- and C3′-endo conformations. No direct evidence of C2′-endo conformation was noted for A5 or the iGiC residues in either duplex. Nonetheless, the dihedral angle restraints used in structural

TABLE 1: Measured Thermodynamic Parameters for RNA Duplex Formation in 1 M NaCl, pH7

	T_M^{-1} vs ln(C_T) plots (eq 1)				av of melt curve fits			
sequence	$-\Delta H°$ (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)
GCiG**GA**iCGC CGiC**AG**iGCG	71.2 ± 1.4	200.9 ± 4.4	8.84 ± 0.05	51.5	72.6 ± 4.4	205.2 ± 13.5	8.95 ± 0.17	51.6
GCiG**GA**iCGCA ACGiC**AG**iGCG	86.4 ± 4.1	240.2 ± 12.3	11.86 ± 0.27	60.9	77.1 ± 5.5	212.2 ± 16.8	11.29 ± 0.31	61.3
GCiC**GA**iGGC CGiG**AG**iCCG	69.8 ± 1.0	195.1 ± 3.2	9.25 ± 0.05	53.7	74.9 ± 4.3	210.9 ± 13.2	9.52 ± 0.22	53.8
GGiC**GA**iGCC CCiG**AG**iCGG	68.7 ± 5.0	191.4 ± 15.2	9.31 ± 0.25	54.3	72.3 ± 4.6	202.4 ± 14.0	9.50 ± 0.33	54.3
GCG**GA**CGC[b,d] CGC**AG**GCG	79.5 ± 9.5	225.0 ± 30.4	9.67 ± 0.39	53.4	71.8± 8.6	201.3 ± 27.2	9.35 ± 0.37	53.7
GGC**GA**GCC[c,d] CCG**AG**CGG	66.1 ± 7.9	181.8 ± 24.5	9.68 ± 0.39	57.0	73.9 ± 8.9	205.6 ± 27.8	10.16 ± 0.41	57.0

[a] At $C_T = 0.1$ mM. [b] Ref.[4] [c] Ref.[3] [d] Experimental errors for $\Delta G°_{37}$, $\Delta H°$, and $\Delta S°$ are estimated to be 4, 12, and 13.5%, respectively, according to ref 40.

TABLE 2: Measured Thermodynamic Parameters for RNA Internal Loop Formation in 1 M NaCl, pH7

sequence	$\Delta G°_{37,loop}$ (kcal/mol)	$\Delta H°_{loop}$ (kcal/mol)	$\Delta S°_{loop}$ (eu)
GCiG**GA**iCGC CGiC**AG**iGCG	−0.89 ± 0.23	−16.2 ± 3.2	−48.9 ± 9.3
GCiG**GA**iCGCA ACGiC**AG**iGCG	−0.51 ± 0.35	−13.4 ± 5.0	−41.4 ± 14.8
GCiC**GA**iGGC[a] CGiG**AG**iCCG	0.42 ± 0.20	−13.7 ± 3.2	−46.2 ± 9.5
GGiC**GA**iGCC CCiG**AG**iCGG	0.93 ± 0.43	−6.3 ± 6.6	−23.6 ± 19.7
GCG**GA**CGC[b] CGC**AG**GCG	−2.47 ± 0.43	−28.4 ± 10.0	−83.5 ± 32.0
GGC**GA**GCC[c] CCG**AG**CGG	−0.71 ± 0.42	−8.9 ± 8.4	−26.5 ± 26.0

[a] Thermodynamic parameters for $\frac{GCiCiGGC}{CGiGiCCG}$ were calculated with nearest-neighbor parameters from refs.[27,40] [b] Calculated from data in refs.[4,40] [c] Calculated from data in refs.[3,40]

modeling allowed both sugar puckers for all loop residues. In the resulting models, all loop residues exhibit C3′-endo conformation.

NMR restrained modeling and energy minimization of (GCiGGAiCGCA)$_2$ reveal a symmetric internal loop of tandem sheared GA pairs closed by Watson−Crick-like iGiC pairs. Such a conformation is consistent with cross-strand NOE contacts of G4 amino-G4*H2′/H1′, G4H1-G4*H2′/H1′, G4 amino-A5* amino, G4 amino-A5*H8, A5 amino-A5*H8, and A5H2-A5*H1′. Here, the asterisks represent cross-strand nucleotides. The strong A5H2-A5*H1′ cross-peak, with NMR distance ∼2.7 Å indicates cross-strand base stacking between two A's in tandem GA pairs. This contrasts with a weak cross-strand A5H2-A5*H1′ cross-peak, with an NMR distance of ∼4.1 Å, that was observed in (GCGGACGC)$_2$, which contains tandem imino GA pairs with substantial intrastrand base stacking.[9] Several exchangeable proton cross-peaks with potential spin diffusion effects were not used to generate distance restraints.

A total of 18 of 20 modeled low-energy structures of (GCiGGAiCGCA)$_2$ were selected for analysis. Two of the modeled structures were deleted on the basis of NMR restraint violations and the calculated total energy values. The average of total energies for the selected 18 structures is −397.8 ± 2.4 kcal/mol with the two deleted ones at −367.0 and −386.7 kcal/mol. The average root-mean-square deviation of all selected structures to the average structure for all atoms is 0.30 ± 0.04 Å. No distance or dihedral angle restraint violations were greater than 0.1 Å or 1°, respectively. The average structure is shown in Figure 5a.

The NMR spectra of (GCiGGAiCGCA)$_2$ contain weak, sometimes broad additional resonances, indicating a second conformation with less than 10% population (e.g., Figure 2). Conformational exchange cross-peaks are observed for iG3H1, G4H1, and G7H1 in the 2D SNOESY spectrum in Figure 3 (top) and for iC6H1′, iC6H5, and G4H1′ in a 2D ROESY spectrum at 5 °C (Figure S3). The chemical shifts of protons in the major and minor conformations are listed in Table S2. There is not enough information to provide any definite conclusions about the minor conformation.

In contrast to the sheared GA pairs observed in (GCiGGAiCGCA)$_2$, NMR restrained modeling and energy minimization of (GGiCGAiGCCA)$_2$ reveal a predominant symmetric internal loop of tandem imino GA pairs closed by Watson−Crick-like iGiC pairs. This conformation is consistent with medium to strong NOE contacts for G4H1-A5*H2 (∼2.7 Å) and G4 amino-A5*H2 (∼2.4 Å), and a weak NOE contact for A5H2-A5*H1′. The GH1 imino and the AH2 protons are expected to be less than 2.5 Å apart in the imino hydrogen-bonded conformation, whereas for sheared GA pairs, this distance is greater than 9 Å. The weak A5H2-A5*H1′ NOE contact is similar to that observed in (GCGGACGC)$_2$, which contains tandem imino GA pairs, and is also similar to that observed between AH2 and the H1′ proton 3′ of the U in a canonical AU pair in A-form RNA. Finally, the chemical shift of the G4H1 imino proton (11.5 ppm, Figure 2) is substantially downfield from that commonly observed for sheared GA pairs, including (GCiGGAiCGCA)$_2$, suggesting involvement in a hydrogen bond.

NMR spectra also indicate that (GGiCGAiGCCA)$_2$ exists in more than one conformation. The average population of the minor conformation at 5 °C is estimated as 21 ± 5% from the area of minor resonances in 1D spectra of the imino proton region (Figure 2) and the volume in 2D TOCSY spectra of pyrimidine H5−H6 cross-peaks (Figure S1). The chemical shifts of protons in the major and minor conformations are listed in Table S4; the protons that exhibit the largest chemical shift differences are in or near the internal loop. NOESY and TOCSY cross-peaks between the same protons in the two conformations (generally near the diagonal) indicate conformational exchange occurs on a time scale comparable to or shorter than the mixing time (Figure 3 and Figure S1). Changes in chemical shift and

TABLE 3: Potential Cross-Strand Hydrogen Bonds and Distances in Angstroms (between Hydrogen and Hydrogen Bond Acceptor) Involving Tandem GA Pairs Observed in the Averaged NMR Structures of (GCiGGAiCGCA)$_2$ and (GGCGAGCC)$_2$[a] and the Crystal Structures of (GGCGAGCC)$_2$'[b] and (GICGAGCC)$_2$[c] Where I is Inosine[d,e]

				(CGAG)$_2$ crystal 1SA9			(CGAG)$_2$ crystal 1SAQ		
HB acceptor	HB donor	(iGGAiC)$_2$ NMR	(CGAG)$_2$ NMR	A	B	C	A	B	C
G4 N3	A5*amino	3.43/3.48	1.95	1.71/1.82	3.23/1.80	1.90	1.89/2.41	2.52/2.65	3.32
G4 2′-OH	A5*amino	3.17/3.06	1.71	2.16/1.97	3.06/2.34	1.72	2.24/1.86	2.51/2.61	4.92
A5 N7	G4*amino	1.90/1.90	1.62	1.38/1.68	1.70/1.83	1.84	2.25/1.61	1.92/2.21	3.47
A5 OP	G4*amino	1.75/1.75	1.97	3.33/2.98	3.34/3.05	2.61	3.56/2.90	3.25/4.27	4.95

[a] PDB: 1YFV; 2-fold symmetry was applied as restraint.[8] [b] PDB: 1SA9. [c] PDB: 1SAQ.[62] [d] When 2-fold symmetry was not applied in the structural modeling, potential hydrogen bond distances are given with the hydrogen bond donors (or acceptors) from the same strand always on the same side of the slash. [e] Additional potential intrastrand hydrogen bonds in the (GCiGGAiCGCA)$_2$ duplex, with all of the hydrogen bond acceptors on the left side of the dash, are A5 O4′-G4 2′OH (2.47/2.41), A5 O5′-G4 2′OH (2.03/2.03), iC6 2′OH-iC6 amino (2.31/2.39), iC6 OP-A5 2′OH (2.03/2.38), G7 O4′-iC6 2′OH (2.26/2.37), and G7 O5′-iC6 2′OH (2.00/2.15). An interaction that may stabilize the imino conformation of (GGiCGAiGCCA)$_2$ is the G4 O4′-iC3 2′OH hydrogen bond (2.09/2.13 Å). This backbone interaction is occasionally observed in RNA.[70,71] Formation of these additional potential hydrogen bonds is probably force-field-dependent because very few restraints are applied for backbones. All potential hydrogen bond distances are shown, although they do not necessarily form in all structures generated.

line width as a function of temperature (Figure S2) reflect changes of the conformational exchange rate. Analysis of the temperature dependence[57,58] of signals from iG6H8 and A5H1′ suggests the exchange rate varies from ~150 s^{-1} at 0 °C to ~20 000 s^{-1} at 50 °C. No NOE contacts are observed among minor conformation resonances, so no distance restraints can be obtained for the minor structure.

A total of 18 of 20 modeled low-energy structures of (GGiCGAiGCCA)$_2$ were selected for analysis. Two modeled structures were deleted on the basis of NMR restraint violations and the calculated total energy values. The average of total energies for the 18 selected structures is −379.8 ± 1.9 kcal/mol, with the deleted ones at −375.3 and −374.6 kcal/mol. The average root-mean-square deviation of all selected structures to the average structure for all atoms is 0.35 ± 0.05 Å. No distance or dihedral angle restraint violations were greater than 0.1 Å or 1°, respectively. The average structure is shown in Figure 5b.

Discussion

The physical−chemical nature of RNA sequences determines the interactions accounting for three-dimensional conformation, stability, and function. The effect of flanking Watson−Crick-like iGiC pairs on the structure and thermodynamics of tandem GA pairs provides a model system for better understanding such relationships.

The most dramatic result from this study is that substituting iGiC pairs for GC pairs adjacent to tandem GA pairs changes the shape of the GA pairs (Figures 5 and 6). In particular, the (GGAC)$_2$ motif has imino GA pairs,[9] whereas (iGGAiC)$_2$ has sheared GA pairs. The (CGAG)$_2$ motif has sheared GA pairs,[8] whereas (iCGAiG)$_2$ has imino GA pairs in its predominant structure. Because the iGiC pairs are isosteric with GC pairs, this indicates that the structure of the GA pairs is not dependent on the possible overlap area with the adjacent base pairs. At least three other types of interaction may explain the observed sequence dependence. It has been suggested that a stabilizing hydrogen bond could form between a nonplanar amino group of the G in an imino GA pair and the carbonyl of the C in a 5′GG/3′CA stack.[59] This interaction is not available in the 5′iGG/3′iCA stack, but a similar interaction might form between a nonplanar G4 amino group and the N3 of iG6 in the 5′iCG/3′iGA stack of (GGiCGAiGCCA)$_2$ when the GA is in an imino conformation (see Figure 6). The structures generated from NMR data do not contain this type of interaction because nonplanar amino groups are not favored by the force field. The second type of interaction that may determine the shape of the GA pairs is the electrostatic interaction between partial charges on the bases. Figure 1 shows the calculated partial charges,[7,46,68] and Figure 6 illustrates the electrostatic potential between base pairs due to the partial charges. Many atoms of each base have large partial charges, and the relative positions of these partial charges between stacked bases will be a determinant of the free energy of a structure. A third possibility is specific solvation and metal ion binding to the various motifs.

The thermodynamics of the RNA internal loops also change when iGiC pairs are substituted for GC pairs. The loop free energy increment becomes less favorable by ~1.6 kcal/mol when iGiC is substituted for GC in the pairs adjacent to GA in both (GGAC)$_2$ and (CGAG)$_2$ (see Table 2). In both cases, the structure of the GA pairs changes, indicating that the thermodynamic stability would be even less favorable if the structure were maintained.

Although the switches between imino and sheared GA pairs are the most dramatic effects observed, the structures presented here also provide new examples of more subtle structural effects. For example, the sugar puckers of the G in the GA pair of the (iGGAiC)$_2$ and (iCGAiG)$_2$ motifs are a mixture of C3′- and C2′-endo, whereas only C3′-endo sugar puckers were observed for the G of GA pairs in (GGAC)$_2$ and (CGAG)$_2$.[8,9] In contrast, the sugar pucker is C2′-endo in the (UGAA)$_2$ motif, and it has been suggested that the sugar pucker and the base−backbone hydrogen-bonding pattern correlate with the stability of the internal loop.[11,60] This is also consistent with the mixed sugar pucker observed here for the (iCGAiG)$_2$ loop and with observations on tandem GA pairs closed by GU pairs.[15] The current results, however, suggest this is not a strict correlation because the (iGGAiC)$_2$ internal loop has a favorable free energy increment, but the G has a mixture of C3′ and C2′-endo sugars. Nonetheless, loop stability and sugar pucker are correlated for the GC-to-iGiC replacements. The replacements result in decreased loop stability and increased C2′-endo population, despite the fact that iGiC and GC pairs have the same number of hydrogen bonds and similar stability.[27]

A variety of hydrogen bonds can form with sheared GA pairs, and there can be a dynamic interchange between them.[61] Table 3 lists some hydrogen bond distances that have been reported in NMR and crystal structures of duplexes with tandem sheared GA pairs.[8,62] The A5 N7-G4* amino hydrogen bond is relatively short in all but one case. In contrast, the G4 N3-A5* amino hydrogen bond is quite variable and particularly long in the (iGGAiC)$_2$ motif. In fact, the distance of roughly 3.4 Å in the latter case is outside the range associated with hydrogen bonds.[63] All the structures in Table 3 have a C3′-endo sugar pucker for

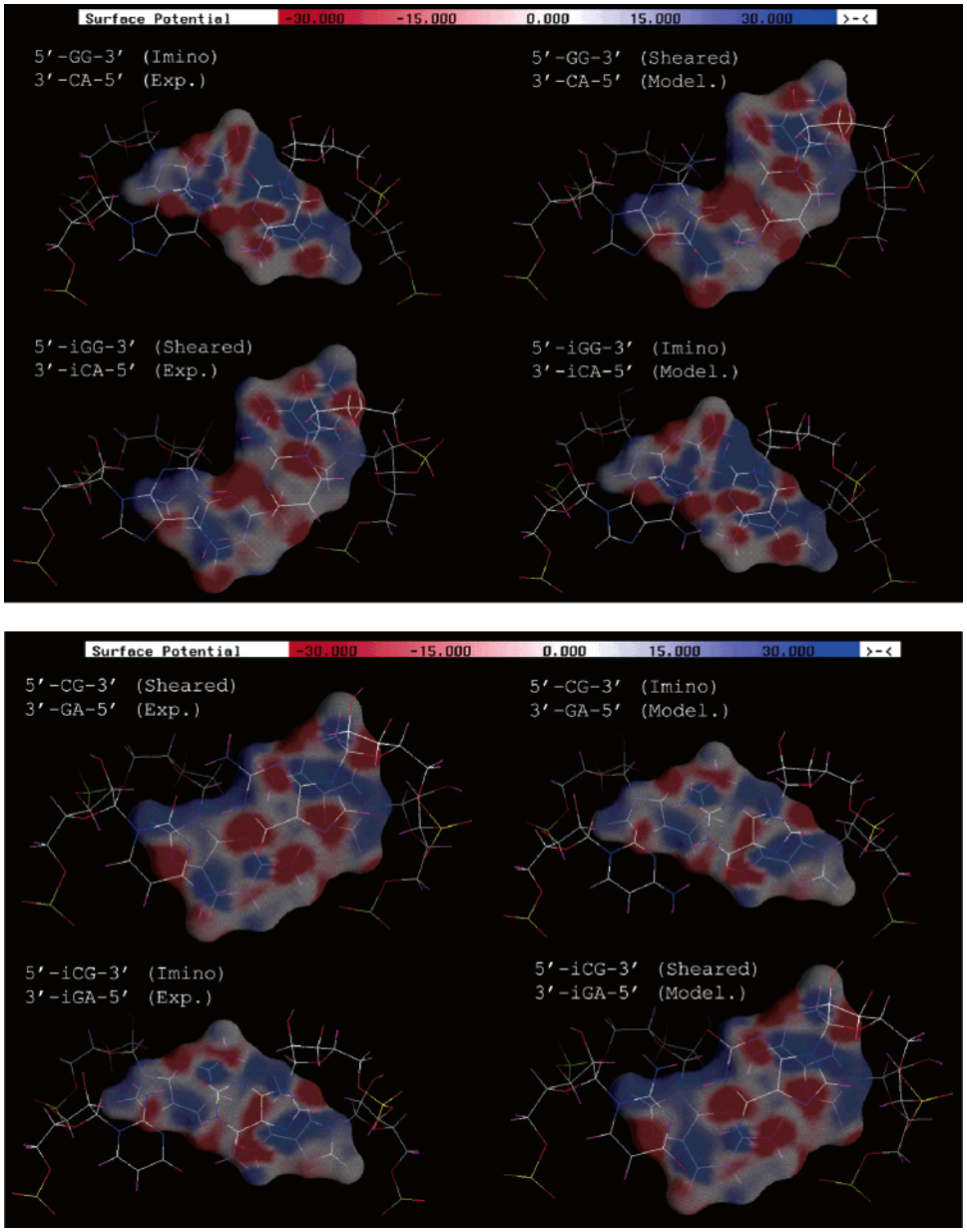

Figure 6. Stacking between GC or iGiC pairs and imino or sheared GA pairs. Structures derived from NMR data are labeled "Exp." Structures labeled "Model" were generated from the NMR structures by transposing the amino and carbonyl groups on the GC or iGiC pair. The top pair is GC or iGiC. The colors represent the electrostatic potential surfaces calculated with the program GRASP[67] using RESP partial charges.[7,46,68] The reference frame for each structure is based on the GA base pair, and the structures were reoriented with the program 3DNA.[69] Electrostatic potential surfaces were created around the GA pair with a probe radius of 1.4 Å and solvent and internal dielectric constants of 80.0 and 2.0, respectively. Red and blue regions represent negative and positive potential, respectively, with the spectrum covering −30 to +30 kT/*e*.

the G in the tandem sheared GA pairs, resulting in similar base-backbone hydrogen bonds. Interestingly, the $(iG\underline{GA}iC)_2$ motif structure reported here has the shortest A5 OP-G4* amino hydrogen bond, perhaps compensating for the long G4 N3-A5* amino "hydrogen bond". The A5 OP-G4* amino hydrogen bond in both solution NMR models is considerably shorter than that in all the crystal structures. This could reflect the force field used for modeling and/or it could reflect packing effects in the crystals.

Although no distance restraints were obtained that could help define the minor conformation of $(GGiC\underline{GA}iGCCA)_2$, three pieces of evidence point to the possibility that the GA pairs transiently sample a sheared-type conformation. In the minor conformation, (1) the iG6H1′ proton is shifted 1.27 ppm upfield relative to the major conformation (Figure S1 and Table S4), consistent with the shift induced by the adenine ring of a sheared GA pair, as observed for iC6H1′ in $(GCiG\underline{GA}iCGCA)_2$ and other structures;[8,14,64] (2) the G4 imino proton is shifted upfield to 9.8 ppm (Figure 2, 3, and Table S4), consistent with an imino proton not involved in a hydrogen bond; and (3) the G4 ribose ring is observed to exist partially in the C2′-endo conformation (Figure S4), consistent with the observation that in several structures involving thermodynamically less stable sheared GA pairs, the G ribose is found to be in the C2′-endo conformation.[11,14,15,60,64] It cannot be determined, however, whether the C2′-endo pucker occurs when the GA pair is in the major or minor conformation or both.

It is also not possible to define the minor conformation of $(GCiG\underline{GA}iCGCA)_2$. Two chemical shift changes complementary to those observed in $(GGiC\underline{GA}iGCCA)_2$ are consistent with a transient imino GA pair. In the minor conformation, (1) the iC6H1′ proton is shifted 1.65 ppm downfield relative to the

major sheared GA conformation (Figure S3 and Table S2), suggesting that it is no longer significantly influenced by the shielding effect of the adenine ring current, and (2) G4H1 is shifted 2.1 ppm downfield (Figure 2, 3, and Table S2) suggesting involvement in a hydrogen bond. Furthermore, the shift of G4H1′ in the minor conformation is similar to the shifts of G4H1′ in the imino conformations of $(GGiC\underline{GA}iGCCA)_2$ and $(GCG\underline{GA}CGC)_2$ (~5.9 ppm). All these data taken together suggest that these duplexes can switch between sheared and imino conformations.

Conclusions

Improved understanding of the sequence dependence of energetics, structures, and dynamics of RNA internal loops may facilitate prediction of three-dimensional as well as secondary structure.[1,2,6,14,59,65,66] Many interactions determine the structures and thermodynamic stabilities of internal loops with tandem GA pairs. The data here provide benchmarks to test computational approaches for understanding the relative importance of the various interactions and thus provide the power to predict structure, stability, and dynamics for other internal loops.

Acknowledgment. This work was supported by NIH grants GM22939 (D.H.T.) and 1R03 TW1068 (D.H.T. and R.K.). We thank Dr. Sandip Sur for maintaining the 500 MHz (^{1}H) NMR instruments.

Supporting Information Available: Tables listing chemical shift assignments and NMR distance restraints; 2D TOCSY, additional 2D NOESY, 2D ROESY, and 1D spectra at a range of temperatures are available. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Mathews, D. H.; Sabina, J.; Zuker, M.; Turner, D. H. *J. Mol. Biol.* **1999**, *288*, 911.
(2) Mathews, D. H.; Disney, M. D.; Childs, J. L.; Schroeder, S. J.; Zuker, M.; Turner, D. H. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7287.
(3) SantaLucia, J., Jr.; Kierzek, R.; Turner, D. H. *Biochemistry* **1990**, *29*, 8813.
(4) Walter, A. E.; Wu, M.; Turner, D. H. *Biochemistry* **1994**, *33*, 11349.
(5) Schroeder, S. J.; Turner, D. H. *Biochemistry* **2001**, *40*, 11509.
(6) Chen, G.; Turner, D. H. *Biochemistry* **2006**, *45*, 4025.
(7) Yildirim, I.; Turner, D. H. *Biochemistry* **2005**, *44*, 13225.
(8) SantaLucia, J., Jr.; Turner, D. H. *Biochemistry* **1993**, *32*, 12612.
(9) Wu, M.; Turner, D. H. *Biochemistry* **1996**, *35*, 9677.
(10) Wu, M.; SantaLucia, J., Jr.; Turner, D. H. *Biochemistry* **1997**, *36*, 4449.
(11) Heus, H. A.; Wijmenga, S. S.; Hoppe, H.; Hilbers, C. W. *J. Mol. Biol.* **1997**, *271*, 147.
(12) Zhang, X. H.; Gaffney, B. L.; Jones, R. A. *J. Am. Chem. Soc.* **1998**, *120*, 6625.
(13) Chin, K.; Sharp, K. A.; Honig, B.; Pyle, A. M. *Nat. Struct. Biol.* **1999**, *6*, 1055.
(14) Chen, G.; Znosko, B. M.; Kennedy, S. D.; Krugh, T. R.; Turner, D. H. *Biochemistry* **2005**, *44*, 2845.
(15) Tolbert, B. S.; Kennedy, S. D.; Schroeder, S. J.; Krugh, T. R.; Turner, D. H. *Biochemistry* **2007**, *46*, 1511.
(16) Villescas-Diaz, G.; Zacharias, M. *Biophys. J.* **2003**, *85*, 416.
(17) Disney, M. D.; Turner, D. H. *Biochemistry* **2002**, *41*, 8113.
(18) Doherty, E. A.; Batey, R. T.; Masquida, B.; Doudna, J. A. *Nat. Struct. Biol.* **2001**, *8*, 339.
(19) Nissen, P.; Ippolito, J. A.; Ban, N.; Moore, P. B.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4899.
(20) Agris, P. F. *Nucleic Acids Res.* **2004**, *32*, 223.
(21) Henry, A. A.; Romesberg, F. E. *Curr. Opin. Chem. Biol.* **2003**, *7*, 727.
(22) Roberts, C.; Bandaru, R.; Switzer, C. *J. Am. Chem. Soc.* **1997**, *119*, 4640.
(23) Switzer, C.; Moroney, S. E.; Benner, S. A. *J. Am. Chem. Soc.* **1989**, *111*, 8322.
(24) Switzer, C. Y.; Moroney, S. E.; Benner, S. A. *Biochemistry* **1993**, *32*, 10489.
(25) Chaput, J. C.; Switzer, C. *J. Am. Chem. Soc.* **2000**, *122*, 12866.
(26) Roberts, C.; Bandaru, R.; Switzer, C. *Tetrahedron Lett.* **1995**, *36*, 3601.
(27) Chen, X.; Kierzek, R.; Turner, D. H. *J. Am. Chem. Soc.* **2001**, *123*, 1267.
(28) Chen, X.; McDowell, J. A.; Kierzek, R.; Krugh, T. R.; Turner, D. H. *Biochemistry* **2000**, *39*, 8970.
(29) Burkard, M. E.; Turner, D. H. *Biochemistry* **2000**, *39*, 11748.
(30) Geyer, C. R.; Battersby, T. R.; Benner, S. A. *Structure* **2003**, *11*, 1485.
(31) Robinson, H.; Gao, Y. G.; Bauer, C.; Roberts, C.; Switzer, C.; Wang, A. H. J. *Biochemistry* **1998**, *37*, 10897.
(32) Yang, X. L.; Sugiyama, H.; Ikeda, S.; Saito, I.; Wang, A. H. J. *Biophys. J.* **1998**, *75*, 1163.
(33) Usman, N.; Ogilvie, K. K.; Jiang, M. Y.; Cedergren, R. J. *J. Am. Chem. Soc.* **1987**, *109*, 7845.
(34) Wincott, F.; Direnzo, A.; Shaffer, C.; Grimm, S.; Tracz, D.; Workman, C.; Sweedler, D.; Gonzalez, C.; Scaringe, S.; Usman, N. *Nucleic Acids Res.* **1995**, *23*, 2677.
(35) Chen, G.; Znosko, B. M.; Jiao, X. Q.; Turner, D. H. *Biochemistry* **2004**, *43*, 12865.
(36) Borer, P. N. Optical properties of nucleic acids, absorption and circular dichroism spectra. In *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*, 3rd ed.; Fasman, G. D., Ed.; CRC Press: Cleveland, OH, 1975; Vol. I; p 589.
(37) Richards, E. G. Use of tables in calculation of absorption, optical rotatory dispersion and circular dichroism of polyribonucleotides. In *Handbook of Biochemistry and Molecular Biology: Nucleic Acids*, 3rd ed.; Fasman, G. D., Ed.; CRC Press: Cleveland, OH, 1975; Vol. 1; pp 596.
(38) McDowell, J. A.; Turner, D. H. *Biochemistry* **1996**, *35*, 14077.
(39) Petersheim, M.; Turner, D. H. *Biochemistry* **1983**, *22*, 256.
(40) Xia, T.; SantaLucia, J., Jr.; Burkard, M. E.; Kierzek, R.; Schroeder, S. J.; Jiao, X.; Cox, C.; Turner, D. H. *Biochemistry* **1998**, *37*, 14719.
(41) Borer, P. N.; Dengler, B.; Tinoco, I., Jr.; Uhlenbeck, O. C. *J. Mol. Biol.* **1974**, *86*, 843.
(42) Znosko, B. M.; Burkard, M. E.; Schroeder, S. J.; Krugh, T. R.; Turner, D. H. *Biochemistry* **2002**, *41*, 14969.
(43) Lukavsky, P. J.; Puglisi, J. D. *Methods* **2001**, *25*, 316.
(44) Varani, G.; Aboulela, F.; Allain, F. H. T. *Prog. Nucl. Magn. Reson. Spectrosc.* **1996**, *29*, 51.
(45) Varani, G.; Tinoco, I., Jr. *Q. Rev. Biophys.* **1991**, *24*, 479.
(46) 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.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(47) Ding, H. Q.; Karasawa, N.; Goddard, W. A. *J. Chem. Phys.* **1992**, *97*, 4309.
(48) DeLano, W. L. *The PyMOL User's Manual*; DeLano Scientific: San Carlos, CA, 2002.
(49) Gralla, J.; Crothers, D. M. *J. Mol. Biol.* **1973**, *78*, 301.
(50) Turner, D. H.; Sugimoto, N.; Freier, S. M. *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 167.
(51) Turner, D. H. Conformational Changes. In *Nucleic Acids: Structures, Properties, and Functions*; Bloomfield, V. A., Crothers, D. M., Tinoco, I., Jr., Eds.; University Science Books: Sausalito, CA, 2000; p 259.
(52) Kerwood, D. J.; Cavaluzzi, M. J.; Borer, P. N. *Biochemistry* **2001**, *40*, 14518.
(53) Mueller, L.; Legault, P.; Pardi, A. *J. Am. Chem. Soc.* **1995**, *117*, 11043.
(54) Schroeder, K. T.; Skalicky, J. J.; Greenbaum, N. L. *RNA* **2005**, *11*, 1012.
(55) Hennig, M.; Fohrer, J.; Carlomagno, T. *J. Am. Chem. Soc.* **2005**, *127*, 2028.
(56) Du, Z. H.; Yu, J. H.; Ulyanov, N. B.; Andino, R.; James, T. L. *Biochemistry* **2004**, *43*, 11959.
(57) Gutowsky, H. S.; Holm, C. H. *J. Chem. Phys.* **1956**, *25*, 1228.
(58) Legault, P.; Pardi, A. *J. Am. Chem. Soc.* **1997**, *119*, 6621.
(59) Sponer, J.; Mokdad, A.; Sponer, J. E.; Spackova, N.; Leszczynski, J.; Leontis, N. B. *J. Mol. Biol.* **2003**, *330*, 967.
(60) Michiels, P. J. A.; Schouten, C. H. J.; Hilbers, C. W.; Heus, H. A. *RNA* **2000**, *6*, 1821.
(61) Jucker, F. M.; Heus, H. A.; Yip, P. F.; Moors, E. H. M.; Pardi, A. *J. Mol. Biol.* **1996**, *264*, 968.
(62) Jang, S. B.; Baeyens, K.; Jeong, M. S.; SantaLucia, J., Jr.; Turner, D. H.; Holbrook, S. R. *Acta Crystallogr. D* **2004**, *60*, 829.
(63) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press, Inc.: New York, 1997, p12.
(64) Chen, G.; Kennedy, S. D.; Qiao, J.; Krugh, T. R.; Turner, D. H. *Biochemistry* **2006**, *45*, 6889.
(65) Strobel, S. A.; Ortoleva-Donnelly, L.; Ryder, S. P.; Cate, J. H.; Moncoeur, E. *Nat. Struct. Biol.* **1998**, *5*, 60−66.
(66) Masquida, B.; Westhof, E. A modular and hierarchical approach for all-atom RNA modeling. In *The RNA World*, 3rd ed.; Gesteland, R. F.,

Cech, T. R., Atkins, J. F., Eds.; Cold Spring Harbor Laboratory Press: Woodbury, NY, 2005; p 659.
(67) Nicholls, A.; Sharp, K. A.; Honig, B. *Proteins* **1991**, *11*, 281−296.
(68) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620−9631.
(69) Lu, X. J.; Olson, W. K. *Nucleic Acids. Res.* **2003**, *31*, 5108−5121.
(70) Schneider, C.; Suhnel, J. *Biopolymers* **1999**, *50*, 287.
(71) Su, L.; Chen, L.; Egli, M.; Berger, J. M.; Rich, A. *Nat. Stuct. Biol.* **1999**, *6*, 285.
(72) Burkard, M. E.; Kierzek, R.; Turner, D. H. *J. Mol. Biol.* **1999**, *290*, 967.