Investigation of the Structural Basis for Thermodynamic Stabilities of Tandem GU Wobble Pairs: NMR Structures of (rGGAGUUCC)$_2$ and (rGGAUGUCC)$_2$[†,‡]

Jeffrey A. McDowell,[§] Liyan He,[||] Xiaoying Chen, and Douglas H. Turner*

Department of Chemistry, University of Rochester, Rochester, New York 14627-0216

Received January 21, 1997; Revised Manuscript Received April 21, 1997[⊗]

ABSTRACT: The symmetric, tandem GU mismatch motifs, $\begin{smallmatrix}5'-AUGU-3'\\3'-UGUA-5'\end{smallmatrix}$ and $\begin{smallmatrix}5'-AGUU-3'\\3'-UUGA-5'\end{smallmatrix}$, which only differ in the mismatch order, have an average difference in thermodynamic stability of 2 kcal/mol at 37 °C. Thermodynamic studies of duplexes containing these motifs indicate the effect is largely localized to the mismatches and adjacent base pairs. The three-dimensional structures of two representative duplexes, (rGGAGUUCC)$_2$ and (rGGAUGUCC)$_2$, were determined by two-dimensional NMR and a simulated annealing protocol. Local deviations are similar to other intrahelical GU mismatches with little effect on backbone torsion angles and a slight overtwisting between the base pair 5′ of the G of the mismatch and the mismatch itself. Comparisons of the resulting stacking patterns along with electrostatic potential maps suggest that interactions between highly negative electrostatic regions between base pairs may play a role in the observed thermodynamic differences.

GU base pairs are the most common noncanonical interaction in RNA and are also involved in many specific functional roles (Watson et al., 1987). The usual GU wobble pair geometry provides two hydrogen-bonding sites, GN2 and UO4, that go unsatisfied in the base pairing and are accessible in the minor and major grooves, respectively, allowing their use as functional contacts. GU wobble pairs are currently known to be responsible for the recognition of alanine tRNA of *Escherichia coli* by its cognate synthetase (Hou & Schimmel, 1988; McClain & Foss, 1988; Gabriel et al., 1996), for providing vital contacts in the folding of the tertiary structure of the group I intron catalytic core (Strobel & Cech, 1995), and for facilitating RNA editing (Simpson & Thiemann, 1995). Many GU wobble pairs in ribosomal RNA (rRNA) are conserved (Gautheret et al., 1995), implying additional functional roles of this motif are yet to be discovered.

Studies of rRNA have shown that GU wobble pairs frequently occur in tandem, with $\begin{smallmatrix}5'-UG-3'\\3'-GU-5'\end{smallmatrix}$ roughly 7-fold more prevalent than $\begin{smallmatrix}5'-GU-3'\\3'-UG-5'\end{smallmatrix}$ (He et al., 1991; Wu et al., 1995; Gautheret et al., 1995). The thermodynamic stabilities of intrahelical, symmetric, tandem GU wobbles have a strong dependence on the mismatch sequence, with $\begin{smallmatrix}5'-UG-3'\\3'-GU-5'\end{smallmatrix}$ > $\begin{smallmatrix}5'-GU-3'\\3'-UG-5'\end{smallmatrix}$, and on the adjacent base pair, with $\begin{smallmatrix}5'G\\3'C\end{smallmatrix}$ > $\begin{smallmatrix}5'C\\3'G\end{smallmatrix}$ > $\begin{smallmatrix}5'U\\3'A\end{smallmatrix}$ ≥ $\begin{smallmatrix}5'A\\3'U\end{smallmatrix}$ (Wu et al., 1995). Imino proton NMR studies have indicated that most GU wobbles in tandem pairs have two hydrogen bonds and have similar topology (He et al., 1991; McDowell & Turner, 1996). Many studies have shown that incorporation of GU wobbles in helixes only modestly affects the overall helical geometry (White et al., 1992; Allain & Varani, 1995a; McDowell & Turner, 1996). The difference in stability must then be attributed to small, local changes within and around the mismatches. Specifically, the stability differences are probably due to interactions between adjacent base pairs, collectively referred to as stacking interactions.

We recently reported an NMR structure for (rGAGGUCUC)$_2$ which suggested that the 3.5 kcal/mol at 37 °C difference in stability between $\begin{smallmatrix}5'-GGUC-3'\\3'-CUGG-5'\end{smallmatrix}$ and $\begin{smallmatrix}5'-CGUG-3'\\3'-GUGC-5'\end{smallmatrix}$ motifs might be attributable to electrostatic interactions between the bases (McDowell & Turner, 1996). This work tests the importance of electrostatic interactions by comparing the solution structures of (rGGAUGUCC)$_2$ and (rGGAGUUCC)$_2$, whose thermodynamic stabilities differ by 2.0 kcal/mol at 37 °C (He et al., 1991). The sequences of these molecules differ only in the order of wobble pairs: $\begin{smallmatrix}5'-UG-3'\\3'-GU-5'\end{smallmatrix}$ and $\begin{smallmatrix}5'-GU-3'\\3'-UG-5'\end{smallmatrix}$, respectively. Adjacent to the tandem GU pairs are AU pairs that have regions of negative electrostatic potential that are more fragmented and smaller in magnitude than their GC and GU counterparts and therefore are expected to contribute less to the electrostatic portion of stacking energy (Hunter, 1993; McDowell & Turner, 1996). By considering structures with $\begin{smallmatrix}5'-AUGU-3'\\3'-UGUA-5'\end{smallmatrix}$ and $\begin{smallmatrix}5'-AGUU-3'\\3'-UUGA-5'\end{smallmatrix}$ motifs, the electrostatic contribution to the thermodynamic difference between $\begin{smallmatrix}5'-UG-3'\\3'-GU-5'\end{smallmatrix}$ and $\begin{smallmatrix}5'-GU-3'\\3'-UG-5'\end{smallmatrix}$ might be largely isolated to the tandem GU motifs.

MATERIALS AND METHODS

RNA Synthesis, Purification, and Sample Preparation. Oligonucleotides were synthesized and purified as described elsewhere (Usman et al., 1987; McDowell & Turner, 1996). All oligomers were checked by HPLC on a C-8 reverse-phase column (Hamilton) with a 50% methanol mobile phase and found to be better than 98% pure. NMR samples were dissolved in 80 mM NaCl, 10 mM sodium phosphates, and 1 mM Na_2EDTA at pH 7.0 and lyophilized to dryness three times with 99.996% D_2O. Samples were then redissolved

[†] This work was supported by NIH Grant GM 22939.

[‡] Coordinates of converged structures along with the NMR-derived restraints have been deposited in the Brookhaven Protein Data Bank [PDB ID codes 1QES and R1QESMR for (rGGAGUUCC)$_2$ and 1QET and R1QETMR for (rGGAUGUCC)$_2$].

* Author to whom correspondence should be addressed.

[§] Current address: Abbott Laboratories, Building AP10, 100 Abbott Park Rd., Abbott Park, IL 60064.

[||] Current address: Chromagen, Inc., 10441-A Roselle St., San Diego, CA 92121-1503.

[⊗] Abstract published in *Advance ACS Abstracts,* June 1, 1997.

S0006-2960(97)00122-0 CCC: $14.00

Table 1: Thermodynamics of Duplex Formation in 1 M NaCl and, in Parentheses, in 0.1 M KCl and 10 mM $MgCl_2$

	melt curve fitting parameters				$1/T_m$ vs ln C_T parameters			
sequence	$\Delta G°_{37}$ (kcal/mol)	$\Delta H°$ (kcal/mol)	$\Delta S°$ [cal/(K·mol)]	T_m[a] (°C)	$\Delta G°_{37}$ (kcal/mol)	$\Delta H°$ (kcal/mol)	$\Delta S°$ [cal/(K·mol)]	T_m[a] (°C)
GGAUGUCC	-8.6 ± 0.2[b]	-78.0 ± 8[b]	-223.4 ± 22[b]	49.0[b]	-8.4 ± 0.2[b]	-73.0 ± 7[b]	-208.4 ± 21[b]	49.0[b]
	(-8.5 ± 0.2)	(-79.3 ± 4.5)	(-228.4 ± 15)	(48.5)	(-8.8 ± 0.1)	(-90.6 ± 3.3)	(-263.8 ± 10)	(47.9)
GGAGUUCC	-6.4 ± 0.1[b]	-68.4 ± 7[b]	-199.6 ± 20[b]	40.5[b]	-6.4 ± 0.1[b]	-73.1 ± 7[b]	-214.9 ± 21[b]	40.2[b]
	(-6.3 ± 0.1)	(-70.1 ± 8.8)	(-205.7 ± 29)	(39.8)	(-6.1 ± 0.1)	(-70.5 ± 3.2)	(-207.6 ± 10)	(39.0)
CCAUGUGG	-7.8 ± 0.1	-71.2 ± 6.1	-204.3 ± 10.1	46.9	-7.8 ± 0.3	-70.5 ± 1.1	-202.1 ± 3.6	46.5
	(-8.1 ± 0.1)	(-76.0 ± 2.0)	(-218.9 ± 6.4)	(47.3)	(-8.3 ± 0.1)	(-84.7 ± 3.1)	(-246.3 ± 9.7)	(47.1)
CCAGUUGG	-5.8 ± 0.1	-60.4 ± 5.4	-176.3 ± 4.2	37.4	-5.7 ± 0.1	-61.1 ± 1.2	-178.6 ± 3.9	37.1
	(-6.0 ± 0.1)	(-64.2 ± 2.9)	(-187.5 ± 9.3)	(38.5)	(-6.0 ± 0.1)	(-59.8 ± 1.0)	(-173.5 ± 3.1)	(38.5)
CCAUGG	-7.5 ± 0.2	-61.9 ± 5.1	-175.3 ± 11.0	46.4	-7.3 ± 0.1	-56.7 ± 2.2	-159.0 ± 6.7	46.3
	(-7.6 ± 0.2)	(-57.2 ± 3.6)	(-160.2 ± 11.6)	(47.7)	(-7.7 ± 0.1)	(-63.7 ± 4.2)	(-180.5 ± 13)	(47.3)

[a] Calculated for an oligomer concentration of 10^{-4} M. [b] He et al., 1991.

under dry nitrogen in 99.996% D_2O with enough volume to give final oligomer concentrations of approximately 3.5 mM.

Thermodynamic Measurements and Analysis. Absorbance versus temperature melting curves were measured at 280 nm with a heating rate of 1 °C/min and collected on a Gilford 250 spectrometer. Thermodynamic parameters were measured using a buffer system of 1.0 M NaCl, 10 mM sodium cacodylate, and 0.50 mM Na_2EDTA at pH 7.0 or 0.1 M KCl, 10 mM $MgCl_2$, 10 mM sodium cacodylate, and 0.50 mM Na_2EDTA at pH 7.0 and calculated using the Meltwin analysis program by methods described elsewhere (McDowell & Turner, 1996).

NMR Spectroscopy and Structure Refinement. All 2D NMR spectra were recorded in the phase-sensitive mode using the States−Haberkorn method (States et al., 1982). NMR parameters, restraint set generation methods, and the constraint satisfaction protocol used were the same as described in McDowell and Turner (1996). All NMR processing and structure calculations were performed using Biosym's Felix, Insight, and Discover software running on Silicon Graphics ONYX and 4D/35 workstations. The quality of the final structures was determined by calculating the average all-atom pairwise root mean squared deviations (RMSD) of the converged structures. Structural accuracy was assessed by considering the number and magnitude of constraint violations and by calculating theoretical NOESY spectra using the final structures and a relaxation matrix approach (RMA) (Boelens et al., 1988, 1989). RMA calculations were performed with an isotropic correlation time of 3 ns and quantitatively compared to the experimental NOESY spectra with mixing times of 100, 150, 200, and 400 ms by calculating three *R*-factors (Gonzalez et al., 1991) for spin pairs with well-determined cross peaks:

$$R_1 = \frac{\sum_{\text{spectra}}\sum_{i,j}|A_{ij}^{\text{calc}}(\tau) - A_{ij}^{\text{exp}}(\tau)|}{\sum_{\text{spectra}}\sum_{i,j}(A_{ij}^{\text{exp}}(\tau))} \quad (1)$$

$$R_\tau = \frac{\sum_{\text{spectra}}\tau\sum_{i,j}|A_{ij}^{\text{calc}}(\tau) - A_{ij}^{\text{exp}}(\tau)|}{\sum_{\text{spectra}}\tau\sum_{i,j}(A_{ij}^{\text{exp}}(\tau))} \quad (2)$$

$$R_2 = \frac{\sqrt{\sum_{\text{spectra}}\sum_{i,j}|A_{ij}^{\text{calc}}(\tau) - A_{ij}^{\text{exp}}(\tau)|^2}}{\sum_{\text{spectra}}\sum_{i,j}(A_{ij}^{\text{exp}}(\tau))^2} \quad (3)$$

where $A_{ij}(\tau)$ is the NOE intensity of the spin pair *ij* at mixing time τ.

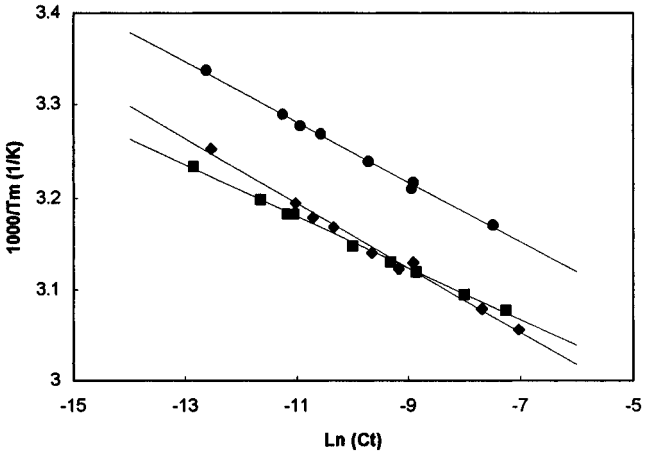

FIGURE 1: Plots of T_m^{-1} vs ln C_T: CCAGUUGG (●), CCAUGUGG (■), and CCAUGG (◆).

RESULTS

Thermodynamic Stability of Tandem GU Mismatches with Adjacent AU Base Pairs. To test the generality of the thermodynamic increments reported by He et al. (1991), thermodynamic parameters for duplex formation were measured for (rCCAUGUGG)$_2$, (rCCAGUUGG)$_2$, and (rCCAUGG)$_2$ (Figure 1, Table 1). The thermodynamic increments for the tandem mismatches (Table 2) are defined as in the example:

$$\Delta G°_{37}(\text{AGUU}) = \Delta G°_{37}(\text{CCAGUUGG}) - \Delta G°_{37}(\text{CCAUGG}) + \Delta G°_{37}(\text{AU}) \quad (4)$$

where $\Delta G°_{37}$(CCAGUUGG) and $\Delta G°_{37}$(CCAUGG) are the free energy changes of duplex formation derived from T_m^{-1} versus ln C_T plots for the strand sequences shown in parentheses and $\Delta G°_{37}$(AU) is the free energy increment (Freier et al., 1986) for the nearest-neighbor interaction interrupted by insertion of the tandem GU mismatch. This value reflects the stability provided by the incorporation of this motif into a duplex. For both pairs of sequences that have been studied, the $^{5'-\text{AUGU}-3'}_{3'-\text{UGUA}-5'}$ motif stabilizes the duplex whereas $^{5'-\text{AGUU}-3'}_{3'-\text{UUGA}-5'}$ destabilizes it.

It has recently been shown that tandem GU wobbles can serve as binding sites for Mg^{2+} (Cate & Doudna, 1996; Cate et al., 1996) and that single GU wobbles can bind Mn^{2+} (Limmer et al., 1993; Ott et al., 1993; Allain & Varani, 1995b). To test whether the thermodynamic measurements at 1 M NaCl are a reasonable approximation for buffers

Table 2: Thermodynamic Increments for Tandem GU Mismatches with Adjacent AU Base Pairs in 1 M NaCl[a]

sequence	$\Delta G°_{37}$ (kcal/mol)	$\Delta H°$ (kcal/mol)	$\Delta S°$ [cal/(K·mol)]	sequence	$\Delta G°_{37}$ (kcal/mol)	$\Delta H°$ (kcal/mol)	$\Delta S°$ [cal/(K·mol)]
GGAUGUCC[b]	−1.9	−25.0	−74.8	GGAGUUCC[b]	+0.1	−25.1	−81.4
CCAUGUGG	−1.4	−19.5	−58.6	CCAGUUGG	+0.7	−10.1	−35.1

[a] Calculated as $\Delta G°_{37}$ for AUGU: $\Delta G°_{37}(\text{AUGU}) = \Delta G°_{37}(\text{GGAUGUCC}) - \Delta G°_{37}(\text{GGAUCC}) + \Delta G°_{37}(\text{AU})$. Thermodynamic parameters were derived from T_m^{-1} vs ln C_T plots. [b] He et al., 1991.

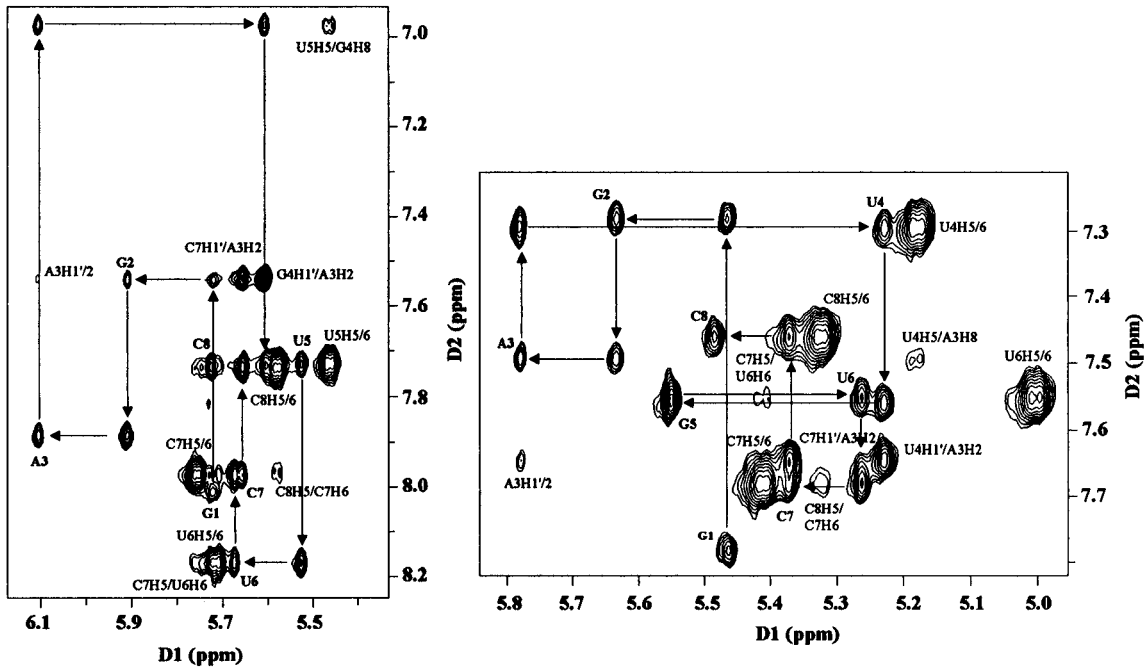

FIGURE 2: 400 ms NOESY spectra of (rGGAGUUCC)$_2$ at 20 °C (left) and (rGGAUGUCC)$_2$ at 30 °C (right) showing the H1′/5−H8/6/2 region. Base$_n$−H1′$_n$ cross peaks are labeled as well as peaks not in the H8/6$_n$−H1′$_n$−H8/6$_{n+1}$ NOE connectivity pathway.

Table 3: Chemical Shifts[a] (ppm) of Nonexchangeable and Imino Protons and Phosphorus Resonances of (rGGAGUUCC)$_2$

	H8/H6	H2/H5	H1′	H2′	H3′	H4′	H5′/H5″	phosphorus
G1	7.98	na	5.68	4.84	4.52	4.05	3.89/3.99	na
G2	7.52	na	5.89	4.64	4.73	5.54	4.60/4.17	−3.00
A3	7.87	7.52	6.09	4.73	4.70	4.52		−2.88
G4	6.96	na	5.59	4.79	4.01	4.52	3.98/4.33	−2.58
U5	7.71	5.44	5.51	4.18	4.62	4.45		−3.56
U6	8.17	5.71	5.66	4.60	4.58	4.48		−3.84
C7	7.97	5.74	5.64	4.31	4.51	4.46		−3.28
C8	7.73	5.57	5.71	4.06	4.19	4.60	4.47/4.14	−3.24

[a] Proton chemical shifts are referenced to an external TSP [3-(trimethylsilyl)propionate] reference at 20 °C. Phosphorus chemical shifts are referenced to internal phosphate buffer.

Table 4: Chemical Shifts[a] (ppm) of Nonexchangeable and Imino Protons and Phosphorus Resonances of (rGGAUGUCC)$_2$

	H8/H6	H2/H5	H1′	H2′	H3′	H4′	H5′/H5″	phosphorus
G1	7.77	na	5.46	4.61	4.33	4.08		na
G2	7.27	na	5.62	4.47	4.42	4.59		−2.82
A3	7.48	7.63	5.77	4.40	4.30	4.42		−2.86
U4	7.28	5.17	5.22	4.13	4.26	4.20		−3.22
G5	7.54	5.56	5.54	4.43	4.08	4.46		−3.19[b]
U6	7.54	4.99	5.25	4.19	4.28	4.47		−3.82
C7	7.66	5.40	5.36	4.04	4.26	4.45		−3.32
C8	7.45	5.32	5.48	3.79	3.94	4.37		−2.95[b]

[a] Proton chemical shifts are referenced to an external TSP [3-(trimethylsilyl)propionate] reference at 30 °C. Phosphorus chemical shifts are referenced to internal phosphate buffer. [b] Unable to distinguish between G5P and C8P from couplings to U4H3′ and C7H3′.

containing Mg^{2+}, thermodynamic parameters were measured in 10 mM $MgCl_2$ and 0.1 M KCl for (rCCAUGUGG)$_2$, (rCCAGUUGG)$_2$, (rCCAUGG)$_2$, (rGGAUGUCC)$_2$, and (rGGAGUUCC)$_2$. These results are listed in parentheses in Table 1 and show that the thermodynamics in 10 mM Mg^{2+} are similar to those in 1 M NaCl.

Nonexchangeable Protons and Phosphorus Resonance Assignments. Assignment of nonexchangeable proton and phosphorus resonances was based on NOESY spectra at five mixing times (60, 100, 150, 200, 400 ms), along with DQF-COSY and $^1H/^{31}P$ HETCOR spectra using methods outlined in Varani and Tinoco (1991). These assignments are summarized in Tables 3 and 4.

The imino proton and A3H2 resonances of (rGGAGUUCC)$_2$ and (rGGAUGUCC)$_2$ have been assigned previously (He et al., 1991). The AH2 assignments provided initial reference points for beginning the total assignment process. In A-form helixes, AH2 protons in an A_n·U_m base pair show cross peaks to the intrastrand H1′$_{n+1}$ and interstrand H1′$_{m+1}$. This permitted assignment of C7H1′ and G4H1′ or U4H1′ for (rGGAGUUCC)$_2$ and (rGGAUGUCC)$_2$, respectively. In addition, each molecule has four pyrimidine nucleotides that each provide an H5/H6 cross peak in this region. Each of these cross peaks can be identified by its splitting and its existence in the DQF-COSY spectra with a characteristic *J*-coupling of ∼6 Hz. Figure 2 shows the aromatic/anomeric region of the 400 ms NOESY spectra of (rGGAGUUCC)$_2$ and (rGGAUGUCC)$_2$ at 20 and 30 °C, respectively, with the usual sequential H8/6$_n$−H1′$_n$−H8/6$_{n+1}$ NOE connectivity pathways (Scheek et al., 1983; Feigon et al., 1983; Petersheim & Turner, 1983; Hare et al., 1983).

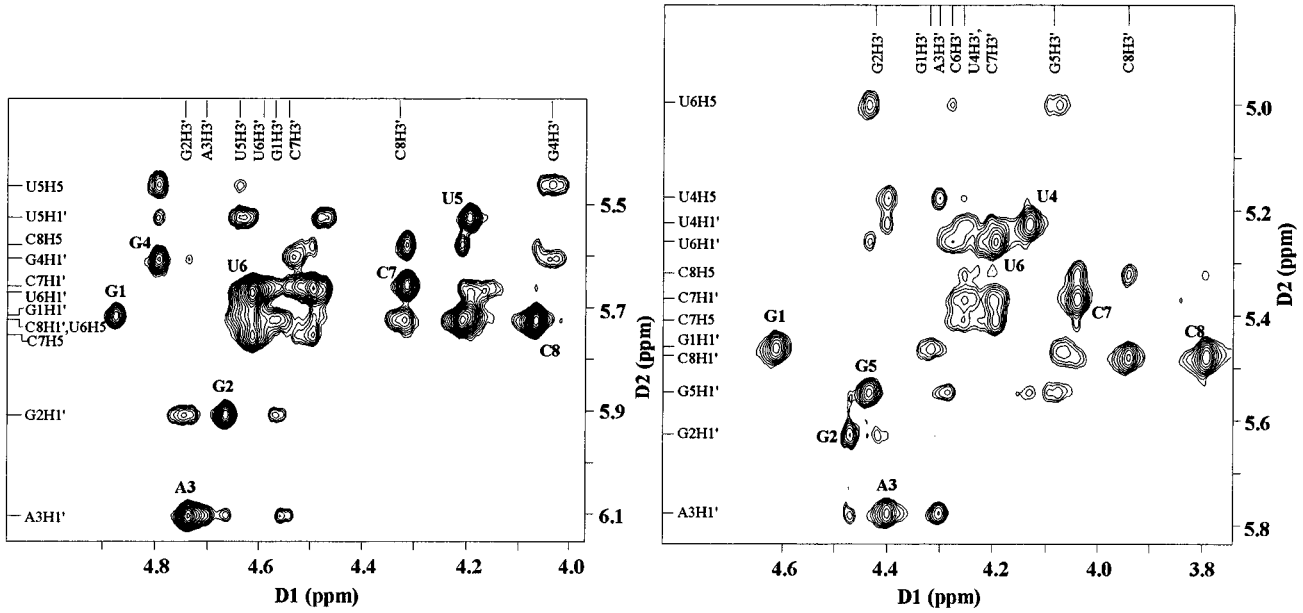

FIGURE 3: 400 ms NOESY spectra of (rGGAGUUCC)$_2$ at 20 °C (left) and (rGGAUGUCC)$_2$ at 30 °C (right) showing the H1′/5−sugar (H2′, H3′, H4′, H5′, H5″) region. Labels designate H1′−H2′ cross peaks.

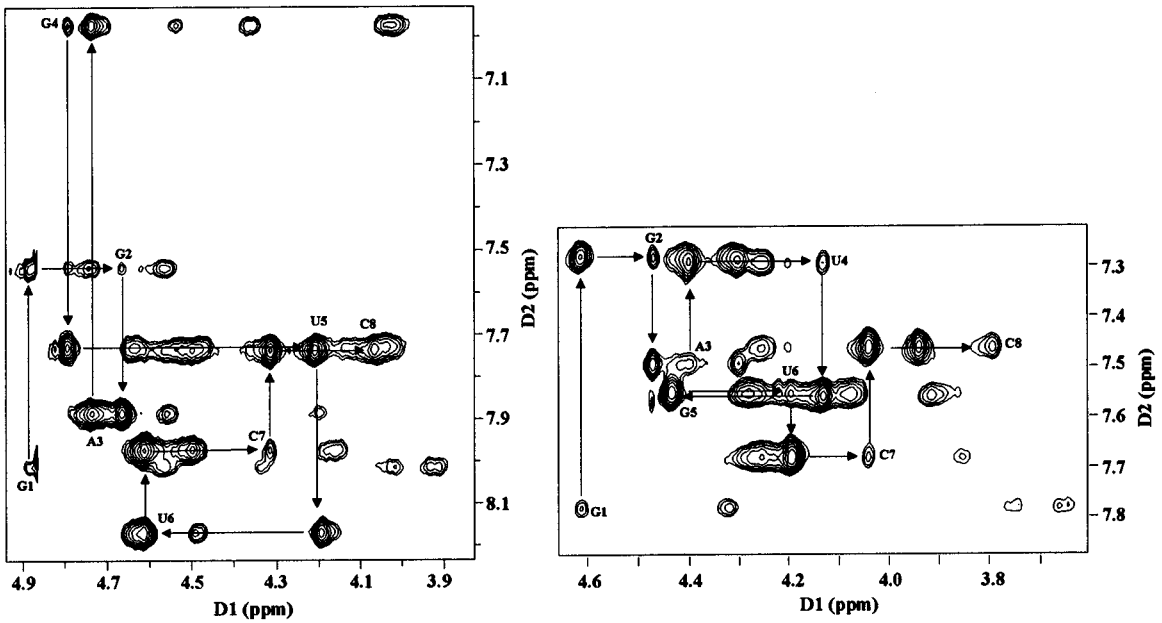

FIGURE 4: 400 ms NOESY spectra of (rGGAGUUCC)$_2$ at 20 °C (left) and (rGGAUGUCC)$_2$ at 30 °C (right) showing the sugar (H2′, H3′, H4′, H5′, H5″)−H8/6/2 region. Base$_n$−H2′$_n$ cross peaks are labeled.

The cross peak patterns and intensities verify that the structures are generally A-form and all nucleotides have *anti* glycosidic torsion angles.

Assignments of H2′ resonances followed from short mixing time (60, 100, 150 ms) NOESY spectra, so spin diffusion would not significantly contribute to NOE intensities (Neuhaus & Williamson, 1989). H1′/H2′ distances vary from 2.3 to 2.4 Å, largely independent of sugar pucker, so the eight strongest peaks were assigned to H1′/H2′ cross peaks. NOESY spectra of these regions at 150 ms mixing time are shown in the Supporting Information (see paragraph at end of paper regarding Supporting Information). Figure 3 shows these regions in the 400 ms NOESY spectra. The H2′ assignments were confirmed by sequential H8/6$_n$−H2′$_n$−H8/6$_{n+1}$ NOE connectivity pathways (Figure 4) which show very strong H2′$_n$−H8/6$_{n+1}$ cross peaks typical of A-form conformations.

H3′ and H4′ proton resonances were tentatively assigned by their NOE's with H1′ protons. H3′/H1′ distances are essentially independent of sugar pucker at ∼3.9 Å and result in an observable NOE in longer mixing time NOESY spectra (200 and 400 ms). These cross peaks are relatively broad due to the strong coupling between H3′ and the backbone phosphate. H4′/H1′ distances vary with sugar pucker but in C3′-endo sugars have a distance of ∼2.9 Å. H3′ cross peaks were confirmed by their coupling with the backbone phosphate observed in ^{1}H/^{31}P HETCOR spectra (Figure 5). This also provided assignments of the phosphorus resonances. H3′ and H4′ assignments were further confirmed using DQF-COSY spectra. H3′ shows a characteristic *J*-coupling with H2′ of 5 Hz, which is independent of sugar pucker. For C3′-endo ribose conformations, this cross peak shows a characteristic phase pattern. All nucleotides for both molecules show this C3′-endo H2′/H3′ cross peak phase

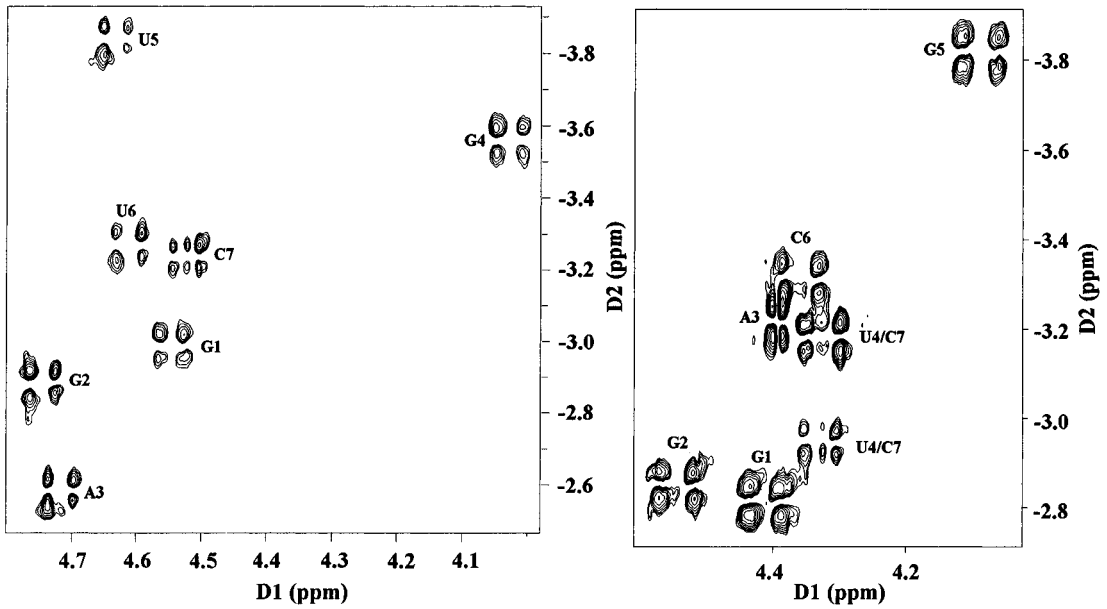

FIGURE 5: Proton-detected $^{1}H/^{31}P$ HETCOR spectra of (rGGAGUUCC)$_2$ at 20 °C (left) and (rGGAUGUCC)$_2$ at 30 °C (right) with the H3′$_n$−P$_{n+1}$ cross peaks labeled.

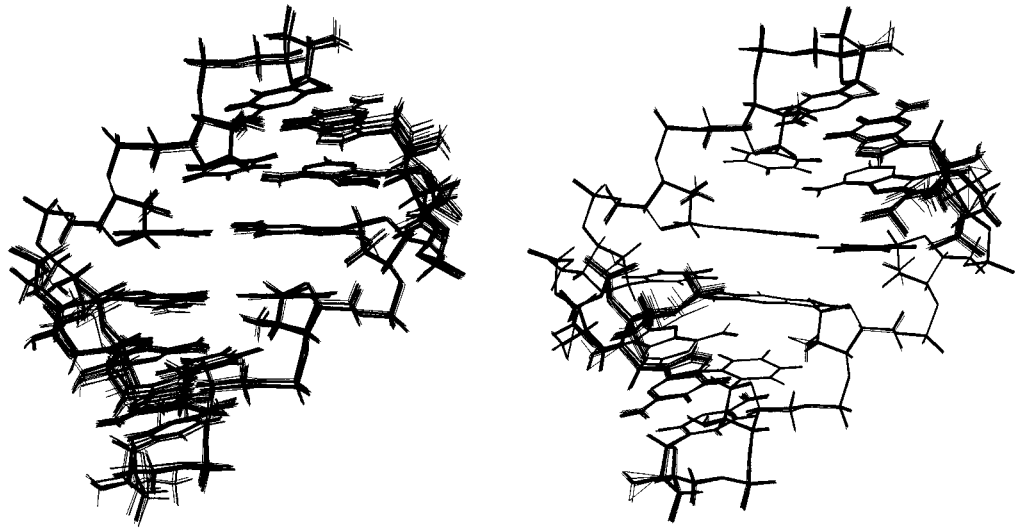

FIGURE 6: Superposition of the 30 converged structures of (rGGAGUUCC)$_2$ (left) and (rGGAUGUCC)$_2$ (right) showing the internal nucleotides. The average RMSD for the pairwise all-atom superposition for the full set of converged structures for each entire molecule is 0.50 and 0.37 Å, respectively. The RMSD calculated for the internal nucleotides is 0.40 and 0.27 Å, respectively.

pattern except for both C8 nucleotides which presumably have considerable dynamics. H4′ proton resonance assignments were confirmed by the large H3′/H4′ coupling.

Structure Calculations. Thirty initial structures for each molecule were used with one A-form, one B-form, and 28 generated with random torsion angle conformations. These, along with the appropriate constraints as listed in Supporting Information, were used in the protocol reported in McDowell and Turner (1996). This procedure resulted in 30 converged structures per molecule as shown in Figure 6. The RMSD values for the entire molecules were 0.50 and 0.37 Å for (rGGAGUUCC)$_2$ and (rGGAUGUCC)$_2$, respectively. Further breakdown of the RMSD values is reported in Table 5.

Each final structure satisfied all distance restraints within 0.15 Å and torsion restraints within <2°. A more rigorous analysis of the structural accuracy is to calculate NMR *R*-factors which indicate not only if the final structure conforms to the proton/proton distance restraint set but also whether it is consistent with the entire proton relaxation network that generates the time-dependent NOESY spectra. Table 6 shows the results of these calculations for a typical final structure compared with three of the starting structures. The low *R*-factors of both molecules reflect good agreement between the calculated and experimental NOE's.

Table 5: All-Atom Pairwise RMSD Values (Å) for Molecules and Base Pairs

	(rGGAGUUCC)$_2$	(rGGAUGUCC)$_2$
full molecule	0.50	0.37
nonterminal nucleotides	0.40	0.27

base pairs	(rGGAGUUCC)$_2$	base pairs	(rGGAUGUCC)$_2$
G1·C8	0.65	G1·C8	0.49
G2·C7	0.42	G2·C7	0.31
A3·U6	0.40	A3·U6	0.22
G4·U5	0.31	U4·G5	0.26
U5·G4	0.32	G5·U4	0.24
U6·A3	0.35	U6·A3	0.22
C7·G2	0.46	C7·G2	0.31
C8·G1	0.64	C8·G1	0.50

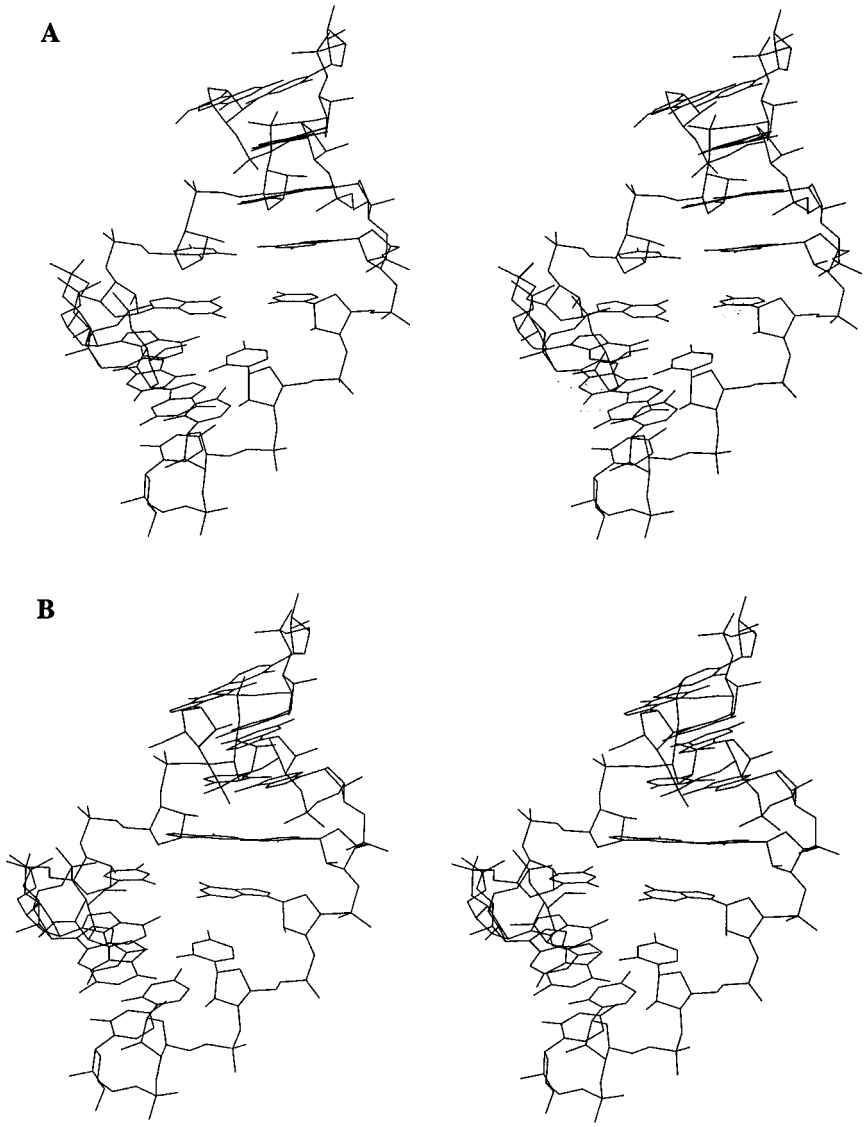

FIGURE 7: Stereoviews of one representative structure of the 30 converged structures of (rGGAGUUCC)$_2$ (A) and (rGGAUGUCC)$_2$ (B).

Table 6: *R*-Factors of the Final and Several Starting Structures from the Structure Determination of (rGGAGUUCC)$_2$ and (rGGAUGUCC)$_2$[a]

	(rGGAGUUCC)$_2$			(rGGAUGUCC)$_2$		
	R_1	R_2	R_r	R_1	R_2	R_r
final structure[b]	0.47	0.22	0.45	0.42	0.43	0.19
A-form	0.83	0.73	0.96	0.66	0.69	0.36
B-form	0.72	0.74	0.60	0.87	0.89	0.90
random structure **1**[c]	1.3	1.4	1.7	1.1	1.3	1.6
random structure **2**[c]	1.3	1.6	2.0	1.3	1.3	2.3

[a] *R*-factors as defined in eqs 1−3. [b] A typical example of the 30 converged structures is shown in Figure 8. [c] Starting structures with randomized torsion angles are included for comparison.

Structural Analysis. Stereoviews of one representative of the converged structures are shown in Figure 7. Figure 8 shows views of the major and minor grooves of the averages of the converged structures in CPK form. The structures are primarily A-form, as indicated by the backbone torsion angles and helical parameters reported in Tables 7−12. Each structure has all nucleotides with N-type sugars and *anti* glycosidic bonds. Local deviations for both molecules are similar to those seen in other intrahelical GU wobble structures (White et al., 1992; Allain & Varani, 1995a; McDowell & Turner, 1996). For (rGGAGUUCC)$_2$, there is ~7° of overtwisting between the base pair 5′ of the G of the GU wobbles and the GU itself. The result of this overtwisting is that there is significantly more intrastrand overlap of the U of a GU wobble toward its 5′ side than its 3′ side and of the G toward its 3′ side than its 5′ side (Mizuno & Sundaralingam, 1978). This stacking pattern is also observed with (rGGAUGUCC)$_2$. Thus, for (rGGAGUUCC)$_2$, the tandem wobble pairs have considerable intrastrand overlap with each other but little overlap with the adjacent Watson−Crick base pairs (Figure 9B,C). In contrast, (rGGAUGUCC)$_2$ has interstrand overlap between the wobble pairs and intrastrand overlap between the wobble pairs and their adjacent Watson−Crick pairs (Figure 9E,F).

DISCUSSION

Tandem GU wobble pairs frequently occur in known secondary structures of RNA, with the $\begin{smallmatrix}5'\text{–}UG\text{–}3'\\3'\text{–}GU\text{–}5'\end{smallmatrix}$ motif roughly 7-fold more common than $\begin{smallmatrix}5'\text{–}GU\text{–}3'\\3'\text{–}UG\text{–}5'\end{smallmatrix}$ (He et al., 1991; Wu et al., 1995; Gautheret et al., 1995). This difference in prevalence of the symmetric, tandem GU motifs may be related to thermodynamic stability. A nearest-neighbor analysis of stabilities of short duplexes found that $\begin{smallmatrix}5'\text{–}UG\text{–}3'\\3'\text{–}GU\text{–}5'\end{smallmatrix}$ is 1.7 kcal/mol more stable at 37 °C than $\begin{smallmatrix}5'\text{–}GU\text{–}3'\\3'\text{–}UG\text{–}5'\end{smallmatrix}$, when duplexes containing the motif $\begin{smallmatrix}5'\text{–}GGUC\text{–}3'\\3'\text{–}CUGG\text{–}5'\end{smallmatrix}$

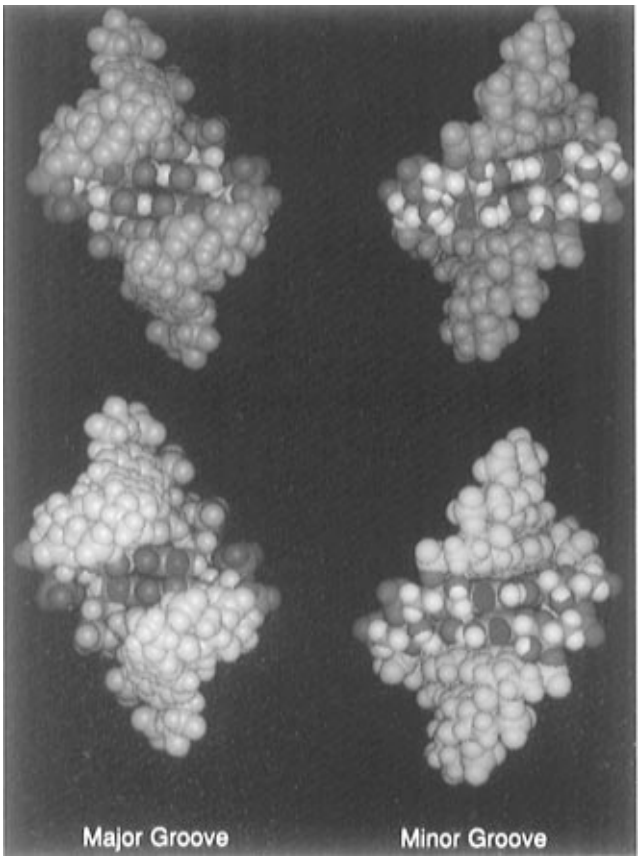

FIGURE 8: CPK models of the averages of the converged structures of (rGGAGUUCC)$_2$ (top) and (rGGAUGUCC)$_2$ (bottom) showing the major and minor grooves.

Table 7: Helical Parameters[a] for Base Pair Steps of the Averaged Structure of (rGGAGUUCC)$_2$

base step	helical twist (deg)	rise (Å)	slide (Å)	roll (deg)
GG/CC	30.9	3.0	−1.7	3.9
GA/CU	30.1	3.2	−1.8	5.9
AG/UU	40.0	2.8	−1.6	10.0
GU/UG	24.0	2.9	−2.6	7.2
UU/GA	40.0	2.8	−1.6	11.0
UC/AG	30.2	3.2	−1.8	5.9
CC/GG	30.7	3.0	−1.7	3.6
av	32.3	3.0	−1.8	6.8
A-form	32.7	2.8	−1.5	−0.4

[a] Parameters were calculated as described by the Cambridge convention (Dickerson, 1989).

are omitted (He et al., 1991). Sequences containing the 5′−GGUC−3′ / 3′−CUGG−5′ motif were omitted because they did not fit the nearest-neighbor model. The 1.7 kcal/mol difference in free energy increments is the largest observed thus far for a simple reversal of base pairs. The additional thermodynamics in Table 1 are consistent with this difference in stability. The duplex (rCCAUGUGG)$_2$ is 2 kcal/mol more stable than (rCCAGUUGG)$_2$ at 37 °C. This is also consistent with the previously reported result that (rGGAUGUCC)$_2$ is 2 kcal/mol more stable than (rGGAGUUCC)$_2$ at 37 °C (He et al., 1991). Evidently, the stability differences are largely localized to the tandem GU motifs and are not dependent on the sequences of the surrounding helixes.

Table 8: Helical Parameters[a] for Base Pairs of the Averaged Structure of (rGGAGUUCC)$_2$

base	propeller twist (deg)	buckle (deg)	displacement (Å)	interstrand C1′$_n$−C1′$_m$ distance[b] (Å)	intrastrand P$_n$−P$_{n+1}$ distance[c] (Å)
G1·C8	−5.1	11.0	−3.8	10.7	
G2·C7	5.0	0.5	−4.7	10.6	5.9
A3·U6	3.0	−6.3	−3.5	10.6	6.1
G4·U5	−7.5	0.1	−3.5	10.8	5.8
U5·G4	−7.5	0.6	−3.5	10.8	5.7
U6·A3	2.8	6.5	−4.7	10.6	6.1
C7·G2	5.1	−0.2	−3.8	10.6	6.0
C8·G1	−4.7	−11.0	−4.5	10.7	
av	−1.1	0.2	−4.0	10.7	5.9
A-form	−13.8	0.2	−4.4	10.9	5.9

[a] Parameters were calculated as described by the Cambridge convention (Dickerson, 1989). [b] Interstand C1′−C1′ distance is between the C1′ atoms of the paired bases. [c] Intrastrand P−P distance is between consecutive phosphates of the same strand. The value for each pair of phosphates is identical for both strands.

Table 9: Backbone and Glycosidic Torsion Angles of the Averaged Structure of (rGGAGUUCC)$_2$

base	α	β	γ	δ	ϵ	ζ	χ
G1	na	na	56.4	77.6	−170.3	−66.7	−162.6
G2	−75.4	−175.9	57.5	77.7	−169.4	−63.0	−158.7
A3	−74.3	−178.5	58.5	82.5	−162.4	−57.9	−165.6
G4	−74.7	−177.7	53.0	75.9	−164.2	−55.8	−164.1
U5	−73.9	178.8	57.8	84.4	−171.6	−76.3	−155.8
U6	−82.0	−172.5	53.9	81.4	−169.6	−61.8	−159.5
C7	−90.4	−176.2	65.6	80.2	−167.6	−61.7	−159.7
C8	−77.5	178.7	57.0	79.5	na	na	−159.1
av	−78.3	−177.6	57.5	79.9	−167.9	−63.3	−160.6
A-form[a]	−68	178	54	82	−153	−71	−158

[a] Saenger, 1984. These values differ from those listed in McDowell and Turner (1996), which were measured from an A-form built with the nucleic acid library from the biopolymer module in Biosym's Insight.

Table 10: Helical Parameters[a] for Base Pair Steps of the Averaged Structure of (rGGAUGUCC)$_2$

base step	helical twist (deg)	rise (Å)	slide (Å)	roll (deg)
GG/CC	28.0	3.0	−1.8	1.1
GA/CU	30.0	2.7	−2.1	7.7
AU/UG	28.0	3.0	−1.8	10.5
UG/GU	33.8	2.8	−1.5	11.0
GU/UA	28.0	3.0	−1.8	10.5
UC/AG	30.4	2.7	−2.1	7.9
CC/GG	28.3	3.1	−1.9	1.5
av	29.5	2.9	−1.9	7.2
A-form	32.7	2.8	−1.5	−0.4

[a] Parameters were calculated as described by the Cambridge convention (Dickerson, 1989).

To investigate the structural basis for the difference in stability between 5′−UG−3′ / 3′−GU−5′ and 5′−GU−3′ / 3′−UG−5′, solution structures of (rGGAUGUCC)$_2$ and (rGGAGUUCC)$_2$ have been determined by NMR. Both structures are similar to related crystal and NMR-derived structures. Specifically, comparisons of stacking overlaps of the 5′−AG−3′ / 3′−UU−5′ and 5′−GU−3′ / 3′−UG−5′ steps from (rGGAGUUCC)$_2$ are similar to the corresponding 5′−GG−3′ / 3′−CT−5′ and 5′−GT−3′ / 3′−TG−5′ stacking pairs in the A-DNA crystal structure

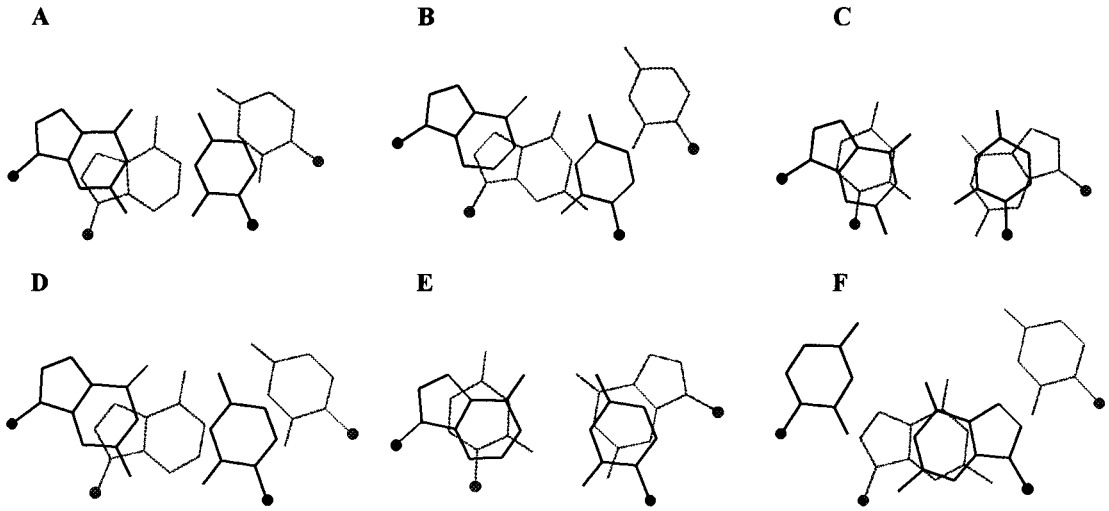

FIGURE 9: Stacking patterns of (A−C) (rGGAGUUCC)$_2$ and (D−F) (rGGAUGUCC)$_2$ showing the stacking geometries of the internal stacking steps: (A) $\begin{smallmatrix}5'-GA-3'\\3'-CU-5'\end{smallmatrix}$, (B) $\begin{smallmatrix}5'-AG-3'\\3'-UU-5'\end{smallmatrix}$, (C) $\begin{smallmatrix}5'-GU-3'\\3'-UG-5'\end{smallmatrix}$, (D) $\begin{smallmatrix}5'-GA-3'\\3'-CU-5'\end{smallmatrix}$, (E) $\begin{smallmatrix}5'-AU-3'\\3'-UG-5'\end{smallmatrix}$, and (F) $\begin{smallmatrix}5'-UG-3'\\3'-GU-5'\end{smallmatrix}$. The first base pair in each sequence is in bold. The $\begin{smallmatrix}5'-AG-3'\\3'-UC-5'\end{smallmatrix}$ steps have essentially the same stacking geometry as A-form helixes (Saenger, 1989). The steps 5′ of the G of the GU mismatch show little intrastrand overlap while those steps 3′ of the mismatched G show considerable overlap. The $\begin{smallmatrix}5'-UG-3'\\3'-GU-5'\end{smallmatrix}$ step (F) also shows an interstrand stack similar to that seen in the A-DNA crystal structure of (dGCGTGCGC)$_2$ (Rabinovich et al., 1988).

Table 11: Helical Parameters[a] for Base Pairs of the Averaged Structure of (rGGAUGUCC)$_2$

base	propeller twist (deg)	buckle (deg)	displacement (Å)	interstrand C1′$_n$−C1′$_m$ distance[b] (Å)	intrastrand P$_n$−P$_{n+1}$ distance[c] (Å)
G1·C8	0.5	10.7	−3.0	10.7	
G2·C7	−0.1	−3.7	−3.8	10.6	5.9
A3·U6	−10.5	9.1	−4.9	10.6	6.1
U4·G5	−5.2	5.9	−4.6	10.8	5.8
G5·U4	−3.3	−5.5	−4.1	10.8	5.7
U6·A3	−10.5	−9.7	−4.0	10.6	6.1
C7·G2	−0.8	3.8	−3.9	10.6	6.0
C8·G1	−0.3	−10.1	−5.1	10.7	
av	−3.8	0.1	−4.2	10.7	5.9
A-form	−13.8	0.2	−4.4	10.9	5.9

[a] Parameters were calculated as described by the Cambridge convention (Dickerson, 1989). [b] Interstand C1′−C1′ distance is between the C1′ atoms of the paired bases. [c] Intrastrand P−P distance is between consecutive phosphates of the same strand. The value for each pair of phosphates is identical for both strands.

Table 12: Backbone and Glycosidic Torsion Angles of the Averaged Structure of (rGGAUGUCC)$_2$

base	α	β	γ	δ	ϵ	ζ	χ
G1	na	na	60.0	81.8	−170.5	−71.0	−163.6
G2	−74.1	−178.9	59.2	82.2	−161.9	−63.8	−170.7
A3	−72.4	179.0	62.5	79.3	−125.7	−71.2	−171.0
U4	−71.3	168.4	55.7	76.1	−171.7	−67.7	−161.5
G5	−76.5	−179.8	56.7	77.0	−166.5	−67.1	−161.5
U6	−80.9	−176.3	53.9	78.7	−167.3	−60.4	−151.2
C7	−75.9	−177.8	55.5	77.4	−166.8	−66.8	−166.6
C8	−75.1	−177.5	52.4	78.0	na	na	−157.6
av	−75.2	179.6	57.0	78.8	−161.5	−66.9	−163.0
A-form	−68	178	54	82	−153	−71	−158

[a] Saenger, 1984. These values differ from those listed in McDowell and Turner (1996), which were measured from an A-form built with the nucleic acid library from the biopolymer module in Biosym's Insight.

of (dGGGGTCCC)$_2$ (Kneale et al., 1985), the NMR structure of (rGAGGUCUC)$_2$ (McDowell & Turner, 1996), and the $\begin{smallmatrix}5'-GGUC-3'\\3'-CUGG-5'\end{smallmatrix}$ motif in the crystal structure of a part of a group I intron (Cate et al., 1996). The $\begin{smallmatrix}5'-AU-3'\\3'-UG-5'\end{smallmatrix}$ and $\begin{smallmatrix}5'-UG-3'\\3'-GU-5'\end{smallmatrix}$ steps of the (rGGAUGUCC)$_2$ structure are likewise nearly identical to the $\begin{smallmatrix}5'-GT-3'\\3'-CG-5'\end{smallmatrix}$ and $\begin{smallmatrix}5'-TG-3'\\3'-GT-5'\end{smallmatrix}$ overlapped base pairs in the A-DNA crystal structure of (dGCGTGCGC)$_2$ (Rabinovich et al., 1988). This includes the observed interstrand stacking between the wobble guanines (Figure 9F). All the GU and GT mismatches also have the same two hydrogen bonds. Evidently, the structures of these A-form helixes are not significantly altered by changing a 2′OH to 2′H or by changing the base pairs, and furthermore, A-form structures in crystals are similar to those in solution. Thus these structures are relatively rigid. This might be expected since the mismatches each have two hydrogen bonds and are constrained by adjacent Watson−Crick base pairs.

Both the $\begin{smallmatrix}5'-UG-3'\\3'-GU-5'\end{smallmatrix}$ and $\begin{smallmatrix}5'-GU-3'\\3'-UG-5'\end{smallmatrix}$ motifs appear to fit into an A-form helix with little distortion. It has been previously suggested that $\begin{smallmatrix}5'-UG-3'\\3'-GU-5'\end{smallmatrix}$ might fit more easily than $\begin{smallmatrix}5'-GU-3'\\3'-UG-5'\end{smallmatrix}$ into A-form, thereby providing a rationale for the difference in thermodynamics (Gautheret et al., 1995). The lack of distortion in the helical regions adjacent to the GU wobble pairs is consistent with the nearest-neighbor model that has been used to analyze the thermodynamics of GU pairs (He et al., 1991).

One major difference revealed by the structures of (rGGAUGUCC)$_2$ and (rGGAGUUCC)$_2$ is a different pattern of overlap of electrostatic potentials, as shown in Figure 10. Hunter (1993) has suggested that electrostatic interactions are responsible for local variations in DNA structure, especially for guanine because it has an unusually broad region of high negative potential. Consideration of overlaps of electrostatic potentials has also suggested that electrostatic interactions may be important for determining structures of tandem GA mismatches (Wu & Turner, 1996) and stabilities of tandem GU wobbles adjacent to GC pairs in RNA (McDowell & Turner, 1996). The qualitative pictures in Figure 10 are consistent with this suggestion in that there is less overlap of negative potentials in the $\begin{smallmatrix}5'-UG-3'\\3'-GU-5'\end{smallmatrix}$ motif than in the less stable $\begin{smallmatrix}5'-GU-3'\\3'-UG-5'\end{smallmatrix}$ motif. Preliminary calculations, however, indicate the absolute magnitudes and signs of the electrostatic interactions are very sensitive to the charge set employed. Moreover, the calculations suggest that

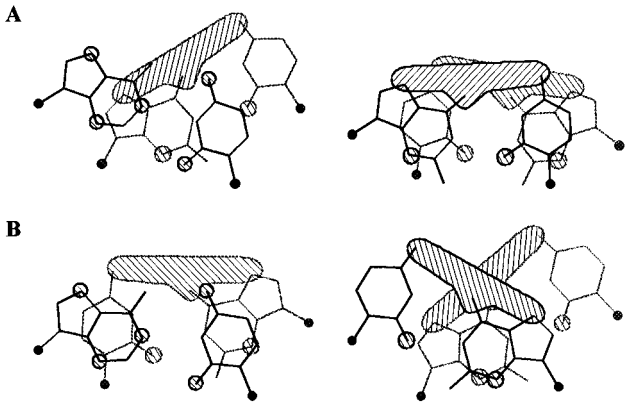

FIGURE 10: Stacking patterns around the mismatches of (A) (rGGAGUUCC)$_2$ and (B) (rGGAUGUCC)$_2$ with regions of high negative potential superimposed as hashed regions. The AU and adjacent GU (A, B left) have similar overlaps of potentials. In contrast, the regions of negative potential are almost entirely overlapped in the $\frac{5'-GU-3'}{3'-UG-5'}$ step and only cross in the $\frac{5'-UG-3'}{3'-GU-5'}$ step. The $\frac{5'-GU-3'}{3'-UG-5'}$ motif is thermodynamically less stable by 2 kcal/mol at 37 °C.

interactions of the adjacent AU base pairs with the GU mismatches may be important even though the apparent electrostatic overlap is minimal. It is also possible that other effects, such as buried surface area (Friedman & Honig, 1995) and polarization, will be important. Thus detailed calculations are required for testing the power of various theories to rationalize and predict the observed thermodynamic and structural properties of tandem GU wobble pairs. The structures presented here should be useful for such tests, especially since these motifs exhibit large differences in thermodynamics while being relatively rigid and having equivalent hydrogen bonding.

SUPPORTING INFORMATION AVAILABLE

One figure of the 150 ms NOESY spectra of (rGGAGUUCC)$_2$ and (rGGAUGUCC)$_2$ showing the H1′/5 to H2′/3′/4′/5′/5″ region with the H1′/H2′ cross peaks labeled and four tables giving distance and torsion angle restraints derived from NMR data (5 pages). Ordering information is given on any current masthead page.

REFERENCES

Allain, F. H.-T., & Varani, G. (1995a) *J. Mol. Biol. 250*, 333−353.

Allain, F. H.-T., & Varani, G. (1995b) *Nucleic Acids Res. 23*, 341−350.

Boelens, R., Koning, T. M. G., & Kaptein, R. (1988) *J. Mol. Struct. 173*, 299−311.

Boelens, R., Vuister, G. W., Koning, T. M. G., & Kaptein, R. (1989) *J. Am. Chem. Soc. 111*, 8525−8526.

Cate, J. H., & Doudna, J. A. (1996) *Structure 4*, 1221−1229.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., & Doudna, J. A. (1996) *Science 273*, 1678−1685.

Feigon, J., Leupin, W., Denny, W. A., & Kearns, D. R. (1983) *Biochemistry 22*, 5943−5951.

Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373−9377.

Friedman, R. A., & Honig, B. (1995) *Biophys. J. 69*, 1528.

Gabriel, K., Schneider, J., & McClain, W. H. (1996) *Science 271*, 195−197.

Gautheret, D., Konings, D., & Gutell, R. R. (1995) *RNA 1*, 807−814.

Gonzalez, C., Rullmann, J. A. C., Bonvin, A. M. J. J., Boelens, R., & Kaptein, R. (1991) *J. Magn. Reson. 91*, 659−664.

Hare, D. R., Wemmer, D. E., Chou, S. H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319−336.

He, L., Kierzek, R., SantaLucia, J., Jr., Walter, A. E., &Turner, D. H. (1991) *Biochemistry 30*, 11124−11132.

Hou, Y. M., & Schimmel, P. (1988) *Nature 333*, 140−145.

Hunter, C. A. (1993) *J. Mol. Biol. 230*, 1025−1054.

Kneale, G., Brown, T., & Kennard, O. (1985) *J. Mol. Biol. 186*, 805−814.

Limmer, S., Hofmann, H.-P., Ott, G., & Sprinzl, M. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 6199−6202.

McClain, W. H., & Foss, K. (1988) *Science 240*, 793−796.

McDowell, J. A., & Turner, D. H. (1996) *Biochemistry 35*, 14077−14089.

Mizuno, H., & Sundaralingam, M. (1978) *Nucleic Acids Res. 6*, 4451−4461.

Neuhaus, D., & Williamson, M. P. (1989*) The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH Publishers, Inc., New York.

Ott, G., Arnold, L., & Limmer, S. (1993) *Nucleic Acids Res. 21*, 5859−5864.

Petersheim, M., & Turner, D. H. (1983) *Biochemistry 22*, 264−268.

Saenger, W. (1984) *Principles of Nucleic Acid Structure*, p 230, Springer-Verlag, New York.

Scheek, R. M., Russo, N., Boelens, R., & Kaptein, R. (1983) *J. Am. Chem. Soc. 105*, 2914−2916.

Simpson, L., & Thiemann, O. H. (1995) *Cell 81*, 837−840.

States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.

Strobel, S. A., & Cech, T. R. (1995) *Science 267*, 675−679.

Usman, N., Ogilvie, K. K., Jiang, M.-V., & Cedergren, R. (1987) *J. Am. Chem. Soc. 109*, 7845−7854.

Varani, G., & Tinoco, I., Jr. (1991) *Q. Rev. Biophys. 24*, 479−532.

Watson, J. D., Hopkins, N. H., Roberts, J. W., Steitz, J. A., & Weiner, A. M. (1987) *Molecular Biology of the Gene*, Benjamin Cummings, Inc., Menlo Park, CA.

White, S. A., Nilges, M., Huang, A., Brünger, A. T., & Moore, P. B. (1992) *Biochemistry 31*, 1610−1621.

Wu, M., McDowell, J. A., & Turner, D. H. (1995) *Biochemistry 34*, 3204−3211.

BI970122C