Investigation of the Structural Basis for Thermodynamic Stabilities of Tandem GU Mismatches: Solution Structure of (rGAGGUCUC)$_2$ by Two-Dimensional NMR and Simulated Annealing[†,‡]

Jeffrey A. McDowell and Douglas H. Turner*

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

Received June 28, 1996; Revised Manuscript Received August 23, 1996[⊗]

ABSTRACT: The duplex (rGAGGUCUC)$_2$ contains the motif $^{5'\text{-GGUC-}3'}_{3'\text{-CUGG-}5'}$, which is unusually stable compared with other symmetric tandem GU mismatches and occurs in the P5 helix of the group I intron of *Tetrahymena thermophila*. The three-dimensional solution structure of (rGAGGUCUC)$_2$ was determined using two-dimensional NMR and a simulated annealing protocol. The structure is remarkably similar to the A-DNA crystal structure of (dGGGGTCCC)$_2$ [Kneale, G., Brown, T., & Kennard, O. (1985) *J. Mol. Biol. 186*, 805−814] which contains the analogous motif $^{5'\text{-GGTC-}3'}_{3'\text{-CTGG-}5'}$. Incorporation of the $^{5'\text{-GGUC-}3'}_{3'\text{-CUGG-}5'}$ motif has little effect on backbone torsion angles and helical parameters compared with standard A-form. The only significant departure from A-form is a slight overtwisting 5′ of the G in the GU mismatch and a displacement of the mismatches toward the minor groove. Inspection of stacking patterns of this structure and comparison with symmetric tandem GT mismatches in A-DNA oligonucleotides from crystal structure data suggest that electrostatics are important in determining motif stability.

The most common base pairs in RNA besides Watson−Crick base pairs are GU mismatches. While their existence was first predicted in order to explain genetic code degeneracies (Crick, 1966), their frequent occurrence in secondary structures (Gutell et al., 1993; Gutell, 1994; Damberger & Gutell, 1994) has led to many questions regarding their structural and functional roles in RNA. Recent statistical studies have found that GU mismatches are often conserved, suggesting they are not always simply replacements of Watson−Crick base pairs but may play either a defined functional or structural role (van Knippenberg et al., 1990; Gautheret et al., 1995).

Many functional roles of GU mismatches have been recognized. Perhaps the most dramatic is the extremely conserved GU mismatch in the P1 helix of group I introns (Michel & Westhof, 1990). Its importance appears to be twofold: allowing the formation of vital tertiary contacts with the intron's catalytic core and stabilizing the ribozyme's transition state through a water bridge or Mg^{2+} ion (Strobel & Cech, 1995) without which ribozyme substrate recognition and fidelity decrease dramatically (Knitt et al., 1994; Pyle et al., 1994). Additionally, the unique wobble pair structure and its effect on the local helical geometry such as major and minor groove topology are functionally important for enzyme recognition. Such importance is observed with tRNA cognate synthetases (Hou & Schimmel, 1988; McClain & Foss, 1988; Gabriel et al., 1996). GU mismatches are also thought to have functional importance in RNA editing (Simpson & Thiemann, 1995). After a guide-RNA (gRNA) mediated block portion of mRNA has been edited, the gRNA−mRNA duplex must be dissociated to allow the adjacent gRNA to associate with an upstream mRNA block and initiate further editing. The dissociation is facilitated by numerous GU mismatches in the gRNA−mRNA duplex.

Structurally, GU mismatches are quite different from Watson−Crick base pairs. C1′−C1′ distances and sugar conformations are similar; the angles between each glycosidic bond and the line segment connecting the C1′ atoms, however, are much different. The result is a loss of the typical pseudodyad structure found in Watson−Crick base pairs. Incorporation of this geometry into a helix causes unusual stacking of adjacent bases. There is significantly more overlap of the U of a GU mismatch toward its 5′ side than its 3′ side and of the G toward its 3′ side than its 5′ side (Mizuno & Sundaralingam, 1978). This effect can be seen in imino proton NMR spectra. The hydrogen bonded, imino protons in a GU mismatch tend to be more shielded, resonating upfield from the resonances of the imino protons of Watson−Crick base pairs due to the difference in ring current shifts arising from this less efficient overall stacking (Geerdes & Hilbers, 1979).

Tandem GU mismatch thermodynamics (He et al., 1991) fit into the same sequence dependence as other tandem mismatches (Wu et al., 1995). Intrahelical tandem mismatch stabilities depend on both the mismatch sequence and the adjacent base pair. The stability depends on the sequence of mismatches in the order $^{5'\text{-UG-}3'}_{3'\text{-GU-}5'} > {}^{5'\text{-GU-}3'}_{3'\text{-UG-}5'}$ and of adjacent base pairs as $^{5'G}_{3'C} > {}^{5'C}_{3'G} > {}^{5'U}_{3'A} \geq {}^{5'A}_{3'U}$. What is unusual about the GU mismatch thermodynamic data is the difference in stability between $^{5'\text{-GGUC-}3'}_{3'\text{-CUGG-}5'}$ and $^{5'\text{-CGUG-}3'}_{3'\text{-GUGC-}5'}$, with the former averaging 3.4 kcal/mol more stability. The motif $^{5'\text{-GGUC-}3'}_{3'\text{-CUGG-}5'}$ is nearly isoenergetic with the motif $^{5'\text{-GUGC-}3'}_{3'\text{-CGUG-}5'}$, whereas $^{5'\text{-CGUG-}3'}_{3'\text{-GUGC-}5'}$ is 3 kcal/mol less stable than $^{5'\text{-CUGG-}3'}_{3'\text{-GGUC-}5'}$. A survey of known secondary structures (Gutell et al., 1993; Gutell, 1994; Damberger & Gutell, 1994) shows that the motif $^{5'\text{-GGUC-}3'}_{3'\text{-CUGG-}5'}$ rarely occurs. The motif does, however,

[†] 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 1GUC and R1GUCMR).

* Author to whom correspondence should be addressed.

[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1996.

S0006-2960(96)01571-1 CCC: $12.00

occur in the P5 helix of the group I intron of *Tetrahymena thermophila* (Michel & Westhof, 1990; Cech et al., 1994; Damberger & Gutell, 1994) and in the signal recognition particle RNA from *Humulus japonicus* (Larsen & Zwieb, 1996). As a step toward understanding the unusual stability of the $^{5'\text{-GGUC-}3'}_{3'\text{-CUGG-}5'}$ motif, this study describes the determination of the solution structure of a $^{5'\text{-GGUC-}3'}_{3'\text{-CUGG-}5'}$-containing helix, $^{5'\text{-GAGGUCUC-}3'}_{3'\text{-CUCUGGAG-}5'}$, and insights this provides into its unusual thermodynamics.

MATERIALS AND METHODS

RNA Synthesis and Purification. RNA oligomers were synthesized using β-cyanoethyl phosphoramidite chemistry (Usman et al., 1987) with an Applied Biosystems DNA/RNA Synthesizer, Model 392. Removal from the CPG support and amino deprotection were by treatment with ethanolic ammonia. Silyl deprotection was accomplished by incubation with freshly made 1 M triethylaminohydrogen fluoride at 55 °C for ~60 h. Samples were desalted by dissolving in 5 mM ammonium acetate at pH 7 and passing through a C-18 column (Waters). Oligomers were purified by TLC on Si500F plates (Baker) with the solvent 55:35:10 1-propanol−ammonia−water. Following purification, oligomers were run through a Dowex 50W-X2 sodium form ion-exchange column to exchange any other cations with sodium. Oligomers were then dialyzed against 0.1 mM EDTA for 24 h and then water for 24 h. Purity was checked by HPLC with a C-8 reverse-phase column (Hamilton) using a 50% methanol mobile phase. All oligomers were better than 98% pure. For NMR, oligomers 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. The oligomers were dissolved under dry nitrogen in 99.996% D_2O for nonexchangeable proton and phosphorus spectra and in 90% H_2O/10% D_2O for exchangeable proton spectra. The final oligomer concentration was approximately 3.5 mM.

Thermodynamic Measurements and Analysis. 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. Oligomer single-strand concentrations, C_T, were calculated from high-temperature absorbances and single-strand extinction coefficients calculated as described previously (Borer, 1975; Richards, 1975). Absorbance versus temperature melting curves were measured at 280 nm with a heating rate of 1 °C/min on a Gilford 250 spectrometer. Thermodynamic parameters for duplex formation were calculated by two methods: (1) by fitting the shape of each curve to a two-state model (Petersheim & Turner, 1983) and (2) by plotting the reciprocal of the melting temperature, T_M^{-1}, versus ln C_T (Borer et al., 1974):

$$\frac{1}{T_M} = \frac{R}{\Delta H^\circ} \ln C_T + \frac{\Delta S^\circ}{\Delta H^\circ} \quad (1)$$

The thermodynamic calculations were performed with a new analysis program, Meltwin. This program was developed in C++ for PC/Windows platforms using a Borland version 4.52 compiler. Thermodynamic calculations are done by fitting the UV-detected thermal melt curves to the equation:

$$\epsilon(T) = (M_{ss}T + B_{ss})\alpha + (M_{ds}T + B_{ds})(1 - \alpha) \quad (2)$$

where $\epsilon(T)$ is the extinction coefficient, ϵ, as a function of temperature, T; the left term of the right expression represents the mole-fraction weighted, linear component of the melt curve due to single strands, and the right term represents the mole-fraction weighted, linear component due to double strands. The mole fraction of single strands, α, is a function of the equilibrium constant, K, and total oligomer concentration, C_T, in the case of bimolecular equilibria and of just the equilibrium constant in unimolecular equilibria. This allows for the extraction of the thermodynamic parameters since

$$K = \exp\left(\frac{\Delta S^\circ}{R} - \frac{\Delta H^\circ}{RT}\right) \quad (3)$$

Fitting of data to eq 2 requires an initial set of parameters which the iterative algorithm refines to calculate the final values. The program makes use of the method of Gralla and Crothers (1973) to estimate an initial enthalpy value, which is then used to calculate the initial entropy value:

$$\Delta H^\circ = \frac{C}{1/T_{1/2} - 1/T_{3/4}} \quad (4)$$

$$\Delta S^\circ = \frac{\Delta H^\circ}{T_{1/2}} + X \quad (5)$$

where C is 4.4, 7.0, or 3.2 cal/(mol·K) and X is ln(C_T), ln(C_T/4), or 0 for self-complementary, non-self-complementary, or unimolecular equilibria, respectively; $T_{1/2}$ and $T_{3/4}$ are the absolute temperatures at the maximum and upper half-maximum of the melt curve derivative, respectively.

NMR Spectroscopy. NMR spectra were collected on a Varian VXR-500S spectrometer with Varian software running on a Sun3 computer. Data processing was performed on Sun4 and SGI computers running Varian and Biosym's Felix software. Two-dimensional spectra were recorded in phase-sensitive mode using the States−Haberkorn method (States et al., 1982). Proton chemical shifts were referenced to an external TSP [3-(trimethylsilyl)propionate] standard. All phosphorus chemical shifts were referenced to phosphate buffer.

The one-dimensional exchangeable proton spectra were acquired using a binomial 1:3:3:1 pulse sequence for HDO peak suppression with a 12 000 Hz sweep width. Pulse delays were calculated to achieve a signal to noise maximum at ~12 ppm, which is in the imino proton resonance region. One-dimensional NOE experiments were performed by irradiation for 3−4 s at low power. Partially overlapped resonances were selectively saturated by positioning the irradiating field 5 Hz off-resonance. To minimize spillover effects, the off-resonance irradiating field of the control spectrum was offset from the resonance potentially exhibiting spillover by the difference of the on-resonance saturation frequency and the resonance frequency potentially exhibiting spillover.

NOESY spectra were acquired at 60, 100, 150, 200, and 400 ms mixing times at 25 and 30 °C. A total of 256−512 FID's were acquired with 4K complex points and 5000 Hz spectral width. For each FID, 32−128 scans were averaged. Any residual HDO resonance was attenuated by low-power presaturation during the recycle delay, which was set to 5 s. Data were apodized using Gaussian or phase-shifted squared sine-bell functions.

DQF-COSY spectra were recorded over spectral widths of 2000 and 4000 Hz at 25 and 30 °C. For each spectrum, 416 FID's were collected with 4K complex points and 80 scans per transient. The recycle delay was set to 2.5 s. Data were zero-filled to 2K real points in the t_1 dimension and apodized using phase-shifted squared sine-bell functions.

^{1}H−^{31}P HETCOR spectra were acquired using the pulse sequence described by Sklenar et al. (1986) with spectral widths of 1100 Hz in the ^{1}H dimension and 1000 Hz in the ^{31}P dimension. A total of 325 FID's with 96 scans per FID were collected in 2K complex points. The recycle delay was set to 2.5 s. Data were apodized using phase-shifted squared sine-bell functions.

Distance Restraint Generation. Interproton distance constraints between nonexchangeable protons were obtained from the intensities of their NOE cross peaks in 2D NOESY spectra at 100, 150, and 200 ms mixing times. Pyrimidine H5/H6 cross peaks which correspond to a fixed distance of 2.45 Å were averaged and used as an internal reference. Isolated, easily quantifiable cross peaks were given limits of ±20%. Other peaks with partial overlap or noise ridge problems were given larger limits up to ±30%. Severely overlapped cross peaks which clearly existed in all acquired NOESY spectra were given a constraint of ≤5.0 Å. Three cross peaks which would have been well resolved but were absent in all acquired NOESY spectra were given restraints of ≥4.0 Å. There is little change in structure when these restraints are omitted. Hydrogen-bonded atoms of the Watson−Crick and GU mismatches were given distance constraints of 1.8 ± 0.2 Å. Structures were also determined without GU mismatch hydrogen-bonding constraints, and no change in structure was observed.

Torsion Angle Restraint Generation. In principle, all nonphosphate backbone torsion angles can be constrained with semiquantitative estimates of *J*-coupling constants obtained from DQF-COSY and ^{1}H−^{31}P HETCOR experiments. The appropriate Karplus equations used to convert the *J*-coupling constants to torsion angles have up to four solutions. By use of multiple *J*-couplings which depend on the same torsion angle, for example, $J_{P-H5'}$ and $J_{P-H5''}$ which both depend on the β (P−O5′−C5′−C4′) torsion angle, it is often possible to determine a torsion angle range consistent with the data (Wijmenga et al., 1993). This is the situation for the β and γ (O5′−C5′−C4′−C3′) torsion angles. Small, equivalent $J_{P-H5'}$ and $J_{P-H5''}$ coupling constants indicate a usual *trans* β angle. Small, equivalent $J_{H4'-H5'}$ and $J_{H4'-H5''}$ coupling constants confirm a *gauche*$^{\pm}$ γ angle. However, a γ angle of *gauche*$^-$ rarely occurs in nucleic acids and can usually be disregarded. Unfortunately, the two, usually small coupling constants that define the γ angle are in an extremely crowded region of the spectrum and can often not be determined. The ϵ (C4′−C3′−O3′−P) torsion angle is determined by $J_{P-H3'}$ and is usually *trans*, giving a coupling constant of 10−13 Hz. This coupling constant cannot distinguish between the *trans* and *gauche*$^-$ conformations. Nucleotides with C3′-*endo* sugar puckers and *gauche*$^-$ ϵ torsion angles have never been observed, however, and are considered forbidden (Wijmenga et al., 1993); therefore, if sugar pucker conformations are clearly C3′-*endo* as indicated by weak or absent H1′/H2′ cross peaks, these coupling constants and sugar puckers are evidence for *trans* ϵ geometry.

The phosphate torsion angles, α (O3′−P−O5′−C5′) and ζ (C3′−O3′−P−O5′), cannot be determined directly from *J*-coupling data but can be somewhat restrained indirectly from ^{31}P chemical shifts. A *trans* conformation for either α or ζ around a phosphorus will result in a large downfield shift of its resonance (Gorenstein, 1984). Therefore, if all resonances have ^{31}P chemical shifts in the usual −3 to −5 ppm range and are within 1.0 ppm of each other, the *trans* conformation of the α and ζ angles can be excluded and the angles restrained to values of 0 ± 150°.

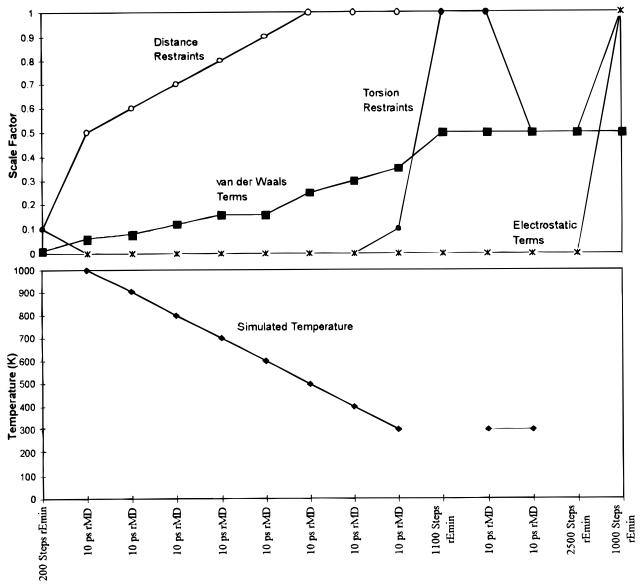

FIGURE 1: Simulated annealing protocol: temperature (◆), van der Waals term factor (■), distance restraint penalty function factor (○), torsion restraint penalty function factor (●), and electrostatic term factor (★).

The sugar pucker defines the final torsion angle, δ (C5′−C4′−C3′−O3′). This angle can be constrained using the H1′/H2′ cross peak. For C3′-*endo* sugar puckers, this coupling is weak (J^3 < 2 Hz), resulting in weak or absent H1′/H2′ cross peaks compared to larger coupling (J^3 ≥ 8 Hz) and cross peaks for C2′-*endo* sugar puckers (Varani & Tinoco, 1991). The sugar pucker can be further confirmed by the H2′/H3′ cross peak pattern which will not show passive coupling of the H2′ due to H1′ splitting in C3′-*endo*.

Constraint Satisfaction Protocol. Structure determination was accomplished by the simulated annealing protocol shown in Figure 1 using Biosym's Discover package running on a Silicon Graphics ONYX workstation. Calculations used the Amber force field (Weiner et al., 1986) with a nonbonded term cutoff of 12 Å. Electrostatic terms were excluded in all but the final minimization step, which used the Amber charge set (Cornell et al., 1995) with phosphate group charges reduced to −0.2 e to simulate the effect of counterions (Harvey et al., 1984) and a distance-dependent dielectric constant to mimic solvent dielectric effects (Weiner et al., 1986). Flat-bottomed quadratic penalty terms for distance and torsion angle restraints with force constants of 25 kcal/(mol·Å^2) and 25 kcal/(mol·rad^2), respectively, were added to the force field. Each restrained energy minimization (rE.Min.) step is calculated either for the maximum number of iterations listed or until the RMS derivative change between successive iterations is ≤0.0001 kcal/(mol·Å). The simulated annealing protocol has six steps (Figure 1):

(a) 200 iterations of rE.Min. with very low van der Waals (vdW) terms and small distance and torsion angle restraint force constants in order to alleviate any extreme hot spots in the initial structure.

(b) 80 ps of restrained molecular dynamics (rMD) with the simulated temperature gradually decreased from 1000 to 300 K and the scaling of vdW terms increased from near zero to 0.35. Distance restraints are gradually introduced during this step. No torsion restraints are included in the initial 70 ps in order to completely satisfy all the distance restraints.

(c) 1100 iterations of rE.Min. with vdW terms scaled by 0.5 and full values of distance and torsion angle restraint force constants. This step is used to alleviate any hot spots in the torsion angles while maintaining the satisfied distance restraint set.

(d) 20 ps of rMD at a simulated temperature of 300 K during which the distance and torsion angle restraint force constant factors were scaled from 1 to 0.5 to simultaneously satisfy both restraint sets.

(e) 2500 iterations of r.E.Min. with distance and torsion angle restraint force constant factors scaled at 0.5. This step eliminates vibrational artifacts from the final structure and has the effect of improving the planarity of the bases.

(f) 1000 iterations of rE.Min. with electrostatic terms and with vdW terms scaled by 0.5 and full constraint force terms. This step is intended to obtain proper electrostatic contacts while maintaining very good agreement with the distance and torsion angle constraint sets. In practice, this step generally produced negligible changes in this set of calculations.

Thirty initial structures were used, one each with A-form and B-form conformations and 28 with randomized backbone torsion angles (excluding δ). The A-form and B-form geometries were generated using Biosym's InsightII nucleic acid library. The randomized torsion angle conformations were generated by modifying A-form structures with a program written in C++ on an SGI ONYX workstation using the SGI Delta-C++ release 5.3 compiler.

In order to address the quality of the final structure, average pairwise root mean squared deviations (RMSD) of the converged structures are calculated. Such a calculation assesses the precision but not necessarily the accuracy of the structure, which is more difficult to consider. While no adequate criteria exist to address accuracy, theoretical NOESY spectra were back-calculated using a relaxation matrix approach (RMA) (Boelens et al., 1988, 1989) with an isotropic correlation time of 3 ns and compared to the experimental NOESY spectra by calculating three *R* factors to estimate the structure's accuracy. Common *R* factors used to compare experimental and calculated NOESY spectra are analogous to those used in crystallography (Gonzalez et al., 1991):

$$R_{\mathrm{A}} = \frac{\sum_{\mathrm{spectra}} W_{ij}(\tau) \sum_{i,j} |A_{ij}^{\mathrm{calc}}(\tau) - A_{ij}^{\mathrm{exp}}(\tau)|}{\sum_{\mathrm{spectra}} W_{ij}(\tau) \sum_{i,j} A_{ij}^{\mathrm{exp}}(\tau)} \quad (6)$$

$$R_{\mathrm{B}} = \sqrt{\frac{\sum_{\mathrm{spectra}} W_{ij}(\tau) \sum_{i,j} |A_{ij}^{\mathrm{calc}}(\tau) - A_{ij}^{\mathrm{exp}}(\tau)|^2}{\sum_{\mathrm{spectra}} W_{ij}(\tau) \sum_{i,j} (A_{ij}^{\mathrm{exp}}(\tau))^2}} \quad (7)$$

where $A_{ij}(\tau)$ is the NOE intensity of the spin pair (ij) at mixing time τ. The weight factors, $W_{ij}(\tau)$, reflect the NOE measurement error, which is a difficult quantity to estimate. As the errors are higher for the lower signal-to-noise ratios of shorter mixing times, in one case $W_{ij}(\tau)$ was set equal to the mixing time. This study used three *R* factors calculated for spin pairs which had well-determined cross peaks for 100, 150, 200, and 400 ms NOESY mixing times:

$$R_1 = \frac{\sum_{\mathrm{spectra}} \sum_{i,j} |A_{ij}^{\mathrm{calc}}(\tau) - A_{ij}^{\mathrm{exp}}(\tau)|}{\sum_{\mathrm{spectra}} \sum_{i,j} A_{ij}^{\mathrm{exp}}(\tau)} \quad (8)$$

$$R_\tau = \frac{\sum_{\mathrm{spectra}} \tau \sum_{i,j} |A_{ij}^{\mathrm{calc}}(\tau) - A_{ij}^{\mathrm{exp}}(\tau)|}{\sum_{\mathrm{spectra}} \tau \sum_{i,j} A_{ij}^{\mathrm{exp}}(\tau)} \quad (9)$$

$$R_2 = \sqrt{\frac{\sum_{\mathrm{spectra}} \sum_{i,j} |A_{ij}^{\mathrm{calc}}(\tau) - A_{ij}^{\mathrm{exp}}(\tau)|^2}{\sum_{\mathrm{spectra}} \sum_{i,j} (A_{ij}^{\mathrm{exp}}(\tau))^2}} \quad (10)$$

R factors were calculated for the final structure as well as for four of the starting structures for comparison.

Potential Surface Calculations. In order to visualize the potential role of electrostatics on the unusual thermodynamics of the $\begin{smallmatrix}5'\text{-GGUC-}3'\\3'\text{-CUGG-}5'\end{smallmatrix}$ motif, electrostatic potential surfaces were calculated for GC, AU, and GU base pairs with the Amber charge set (Cornell et al., 1995) and a dielectric constant of 1 using a program written in C++ on an SGI ONYX workstation using the SGI Delta-C++ release 5.3 compiler. Base pair geometries were from Biosym's InsightII nucleic acid library for Watson−Crick base pairs and from the average final structure from this study for the GU mismatch. Nucleotide sugars were replaced by methyl groups. Methyl hydrogens were given charges between 0.04 and 0.06 e such that the overall base charge was zero. The potential surface was calculated at a distance of 1.4 Å, half the usual A-form rise, above the average base pair plane. GU mismatches have two hydrogen-bonding sites, G−N2 and U−O4, which are left unpaired in the usual wobble pair geometry. These sites when exposed to solvent will be satisfied by bound water molecules. The crystal structure of Holbrook et al. (1991) identified the geometries of these bound water molecules. As these water molecules could affect the contributions of the non-hydrogen-bonded sites to the potential surface, calculations were performed with and without these identified, bound water molecules. The bound waters, however, resulted in only slight attenuation of the potential around the exocyclic amine of G in the mismatch.

RESULTS

Patterns in Thermodynamic Stability of Tandem GU Mismatches Do Not Depend on Base Pairs beyond Nearest Neighbors or on the Number of Hydrogen-Bonded Imino Protons. The free energy increment, ΔG°_{37}, for a tandem GU mismatch is defined as

$$\Delta G^\circ_{37}(\mathrm{G\underline{UG}C}) = \Delta G^\circ_{37}(\mathrm{AUG\underline{UG}CAU}) - \Delta G^\circ_{37}(\mathrm{AUGCAU}) + \Delta G^\circ_{37}(\mathrm{GC})$$

Here $\Delta G^\circ_{37}(\mathrm{AUG\underline{UG}CAU})$ and $\Delta G^\circ_{37}(\mathrm{AUGCAU})$ are the free energy changes of duplex formation derived from $T_{\mathrm{m}}{}^{-1}$ versus ln C_{T} plots for the strand sequences shown in parentheses and $\Delta G^\circ_{37}(\mathrm{GC})$ is the free energy increment (Freier et al., 1986) for the nearest-neighbor interaction interrupted by insertion of the tandem GU mismatch. He et al. (1991) found that the free energy increments of tandem GU mismatch motifs are very sequence dependent, ranging from 0.1 to −4.8 kcal/mol for the A$\underline{\mathrm{GU}}$U to G$\underline{\mathrm{UG}}$C motifs,

Table 1: Thermodynamics of Duplex Formation

	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)
GAGUGCUC[b]	-9.2 ± 0.3	-77.4 ± 2.7	-220.0 ± 8.7	51.8	-9.4 ± 1.1	-83.0 ± 5.7	-237.4 ± 14.9	51.6
GGUUGACC[b]	-8.3 ± 0.2	-77.4 ± 2.6	-222.8 ± 8.7	47.7	-8.3 ± 0.9	-78.3 ± 5.3	-225.9 ± 14.1	47.6
GGCUGGCC[b]	-13.3 ± 0.7	-90.3 ± 1.2	-248.4 ± 3.4	65.4	-13.1 ± 1.3	-87.2 ± 6.2	-238.8 ± 15.6	65.9
GAGGUCUC[b,c]	-8.6 ± 0.1	-77.9 ± 1.7	-223.3 ± 5.4	49.2	-8.8 ± 0.1	-82.8 ± 2.0	-238.7 ± 6.3	49.1
GAGGUCUC[d]	-8.3 ± 0.1	-81.1 ± 3.5	-234.8 ± 11.0	47.2	-8.4 ± 0.1	-85.1 ± 2.3	-247.6 ± 7.3	47.1

[a] Calculated for oligomer concentration of 10^{-4} M. [b] Measured in 1 M NaCl. [c] Wu et al. (1995). [d] Measured in 0.1 M KCl and 10 mM $MgCl_2$.

Table 2: Thermodynamic Increments for Tandem GU Mismatches[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)]
AUGUGCAU[b]	−4.8	−29.6	−80.0	CGGGUCCG[c]	−4.7	−41.5	−118.8
GAGUGCUC	−4.6	−29.3	−79.3	GAGGUCUC[d]	−4.2	−34.7	−98.3
				CUGGUCAG[c]	−3.4	−29.9	−85.4
GCUGGC[b]	−3.9	−36.6	−105.7	GUCGUGAC[c]	−1.0	−23.5	−72.7
GGCUGGCC	−4.0	−30.5	−85.8	GGCGUGCC[c]	−0.4	−14.1	−44.2
CGUUGACG[c]	−2.7	−35.0	−104.1	CCUGUAGG[c]	−0.1	−25.1	−80.9
GGUUGACC	−2.1	−30.6	−96.0				
GGAUGUCC[c]	−1.9	−25.0	−74.8	GGAGUUCC[c]	+0.1	−25.1	−81.4

[a] Calculated as $\Delta G°_{37}$ for GUGC: $\Delta G°_{37}$(GUGC) = $\Delta G°_{37}$(AUGUGCAU) = $\Delta G°_{37}$(AUGCAU) + $\Delta G°_{37}$(GC). Thermodynamic parameters derived from T_m^{-1} vs ln C_T plots. [b] Sugimoto et al. (1987). [c] He et al. (1991). [d] Wu et al. (1995).

respectively. To check the large sequence dependence, thermodynamic parameters were measured for three additional duplexes with tandem GU mismatches, (rGAGUGCUC)$_2$, (rGGCUGGCC)$_2$, and (rGGUUGACC)$_2$. In addition, to check the validity of using 1 M NaCl as a substitute for more physiological conditions, thermodynamic parameters for (rGAGGUCUC)$_2$ in 0.1 M KCl and 10 mM $MgCl_2$ were measured and compared with previous results in 1 M NaCl (Wu et al., 1995). The results are reported in Table 1, and a summary of thermodynamic increments for tandem GU mismatches in all sequences studied thus far is given in Table 2. The results in Table 2 indicate that the free energy increments measured for the tandem GU motifs are relatively independent of base pairs extending beyond those immediately adjacent to the GU mismatches. For example, the free energy increments measured for GUGC in AUGUGCAU and GAGUGCUC are −4.8 and −4.6 kcal/mol, respectively. Comparison of the free energy increments of −3.9 and −4.0 kcal/mol, respectively, for GCUGGC and GGCUGGCC indicates that addition of two base pairs also makes little difference. Thus patterns previously observed are not dependent on sequence beyond the tandem GU motif. In addition, the thermodynamic parameters for (rGAGGUCUC)$_2$ are essentially the same in 1 M NaCl and in 0.1 M KCl and 10 mM $MgCl_2$. A molecular dynamics study has suggested that O4 of thymine in a GT pair can associate with a cation (Shibata et al., 1991). If that occurs for (rGAGGUCUC)$_2$, the thermodynamic stability is likely independent of the cation.

Figure 2A shows imino proton NMR spectra for (rGAGUGCUC)$_2$, (rGGCUGGCC)$_2$, (rGGUUGACC)$_2$, and (rGAGGUCUC)$_2$ at 25 °C. All spectra have five resonances between 10 and 15 ppm, suggesting that all five imino protons are hydrogen bonded. Resonance assignments are based on 1-D NOE spectra as described below for (rGAGGUCUC)$_2$.

The above results suggest that the differences in stabilities of tandem GU mismatches depend on local interactions between the GU mismatches and between the GU mismatches and adjacent Watson−Crick base pairs. To provide an initial test for this hypothesis, the three-dimensional structure of (rGAGGUCUC)$_2$ has been determined by 2-D NMR and simulated annealing. This sequence was chosen because the free energy increment of the GGUC motif averages 3.4 kcal/mol more favorable than that for CGUG. Moreover, the GGUC motif is anomalous in that it has roughly the same free energy increment as GUGC, whereas CGUG, UGUA, and AGUU are at least 2 kcal/mol less stable than their corresponding UG motifs. The GGUC motif also occurs in the secondary structure of the group I intron from *T. thermophila* and the signal recognition particle RNA from *H. japonicus*.

Assignment of AH2 and Exchangeable Imino Protons. Imino proton resonances were assigned at both 0.5 and 25 °C. The one-dimensional imino proton spectrum and NOE difference spectra of (rGAGGUCUC)$_2$ at 0.5 °C are shown in Figure 2B. Five resonances exist between 11 and 14.5 ppm which correspond to the five imino protons of the sequence, G1-H1, G3-H1, G4-H1, U5-H3, and U7-H3. For spectra measured from 0 to 50 °C, the resonance shown at 12.18 ppm is always the broadest and also broadens at a lower temperature than the other resonances due to exchange with solvent. This is consistent with a terminal nucleotide, and thus the resonance at 12.18 ppm is assigned to G1-H1. The resonance at 14.42 ppm shows an NOE with the resonance at 13.10 ppm as well as a peak at 7.60 ppm in the amino/base proton region. The peak at 7.60 ppm is a nonexchangeable proton with a long T_1 of 5.5 s. These data are consistent with an A-H2 peak, since A-H2's typically have long T_1's and the imino proton in the U of an AU base pair is very close to the complementary A-H2 (≈2.85 Å) giving rise to a large NOE. Thus, the resonance at 7.60 ppm is assigned to A2-H2, and the peak at 14.42 ppm is assigned to U7-H3. This downfield shift of U7-H3 is consistent with the expected shift of an imino proton in an AU base pair. The peak at 13.10 ppm is assigned to G3-H1.

GU mismatches have characteristic imino proton chemical shifts and NOE signatures. Due to more shielding compared

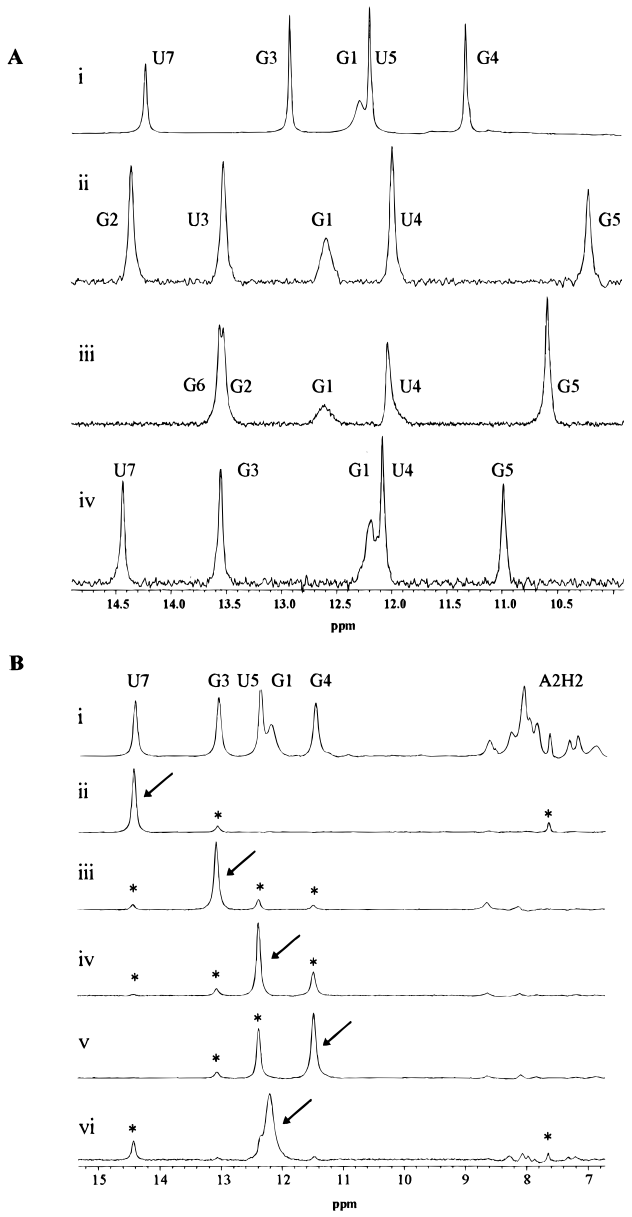

FIGURE 2: (A) Imino proton (uridine H3/guanine H1) NMR spectra of (i) (rGAGGUCUC)$_2$, (ii) (rGGUUGACC)$_2$, (iii) (rGGCUGGCC)$_2$, and (iv) (rGAGUGCUC)$_2$ at 25 °C. Labels indicate assigned imino proton resonances. In each sequence, all five imino protons have an assigned resonance, suggesting all are hydrogen bonded. (B) Imino proton spectrum of (rGAGGUCUC)$_2$ (i) and NOE difference spectra (ii−vi) at 0.5 °C. Nucleotide labels indicate imino proton resonances. Arrows indicate saturated resonance. Observed NOE's listed in Table 3 are indicated by asterisks.

with intrahelical Watson−Crick base pairs, presumably caused by a decrease in stacking of the G to its 5′ side affecting the ring current deshielding, the imino proton of G is upfield shifted from Watson−Crick base pair imino protons (Geerdes & Hilbers, 1979). The imino proton of G in a GU mismatch is usually found around 11 ppm. The imino protons in a GU mismatch are within 2.5 Å, resulting in a mutual NOE (Johnson & Redfield, 1981). Since the G is not typically stacked to its 5′ side, however, the NOE from the G imino proton to the imino proton of the 5′ base pair will be weak or missing (van Knippenberg et al., 1990). Irradiating the peak at 11.50 ppm shows a strong NOE to the 12.35 ppm peak and a weak NOE to the peak at 13.10 ppm previously assigned to G3-H1. This is consistent with the 11.50 ppm peak being G4-H1. At 25 °C, the 12.35 ppm peak shows only an NOE to the peak at 11.50 ppm and is

Table 3: Fractional NOE Transfer[a] of Imino Protons in (rGAGGUCUC)$_2$

irradiated resonance	NOE	fractional NOE transfer 0.5 °C	fractional NOE transfer 25.0 °C
G1H1	U7H3	0.12	w
	A2H2	w[b]	w
G3H1	G4H1	0.06	0.06
	U5H3	0.13	0.11
	U7H3	0.06	0.05
G4H1	G3H1	0.06	w
	U5H3	0.59	0.17
U5H3	G3H1	0.08	w
	G4H1	0.35	0.12
	U7H3	w	w
U7H3	G3H1	0.14	0.04
	A2H2	0.07	0.10

[a] Integrated area of NOE peak divided by integrated area of saturated peak. [b] w = fractional transfer less than 0.04.

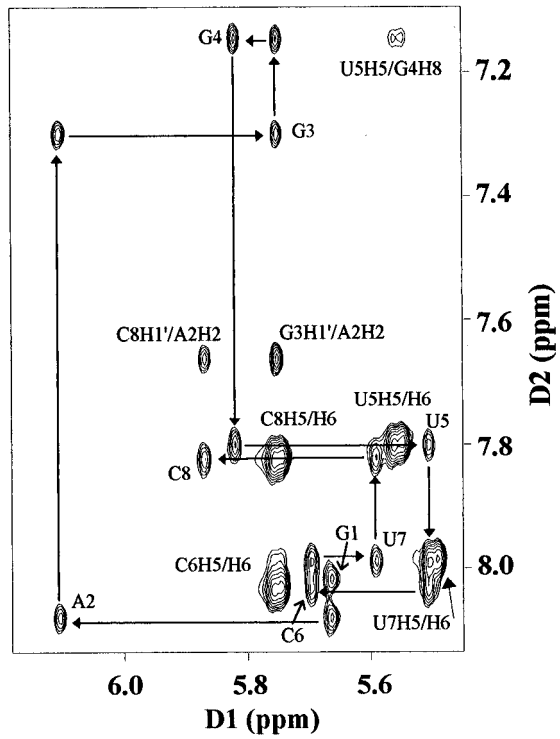

FIGURE 3: 400 ms NOESY spectrum of (rGAGGUCUC)$_2$ at 30 °C showing the H1′/5−H8/6/2 region. Base$_n$−H1′$_n$ cross peaks are labeled as well as peaks not in the H8/6$_n$−H1′$_n$−H8/6$_{n+1}$ NOE connectivity pathway.

therefore assigned to U5-H3. Table 3 lists fractional transfers observed in the NOE spectra at both 0.5 and 25 °C.

Assignment of Nonexchangeable Protons and Phosphorus. Chemical shifts of all assigned proton and phosphorus resonances are reported in Table 4. The base (H8/H6/H2/H5) and sugar H1′ protons were assigned from the 400 ms NOESY spectrum at 30 °C shown in Figure 3. By starting with the A2-H2 as well as the U5H5/G4H8 cross peak, which can easily be identified by its size, splitting, and connectivity to an H5/H6 cross peak (G4-U5 is the only purine−pyrimidine step in the sequence), all peaks in the sequential H8/6(n)−H1′(n)−H8/6(n+1) NOE connectivity pathway (Scheek et al., 1983; Feigon et al., 1983; Petersheim & Turner, 1983; Hare et al., 1983) were assigned. H8/6−H1′ distances are independent of sugar pucker but dependent on the glycosidic torsion angle, χ. This dependence is such that, given the experimental accuracy of NOE measurements, it

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

	H8/H6	H2/H5	H1′	H2′	H3′	H4′	H5′/H5″	imino (GH1/UH3)	phosphorus
G1	8.02	na[b]	5.66	4.83	4.68	4.34	4.06/3.94	12.18	na
A2	8.08	7.66	6.11	4.78	4.81	4.59		na	−2.81
G3	7.30	na	5.76	4.49	4.52	4.50	4.14/4.60	13.10	−2.70
G4	7.15	na	5.82	4.84	4.21	4.53		11.50	−2.60
U5	7.80	5.56	5.51	4.21	4.58	4.46	4.12/4.57	12.35	−3.48
C6	8.03	5.76	5.70	4.37	4.56	4.47		na	−3.49
U7	7.99	5.50	5.59	4.38	4.58	4.45		14.42	−3.45
C8	7.82	5.75	5.87	4.05	4.22	4.37		na	−3.18

[a] Proton chemical shifts are referenced to an external TSP [3-(trimethylsilyl)propionate] reference at 30 °C except for imino protons which are at 0.5 °C. Phosphorus chemical shifts are referenced to internal phosphate buffer. [b] na = not applicable.

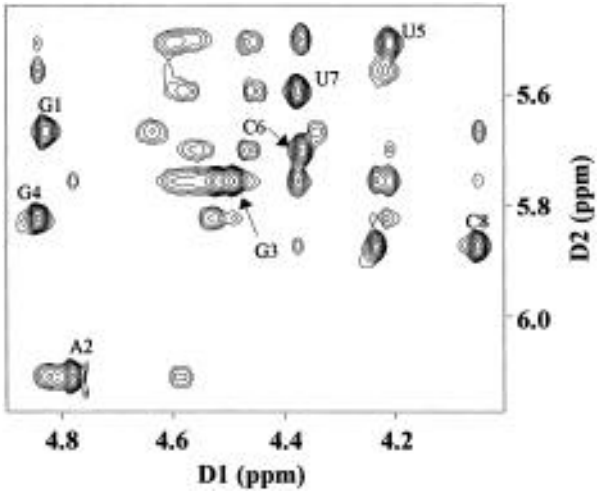

FIGURE 4: 400 ms NOESY spectrum of (rGAGGUCUC)$_2$ at 30 °C showing the H1′/5−sugar (H2′,H3′,H4′,H5′,H5″) region. Labels designate H1′−H2′ cross peaks.

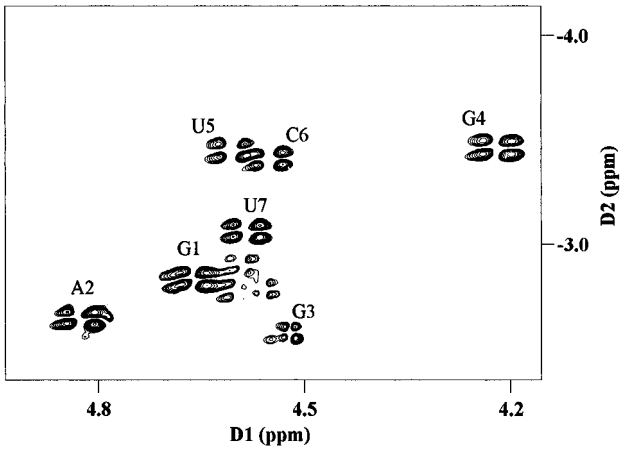

FIGURE 5: Proton-detected ^{1}H/^{31}P HETCOR spectrum of (rGAGGUCUC)$_2$ at 30 °C with the H3′$_n$−P$_{n+1}$ cross peaks labeled.

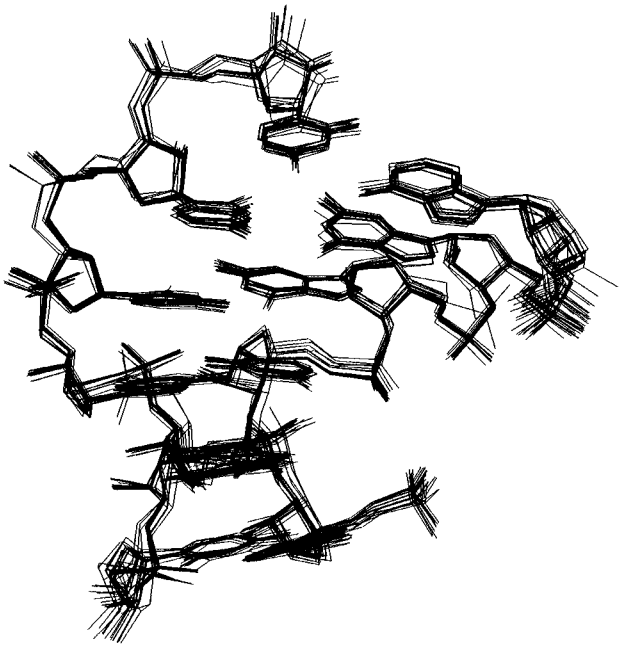

FIGURE 6: Superposition of a random 16 of the 30 converged structures of (rGAGGUCUC)$_2$ showing the internal nucleotides. The average RMSD for the pairwise all-atom superposition for the full set of converged structures of the entire molecule is 0.65 Å. The RMSD calculated for the internal nucleotides is 0.45 Å.

can only be established if a nucleotide is *syn* or *anti* (Wijmenga et al., 1993). The cross peak pattern and intensities verify that the structure is generally A-form and all nucleotides have *anti* glycosidic torsion angles. The H5/H6 cross peaks of the pyrimidines (U5, C6, U7, C8) can be identified by their splitting and confirmed by strong cross peaks in the DQF-COSY spectrum.

The sugar protons (H2′, H3′, H4′, some H5′/5″) were assigned using a combination of NOESY and COSY spectra (Varani & Tinoco, 1991). Using the H1′ assignments, the H2′ protons were assigned by the large H1′/H2′ cross peaks in short mixing time (≤100 ms) NOESY spectra. H1′−H2′ distances vary from 2.3 to 2.4 Å, essentially independent of sugar pucker, so no assumption is made on conformation. H1′−H4′ distances are very dependent on sugar pucker, varying from 2.5 to >3.7 Å. For C3′-*endo* sugar puckers, this distance is ∼2.9 Å, and consequently these cross peaks were present in short mixing time NOESY spectra for (rGAGGUCUC)$_2$. H1′−H3′ distances are also independent of sugar pucker at ∼3.9 Å. These cross peaks are, however, usually broader than the others due to the strong H3′/P coupling. H3′ resonances were therefore tentatively assigned by longer mixing time (200−400 ms) NOESY spectra. Figure 4 shows a 400 ms NOESY spectrum of the H1′/sugar region. H2′, H3′, and H4′ assignments were confirmed by COSY spectra. Specifically, H2′/H3′ coupling for C3′-*endo* sugars gives rise to a medium-sized COSY cross peak with a distinctive cross peak pattern. For spectra without phosphorus decoupling, a C3′-*endo* H2′/H3′ cross peak has the phase pattern, $^{++-\ -}_{-\ -++}$ ← H2′. H3′/H4′ protons are strongly coupled and result in large COSY cross peaks. H4′/H5′(″) couplings depend on the γ backbone angle and are weak (J^3 < 3 Hz) for A- and B-form values of γ (36−54°).

For A- and B-form values of ϵ (≈−155°), the H3′ proton is strongly coupled to the backbone phosphates (J^3 ≥ 10 Hz) (Varani & Tinoco, 1991). Figure 5 shows a proton-detected ^{1}H−^{31}P HETCOR spectrum with the intense H3′-(n)/P(n+1) cross peaks labeled. When present, these cross

Table 5: Distance Restraints (Å) Used in the Simulated Annealing Protocol[a]

intranucleotide	H8/6−H1′	H8/6−H2′	H8/6−H3′	H8/6−H4′	H1′−H2′	H1′−H3′	H1′−H4′
G1	3.4 ± 0.8	3.0 ± 0.7	≤5	3.8 ± 0.9	3.2 ± 0.5	3.4 ± 0.8	3.5 ± 0.8
A2	3.7 ± 0.9	3.2 ± 0.7	NR[b]	NR	2.6 ± 0.6	2.9 ± 0.7	3.7 ± 0.9
G3	4.1 ± 1.0	3.2 ± 0.9	3.2 ± 0.7	≤5	2.7 ± 0.8	3.7 ± 0.9	2.7 ± 0.6
G4	3.5 ± 0.8	3.7 ± 0.9	3.1 ± 0.7	NR	2.6 ± 0.6	3.4 ± 0.8	3.2 ± 0.7
U5	3.3 ± 0.8	2.7 ± 0.8	≤5	2.9 ± 0.7	2.6 ± 0.6	3.2 ± 0.7	3.4 ± 0.8
C6	3.2 ± 0.9	3.3 ± 0.8	≤5	3.2 ± 0.8	2.6 ± 0.6	3.3 ± 0.8	3.2 ± 0.8
U7	3.3 ± 0.8	NR	≤5	≤5	2.5 ± 0.6	3.3 ± 0.8	3.3 ± 0.8
C8	3.4 ± 0.8	3.0 ± 0.7	2.4 ± 0.6	NR	2.6 ± 0.6	2.9 ± 0.7	4.0 ± 0.9

internucleotide	H8/6−H1′	H8/6−H2′	H8/6−H3′	H8/6−H4′	H1′−H2′	H2′−H5	H3′−H5
G1−A2	3.6 ± 0.8	2.3 ± 0.5	2.6 ± 0.6	≥4.0	NR	NR	NR
A2−G3	3.8 ± 0.9	2.6 ± 0.6	3.0 ± 0.7	≥4.0	3.7 ± 0.9	NR	NR
G3−G4	3.9 ± 0.9	2.6 ± 0.6	≥4.0	≤5	3.9 ± 0.9	NR	NR
G4−U5	3.6 ± 0.8	2.6 ± 0.6	NR	NR	3.8 ± 0.9	3.4 ± 0.8	3.0 ± 0.7
U5−C6	3.0 ± 0.8	2.5 ± 0.6	NR	≤5	3.6 ± 0.9	NR	NR
C6−U7	3.3 ± 1.0	NR	NR	NR	NR	3.1 ± 0.6	3.4 ± 0.6
U7−C8	3.1 ± 0.9	2.5 ± 0.6	NR	NR	NR	NR	NR

internucleotide	H5−H5	H8/6−H8/6	H8/6−H5	internucleotide	A2H2
G1−A2	NR	NR	NR	G3H1′	3.2 ± 0.8
A2−G3	NR	4.4 ± 1.1	NR	C8′H1′	3.5 ± 0.8
G3−G4	NR	4.6 ± 1.1	NR	A2H1′	4.7 ± 1.1
G4−U5	NR	3.9 ± 0.9	3.6 ± 0.8	U7′H1′	5.2 ± 1.3
U5−C6	NR	NR	NR		
C6−U7	2.9 ± 0.7	NR	NR		
U7−C8	NR	NR	NR		

[a] Base pairs were also constrained by hydrogen-bonding restraints of 1.8 ± 0.2 Å. [b] NR = not restrained.

Table 6: Torsion Angle Restraints (deg) Used in Simulated Annealing Protocol

	α	β	γ	δ	ε	ζ	χ
G1	na[a]	na	NR[b]	80 ± 15	155 ± 30	0 ± 150	179.5 ± 50
A2	0 ± 150	180 ± 30	NR	80 ± 15	155 ± 30	0 ± 150	179.5 ± 50
G3	0 ± 150	NR	47.43 ± 25	80 ± 15	NR	0 ± 150	179.5 ± 50
G4	0 ± 150	180 ± 30	47.43 ± 25	80 ± 15	155 ± 30	0 ± 150	179.5 ± 50
U5	0 ± 150	180 ± 30	NR	80 ± 15	155 ± 30	0 ± 150	179.5 ± 50
C6	0 ± 150	180 ± 30	NR	80 ± 15	155 ± 30	0 ± 150	179.5 ± 50
U7	0 ± 150	180 ± 30	NR	80 ± 15	155 ± 30	0 ± 150	179.5 ± 50
C8	0 ± 150	180 ± 30	NR	80 ± 15	na	na	179.5 ± 50
A-form	−62	180	47.5	84	−152	−74	−166

[a] na = not applicable. [b] NR = not restrained.

Table 7: All-Atom Pairwise RMSD Values Calculated for Base Pairs in the Final Structure of (rGAGGUCUC)$_2$

base pair	RMSD (Å)	base pair	RMSD (Å)
G1·C8	0.51	U5·G4	0.39
A2·U7	0.57	C6·G3	0.46
G3·C6	0.40	U7·A2	0.47
G4·U5	0.37	C8·G1	0.55

peaks confirm the H3′ assignments, eliminate any possibility of an H3′ misassignment, and assign the phosphorus resonances. Examination of the phosphorus spectrum reveals that all seven phosphorus resonances are within 0.9 ppm, indicating that no α or ζ torsion angle is *trans* (Gorenstein, 1984). ^{1}H−^{31}P HETCOR spectra usually show weak cross peaks due to P(*n*)/H5′(′′) (*n*) coupling which can help to assign these protons. Lack of stereospecificity of H5′/H5′′ assignments prohibits their use in distance restraints.

Structure Calculations. A total of 108 distance restraints and 44 torsion angle restraints per strand were used in the structure calculations (Tables 5 and 6). Figure 6 shows the superposition of a random 16 of the 30 converged structures. The average RMSD for the pairwise all-atom superposition of all 30 converged structures for the entire molecule is 0.65 Å. Individual base pair RMSD values are reported in Table 7. The average RMSD for the pairwise all-atom superposition of the structures excluding the terminal base pairs is 0.45 Å, indicating excellent convergence and precision.

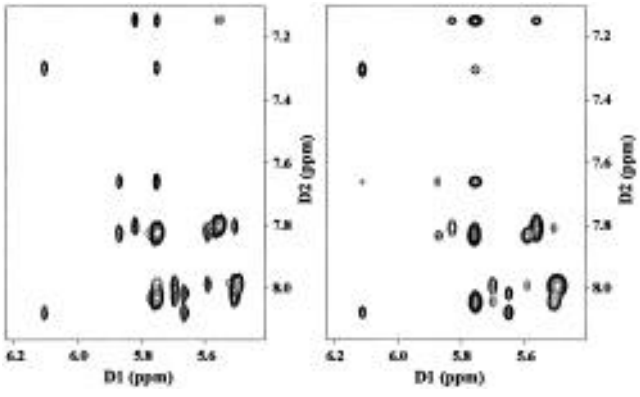

FIGURE 7: Experimental (left) and calculated (right) 400 ms NOESY spectra of (rGAGGUCUC)$_2$ showing the base (H8, H6, H2, H5) and sugar (H1′) region. The NOESY spectrum was calculated using a relaxation matrix approach (Boelens et al., 1988, 1989) using the average final structure and an isotropic correlation time of 3 ns.

Figure 7 shows the base (H8/H6/H2/H5) and sugar H1′ proton region from the experimental and calculated 400 ms NOESY spectra. Comparison suggests good qualitative agreement between the experimental and calculated spectra.

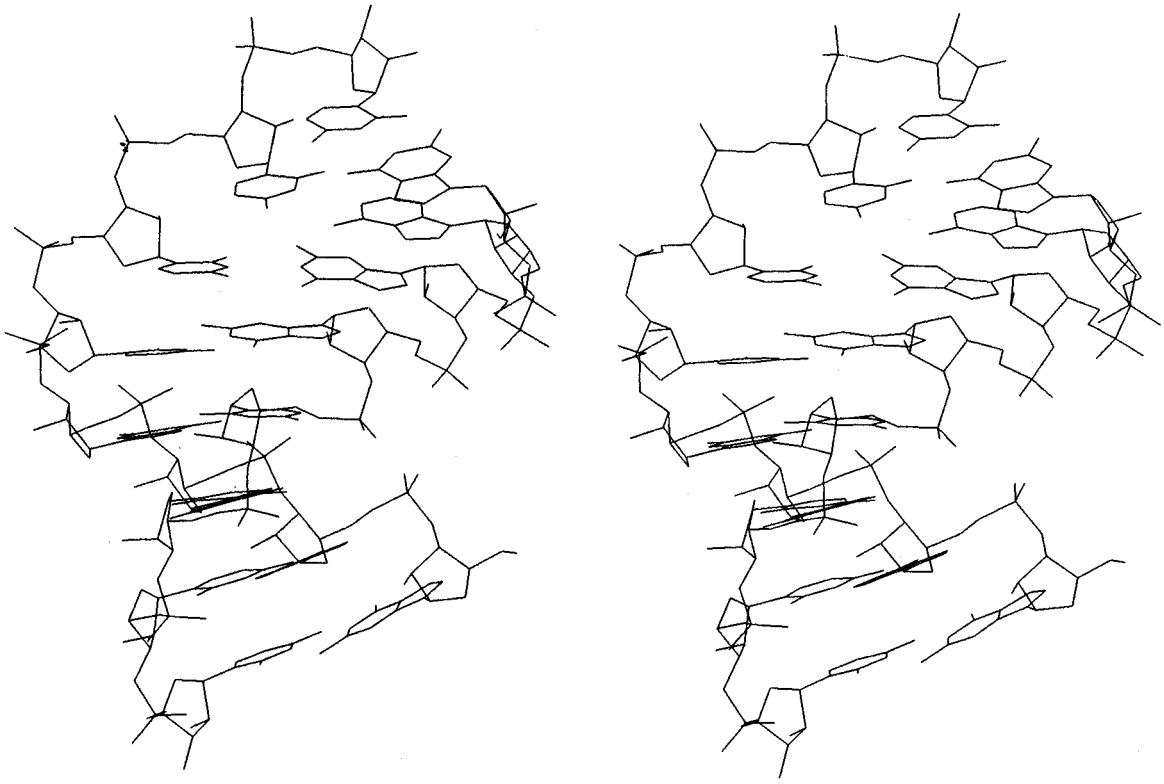

FIGURE 8: Stereoview of the average of the 30 converged structures of (rGAGGUCUC)$_2$.

Table 8: *R* Factors of the Final and Several Starting Structures from the Structure Determination of (rGAGGUCUC)$_2$

	R_1[a]	R_2	R_r
final structure[b]	0.53	0.64	0.52
A-form	1.12	1.51	1.03
B-form	1.04	1.30	0.97
random structure 1[c]	1.37	1.89	1.35
random structure 2	1.24	1.67	1.18

[a] *R* factors as defined in eqs 8−10. [b] Average of the 30 converged structures shown in Figure 7. [c] Starting structures with randomized torsion angles are included for comparison.

Table 9: Helical Parameters[a] for Base Pair Steps of the Averaged Structure of (rGAGGUCUC)$_2$

base step	helical twist (deg)	rise (Å)	slide (Å)	roll (deg)
GA / CU	30.2	3.0	−2.0	4.9
AG / UC	29.7	3.0	−2.3	8.4
GG / CU	37.8	2.7	−1.5	6.5
GU / UG	30.7	3.0	−2.1	3.6
UC / GG	37.3	2.7	−1.5	5.4
CU / GA	29.6	3.0	−2.3	8.9
UC / AG	30.2	3.0	−2.1	6.5
average	32.2	2.9	−2.0	6.3
A-form	32.7	2.8	−1.5	−0.4

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

The three *R* factors discussed in Materials and Methods (eqs 8−10) calculated for all well-resolved cross peaks for the experimental and calculated NOESY spectra are reported in Table 8 for the average of the 30 converged structures and 4 of the starting structures. These values indicate good agreement between the experimental and calculated spectra of the final structure. In addition, the final structure agreed with all distance restraints within 0.2 Å and torsion angle restraints within 2°. Thus the final structure is consistent with the experimental data, suggesting accuracy in the structure.

A stereoscopic view of the average of the converged structures is shown in Figure 8. Helical parameters calculated according to the Cambridge convention (Dickerson, 1989) and backbone torsion angles are reported in Tables 9−11. These indicate that the structure is primarily A-form with some small local deviations. The structure has all C3′-*endo* sugar puckers, and all bases are *anti* relative to their sugars. The helix has 11.2 base pairs per helical turn compared with 11.0 for A-form helixes. Twist angles surrounding the GU mismatches show a moderate 5° deviation from A-form. The helical displacement which is the most important parameter for distinguishing groove geometry, indicates that this structure is close in this respect to A-form with slightly greater displacement of the GU mismatches. The consequence is a deeper, narrower major groove and more shallow, wider minor groove at the mismatch as shown in Figure 9. The inclination is positive for all base pairs, and the helical rise for each base pair step is 3.0 Å or less. Thus both inclination and helical rise are similar to that of A-form. Intrastrand P−P distances as well as interstrand C1′−C1′ distances are also consistent with A-form helixes.

Examination of the backbone torsion angles in Table 11 shows little deviation from A-form geometry. These results show that the tandem mismatch, $^{5'\text{-GU-}3'}_{3'\text{-UG-}5'}$, can be incorporated into a helix without steric conflicts and without significant deviation in the conformation of the backbone from A-form.

Potential Surface Calculations. Contour plots of calculated potentials for GC, AU, and GU base pairs are shown in Figure 10. Results are qualitatively similar to those of Hunter (1993). AU is included to show the dramatic effect G has on a base pair's potential map. The potentials

Table 10: Helical Parameters[a] for Base Pairs of the Averaged Structure of (rGAGGUCUC)$_2$

base pair	propeller twist (deg)	buckle (deg)	inclination (deg)	displacement (Å)	interstrand $C1'_n-C1'_m$ distance[b] (Å)	intrastrand P_n-P_{n+1} distance[c] (Å)
G1·C8	−10.4	−5.3	11.0	−4.5	10.7	
A2·U7	0.4	−7.9	14.0	−3.8	10.6	5.9
G3·C6	−9.5	−7.5	13.3	−3.7	10.6	6.1
G4·U5	−9.8	2.1	11.1	−5.0	10.8	5.8
U5·G4	−9.7	−2.2	11.1	−5.0	10.8	5.7
C6·G3	−8.3	6.4	13.1	−3.7	10.6	6.1
U7·A2	−1.7	9.3	13.2	−3.8	10.6	6.0
C8·G1	−7.5	8.2	12.0	−4.5	10.7	
average	−7.1	0.4	12.3	−4.3	10.7	5.9
A-form	−13.8	0.2	16.9	−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.

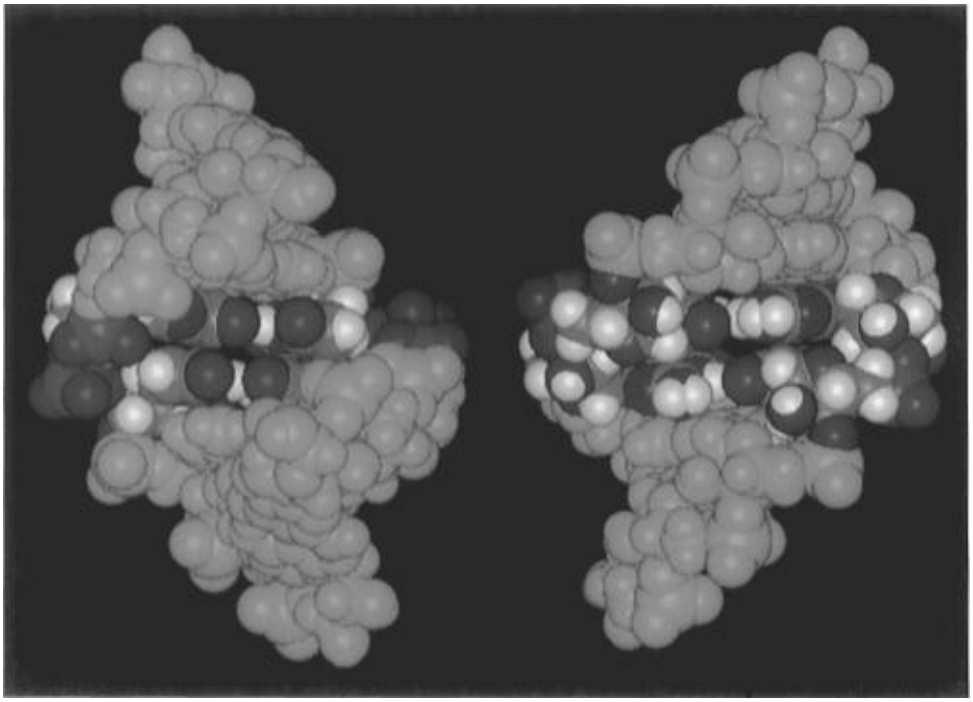

FIGURE 9: CPK models of the final structure of (rGAGGUCUC)$_2$ showing the (left) major and (right) minor grooves. The slightly greater displacement of the GU mismatches results in a deeper, narrower major groove and more shallow, wider minor groove.

Table 11: Backbone and Glycosidic Torsion Angles of the Averaged Structure of (rGAGGUCUC)$_2$

base	α	β	γ	δ	ε	ζ	χ
G1	na	na	65.3	79.7	−166.8	−63.3	−156.6
A2	−91.7	179.9	71.9	78.9	−162.7	−64.2	−164.6
G3	−76.7	178.4	59.1	78.4	−178.0	−76.0	−165.2
G4	−74.1	−174.6	59.0	82.7	−159.4	−53.2	−165.9
U5	−76.9	−176.9	52.0	77.1	−166.8	−61.3	−154.9
C6	−102.6	177.2	81.3	81.3	−162.7	−62.3	−168.0
U7	−82.8	−175.6	62.2	78.7	−169.0	−64.2	−165.8
C8	−82.4	−178.4	62.1	77.2	na	na	−166.5
average	−73.4	181.4	64.1	79.3	−166.5	−63.5	−163.4
A-form	−62	180	47.5	84	−152	−74	−166

generated by AU are more fragmented, and the regions of negative potential are smaller than for GC and GU base pairs. This implies that the electrostatic potential for AU base pairs contributes less to its stacking energy than for GC or GU. GC base pairs show a larger, intranucleotide area of strongly negative potential connecting G-N7 and G-O6, and GU mismatches show a similarly strong negative region which is internucleotide between G-N7, G-O6, and U-O4. The large areas of negative potential in GC and GU presumably affect stacking interactions between these base pairs. Figure 10C shows the GU potential calculated without bound water molecules. The calculation with bound waters resulted in an identical potential map with slight attenuation around the non-hydrogen-bonded U-O4 and G-N2 but did not affect the calculation significantly enough for further consideration.

DISCUSSION

Little is known about interactions determining stability and structure in non-Watson−Crick paired regions of RNA. GU mismatches are the most common non-Watson−Crick base pair in RNA. Their role as sites for enzyme and functionally important tertiary contacts is also being recognized (Strobel & Cech, 1995; McClain & Foss, 1988; Hou & Schimmel, 1988). Therefore, the understanding of GU mismatch structure and stability is important for understanding and even controlling RNA function. Tandem GU mismatches are a common motif in RNA (Gautheret et al., 1995). Measured stability increments at 37 °C for tandem GU motifs range from −4.8 to 0.1 kcal/mol for $\frac{5'\text{-GUGC-}3'}{3'\text{-CGUG-}5'}$ and $\frac{5'\text{-AGUU-}3'}{3'\text{-UUGA-}5'}$, respectively (see Table 2). Despite the large variation in

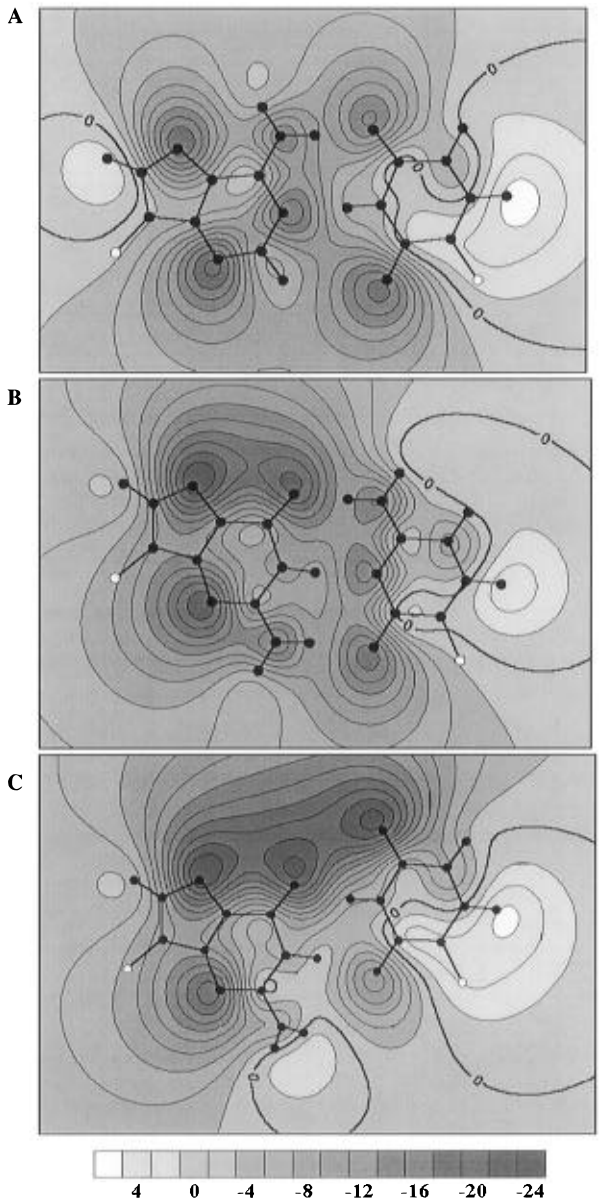

FIGURE 10: Potential maps of (A) AU, (B) GC, and (C) GU base pairs. Calculations were performed using the Amber charge set (Cornell et al., 1995), and geometries were taken from the Biosym nucleic acid library for AU and GC and the average final structure from this study for GU. The potential was computed at a distance of 1.4 Å, one-half the average A-form rise. The hollow atom indicates the C1′ position. Methyl hydrogens included in the calculation are not displayed. The potential scale is given in units of $e/4\pi\epsilon_0 = 14.4$ V.

thermodynamic stabilities, tandem GU mismatches are expected to have two hydrogen bonds per mismatch and to induce only minor distortions in the backbone. Therefore, they provide a convenient system for investigating other factors important for determining stabilities of non-Watson−Crick regions in RNA.

When the nearest-neighbor model is used to fit stabilities of duplexes with GU mismatches, a reasonable fit is obtained only when sequences containing the motif $\begin{smallmatrix}5'\text{-GGUC-}3'\\3'\text{-CUGG-}5'\end{smallmatrix}$ or the motifs $\begin{smallmatrix}5'\text{-CGUG-}3'\\3'\text{-GUGC-}5'\end{smallmatrix}$, $\begin{smallmatrix}5'\text{-AGUU-}3'\\3'\text{-UUGA-}5'\end{smallmatrix}$, and $\begin{smallmatrix}5'\text{-UGUA-}3'\\3'\text{-AUGU-}5'\end{smallmatrix}$ are excluded (He et al., 1991). Similarly, in comparison with other $\begin{smallmatrix}5'\text{-GU-}3'\\3'\text{-UG-}5'\end{smallmatrix}$ mismatch-containing sequences, $\begin{smallmatrix}5'\text{-GGUC-}3'\\3'\text{-CUGG-}5'\end{smallmatrix}$-containing sequences show remarkable stability, averaging 3.4 kcal/mol more favorable than sequences with $\begin{smallmatrix}5'\text{-CGUG-}3'\\3'\text{-GUGC-}5'\end{smallmatrix}$ motifs. Moreover, the $\begin{smallmatrix}5'\text{-GGUC-}3'\\3'\text{-CUGG-}5'\end{smallmatrix}$, motif occurs in the intensely studied group I intron from *T. thermophila* (Michel & Westhof, 1990; Cech et al., 1994; Cech, 1990) and in the signal recognition particle RNA from *H. japonicus* (Larsen & Zwieb, 1996). Thus the structure of $\begin{smallmatrix}5'\text{-GGUC-}3'\\3'\text{-CUGG-}5'\end{smallmatrix}$, is of interest for several reasons.

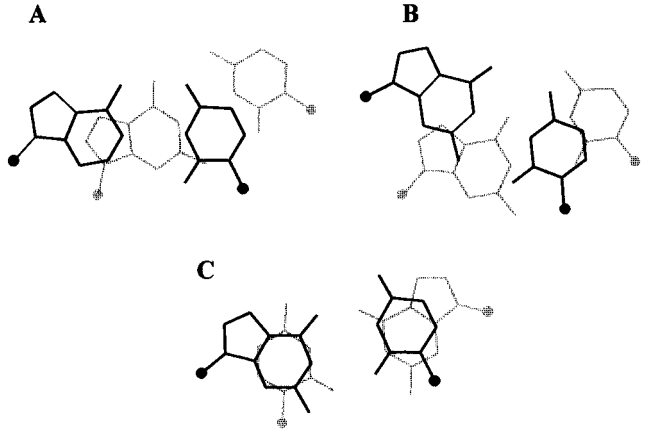

FIGURE 11: Stacking patterns of the (A) $\begin{smallmatrix}5'\text{-AG-}3'\\3'\text{-UC-}5'\end{smallmatrix}$, (B) $\begin{smallmatrix}5'\text{-GG-}3'\\3'\text{-CU-}5'\end{smallmatrix}$, and (C) $\begin{smallmatrix}5'\text{-GU-}3'\\3'\text{-UG-}5'\end{smallmatrix}$ steps. The first base pair in each sequence is in bold. The $\begin{smallmatrix}5'\text{-AG-}3'\\3'\text{-UC-}5'\end{smallmatrix}$ step has essentially the same stacking geometry as A-form helices (Saenger, 1989). $\begin{smallmatrix}5'\text{-GG-}3'\\3'\text{-CU-}5'\end{smallmatrix}$ shows deviation from A-form with almost no overlap with the mismatch. The tandem mismatches have considerable overlap.

The structure displayed in Figure 8 and its backbone torsion angles reported in Table 11 reveal that incorporating the $\begin{smallmatrix}5'\text{-GGUC-}3'\\3'\text{-CUGG-}5'\end{smallmatrix}$ motif into a helix has little effect on the local backbone geometry. Helical parameters including adjacent, intrastrand phosphate distances and interstrand C1′−C1′ distances have values very close to A-form. The only significant differences are in the slight overtwisting at the $\begin{smallmatrix}5'\text{-GG-}3'\\3'\text{-CU-}5'\end{smallmatrix}$ steps and the helical displacement of the two GU mismatches, resulting in the G of each mismatch displaced somewhat toward the minor groove. Helical distortion due to incorporating the tandem mismatch can be seen by comparison of the base stacking in the structure. Figure 11 shows the stacking interactions of each nonterminal base pair step. The $\begin{smallmatrix}5'\text{-AG-}3'\\3'\text{-UC-}5'\end{smallmatrix}$ step is nearly identical to A-form, Watson−Crick base pair stacking (Saenger, 1989). However, the base pair step adjacent to the mismatches, $\begin{smallmatrix}5'\text{-GG-}3'\\3'\text{-CU-}5'\end{smallmatrix}$, shows deviation from A-form with very little overlap as a result of the 5° overtwisting of this step. Between the tandem mismatches, $\begin{smallmatrix}5'\text{-GU-}3'\\3'\text{-UG-}5'\end{smallmatrix}$, there is considerable overlap with the six-membered purine rings overlying the pyrimidine rings.

Other structures containing GU or GT mismatches show similar patterns. The crystal structure of the A-DNA helix $\begin{smallmatrix}5'\text{-GGGGTCCC-}3'\\3'\text{-CCCTGGGG-}5'\end{smallmatrix}$ by Kneale et al. (1985) also shows little effect on the backbone torsion angles and overtwisting of the helix at the step 5′ of the G in the GT mismatches. The A-DNA motif $\begin{smallmatrix}5'\text{-GGTC-}3'\\3'\text{-CTGG-}5'\end{smallmatrix}$ in the sequence of Kneale et al. (1985) shows stacking interactions nearly identical to those shown in Figure 11. The structure is consistent with previous observations of the geometry of single GU mismatches suggesting that the displacement of the G toward the minor groove results in decreased stacking of the G toward its 5′ side (Mizuno & Sundaralingam, 1978). This rationalized the observation that G's in GU pairs are preferentially at the base of loops 5′ to the adjacent helix in many secondary structures. NMR structures of RNAs containing GU mismatches are also consistent with this model (White et al., 1992; Allain et al., 1995).

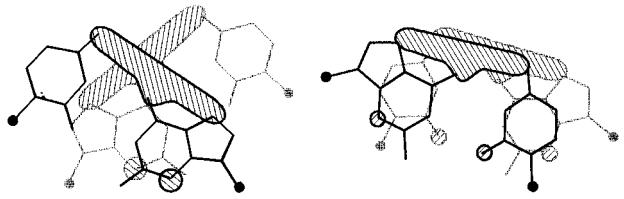

FIGURE 12: Stacking patterns of the tandem mismatches of the $\begin{smallmatrix}5'\text{-TG-}3'\\3'\text{-GT-}5'\end{smallmatrix}$ step (left) from the A-DNA crystal structure of Rabinovich et al. (1988) and the $\begin{smallmatrix}5'\text{-GU-}3'\\3'\text{-UG-}5'\end{smallmatrix}$ step (right). Included are representations of the regions of high negative potential from Figure 10. The $\begin{smallmatrix}5'\text{-GU-}3'\\3'\text{-UG-}5'\end{smallmatrix}$ motif shows considerably more overlap of these regions than the $\begin{smallmatrix}5'\text{-TG-}3'\\3'\text{-GT-}5'\end{smallmatrix}$ motif. In accord, $\begin{smallmatrix}5'\text{-GUGC-}3'\\3'\text{-CGUG-}5'\end{smallmatrix}$ motifs (see Table 2) are 0.7 kcal/mol more stable than $\begin{smallmatrix}5'\text{-GGUC-}3'\\3'\text{-CUGG-}5'\end{smallmatrix}$ motifs.

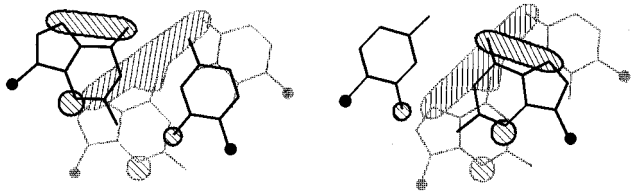

FIGURE 13: Stacking patterns of the $\begin{smallmatrix}5'\text{-GG-}3'\\3'\text{-CU-}5'\end{smallmatrix}$ step (left) and the construct, $\begin{smallmatrix}5'\text{-CG-}3'\\3'\text{-GU-}5'\end{smallmatrix}$ (right), made by replacing the GC pair in the $\begin{smallmatrix}5'\text{-GG-}3'\\3'\text{-CU-}5'\end{smallmatrix}$ step with the isosteric CG base pair. Included are representations of the regions of high negative potential from Figure 10. The $\begin{smallmatrix}5'\text{-GG-}3'\\3'\text{-CU-}5'\end{smallmatrix}$ motif shows little overlap of these regions while in contrast $\begin{smallmatrix}5'\text{-CG-}3'\\3'\text{-GU-}5'\end{smallmatrix}$ exhibits considerable overlap.

The similarity between the $\begin{smallmatrix}5'\text{-GGTC-}3'\\3'\text{-CTGG-}5'\end{smallmatrix}$ motif in the crystal structure (Kneale et al., 1985) and the $\begin{smallmatrix}5'\text{-GGUC-}3'\\3'\text{-CUGG-}5'\end{smallmatrix}$ motif in the NMR structure suggests the crystal structure for $\begin{smallmatrix}5'\text{-GCGTGCGC-}3'\\3'\text{-CGCGTGCG-}5'\end{smallmatrix}$ in the A-form (Rabinovich et al., 1988) may also be similar to its RNA counterpart. Figure 12 compares stacking between tandem mismatches in the $\begin{smallmatrix}5'\text{-TG-}3'\\3'\text{-GT-}5'\end{smallmatrix}$ motif in the A-DNA crystal structure with the $\begin{smallmatrix}5'\text{-GU-}3'\\3'\text{-UG-}5'\end{smallmatrix}$ motif. Outlines of regions of high negative potential based on the calculations illustrated in Figure 10 are also shown. There is apparently more unfavorable overlap of negative potentials for the $\begin{smallmatrix}5'\text{-GU-}3'\\3'\text{-UG-}5'\end{smallmatrix}$ motif. In RNA tandem mismatches, the $\begin{smallmatrix}5'\text{-GUGC-}3'\\3'\text{-CGUG-}5'\end{smallmatrix}$ motif is 0.7 kcal/mol more stable than the $\begin{smallmatrix}5'\text{-GGUC-}3'\\3'\text{-CUGG-}5'\end{smallmatrix}$ motif (Table 2). With other adjacent Watson−Crick pairs, the $\begin{smallmatrix}5'\text{-UG-}3'\\3'\text{-GU-}5'\end{smallmatrix}$ motif is favored by an average of 2.5 kcal/mol (Table 2). The comparisons suggest that electrostatics may be an important determinant of stabilities of mismatches in RNA.

As shown in Table 2, the $\begin{smallmatrix}5'\text{-GGUC-}3'\\3'\text{-CUGG-}5'\end{smallmatrix}$ motif is an average of 3.4 kcal/mol more stable than $\begin{smallmatrix}5'\text{-CGUG-}3'\\3'\text{-GUGC-}5'\end{smallmatrix}$. Figure 13 compares the overlap of negative potentials for the $\begin{smallmatrix}5'\text{-GG-}3'\\3'\text{-CU-}5'\end{smallmatrix}$ step in the NMR structure reported here with that expected for a $\begin{smallmatrix}5'\text{-CG-}3'\\3'\text{-GU-}5'\end{smallmatrix}$ step if the GC base pair in the NMR structure is inverted to the isosteric CG pair. Again, the predicted overlap is more unfavorable for the less stable $\begin{smallmatrix}5'\text{-CGUG-}3'\\3'\text{-GUGC-}5'\end{smallmatrix}$ sequence. Inspection of Figures 12 and 13 indicates that the modeled $\begin{smallmatrix}5'\text{-CGUG-}3'\\3'\text{-GUGC-}5'\end{smallmatrix}$ motif has unfavorable electrostatic interactions at both the $\begin{smallmatrix}5'\text{-CG-}3'\\3'\text{-GU-}5'\end{smallmatrix}$ and $\begin{smallmatrix}5'\text{-GU-}3'\\3'\text{-UG-}5'\end{smallmatrix}$ steps. This suggests that the real structure of this motif may shift to avoid these interactions. There is some evidence to support this suggestion. He et al. (1991) measured 1-D imino proton NOE's for GUCGUGAC at 0.5 °C. The fractional transfers between the G4 and U5 imino protons averaged only 0.17, while the fractional transfer between either G4 or U5 and G6 was less than 0.04. In contrast, for GAGGUCUC at 0.5 °C, Table 3 shows that the fractional transfer between the G4 and U5 imino protons averages 0.47, while the fractional transfer between either G4 or U5 and G3 averages 0.07. Thus the local structures of the $\begin{smallmatrix}5'\text{-CGUG-}3'\\3'\text{-GUGC-}5'\end{smallmatrix}$ and $\begin{smallmatrix}5'\text{-GGUC-}3'\\3'\text{-CUGG-}5'\end{smallmatrix}$ motifs appear different. NMR results have shown that the local structures of $\begin{smallmatrix}5'\text{-GGAC-}3'\\3'\text{-CAGG-}5'\end{smallmatrix}$ and $\begin{smallmatrix}5'\text{-CGAG-}3'\\3'\text{-GAGC-}5'\end{smallmatrix}$ are dramatically different, and it has been suggested that this may be due to electrostatic interactions (Wu & Turner, 1996). The available results provide a working hypothesis that electrostatic interactions are an important determinant of both stabilities and structures of tandem mismatches in RNA.

REFERENCES

Allain, F. H.-T., & Varani, G. (1995) *J. Mol. Biol. 250*, 333−353.
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.
Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., Vol. I, p 597, CRC Press, Cleveland, OH.
Borer, P. N., Dengler, B., & Tinoco, I., Jr. (1974) *J. Mol. Biol. 86*, 843−853.
Cech, T. R. (1990) *Annu. Rev. Biochem. 59*, 543−568.
Cech, T. R., Damberger, S., & Gutell, R. R. (1994) *Nat. Struct. Biol. 1*, 273−280.
Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., & Kollman, P. A. (1995) *J. Am. Chem. Soc. 117*, 5179−5197.
Crick, F. H. C. (1966) *J. Mol. Biol. 19*, 548−555.
Damberger, S. H., & Gutell, R. R. (1994) *Nucleic Acids Res. 22*, 3508−3510.
Dickerson, R. E. (1989) *Nucleic Acids Res. 17*, 1797−1803.
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.
Gabriel, K., Schneider, J., & McClain, W. H. (1996) *Science 271*, 195−197.
Gautheret, D., Konings, D., & Gutell, R. R. (1995) *RNA 1*, 807−814.
Geerdes, H. A. M., & Hilbers, C. W. (1979) *FEBS Lett. 107*, 125−128.
Gonzalez, C., Rullmann, J. A. C., Bonvin, A. M. J. J., Boelens, R., & Kaptein, R. (1991) *J. Magn. Reson. 91*, 659−664.
Gorenstein, D. G. (1984) *Phosphorus-31 NMR: Principles and Applications*, Academic Press, New York.
Gralla, J., & Crothers, D. M. (1973) *J. Mol. Biol. 78*, 301−319.
Gutell, R. R. (1994) *Nucleic Acids Res. 22*, 3502−3507.
Gutell, R. R., Gray, M. W., & Schnare, M. N. (1993) *Nucleic Acids Res. 21*, 3055−3074.
Hare, D. R., Wemmer, D. E., Chou, S. H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319−336.
Harvey, S. C., Mao, B., & McCammon, J. A. (1984) *Science 223*, 1189−1191.
He, L., Kierzek, R., SantaLucia, J., Jr., Walter, A. E., & Turner, D. H. (1991) *Biochemistry 30*, 11124−11132.
Holbrook, S. R., Cheong, C., Tinoco, I., Jr., & Kim, S.-H. (1991) *Nature 353*, 579−581.
Hore, P. J. (1983) *J. Magn. Reson. 55*, 283−300.
Hou, Y. M., & Schimmel, P. (1988) *Nature 333*, 140−145.
Hunter, C. A. (1993) *J. Mol. Biol. 230*, 1025−1054.
Johnson, P. D., & Redfield, A. G. (1981) *Biochemistry 20*, 1147−1156.
Kneale, G., Brown, T., & Kennard, O. (1985) *J. Mol. Biol. 186*, 805−814.
Knitt, D. S., Narlikar, G. J., & Herschlag, D. (1994) *Biochemistry 33*, 13864−13879.
Larsen, N., & Zwieb, C. (1996) *Nucleic Acids Res. 24*, 80−81.
McClain, W. H., & Foss, K. (1988) *Science 240*, 793−796.
Michel, F., & Westhof, E. (1990) *J. Mol. Biol. 216*, 585−610.
Mizuno, H., & Sundaralingam, M. (1978) *Nucleic Acids Res. 6*, 4451−4461.

Petersheim, M., & Turner, D. H. (1983) *Biochemistry 22*, 264−268.
Pyle, A.-M., Moran, S., Strobel, S. A., Chapman, T., Turner, D. H., & Cech, T. R. (1994) *Biochemistry 33*, 13856−13863.
Rabinovich, D., Taran, T., Eisenstein, M., & Shakked, Z. (1988) *J. Mol. Biol. 200*, 151−161.
Richards, E. G. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) 3rd ed., Vol. I, p 579, CRC Press, Cleveland, OH.
Saenger, W. (1989) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
Scheek, R. M., Russo, N., Boelens, R., & Kaptein, R. (1983) *J. Am. Chem. Soc. 105*, 2914−2916.
Shibata, M., Zielinski, T. J., & Rein, R. (1991) *Biopolymers 31*, 211−232.
Simpson, L., & Thiemann, O. H. (1995) *Cell 81*, 837−840.
Sklenar, V., Miyashiro, H., Zon, G., Miles, H. T., & Bax, A. (1986) *FEBS Lett. 208*, 94−98.
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.
Sugimoto, N., Kierzek, R., Freier, S. M., & Turner, D. H. (1986) *Biochemistry 26*, 5755−5759.
Usman, N., Ogilvie, K. K., Jiang, M.-V., & Cedergren, R. (1987) *J. Am. Chem. Soc. 109*, 7845−7854.
van Knippenberg, P. H., Formenoy, L. J., & Heus, H. A. (1990) *Biochim. Biophys. Acta 1050*, 14−17.
Varani, G., & Tinoco, I., Jr. (1991) *Q. Rev. Biophys. 24*, 479−532.
Weiner, S. J., Kollman, P. A., Nguyen, D. T., & Case, D. T. (1986) *J. Comput. Chem. 7*, 230−252.
White, S. A., Nilges, M., Huang, A., Brünger, A. T., & Moore, P. B. (1992) *Biochemistry 31*, 1610−1621.
Wijmenga, S. S., Mooren, M. M. W., & Hilbers, C. W. (1993) in *NMR of Macromolecules: A Practical Approach* (Roberts, G. C. K., Ed.) Oxford University Press, Inc., New York.
Wu, M., & Turner, D. H. (1996) *Biochemistry 35*, 9677−9689.
Wu, M., McDowell, J. A., & Turner, D. H. (1995) *Biochemistry 34*, 3204−3211.

BI9615710