Sheared A_{anti}·A_{anti} Base Pairs in a Destabilizing 2 × 2 Internal Loop: The NMR Structure of 5′(rGGCAAGCCU)$_2$[†,‡]

Brent M. Znosko,[§] Mark E. Burkard,[§] Susan J. Schroeder,[§] Thomas R. Krugh,[§] and Douglas H. Turner*[,§,||]

Department of Chemistry, University of Rochester, Rochester, New York 14627-0216, and Department of Pediatrics and Center for Human Genetics and Molecular Pediatric Disease, School of Medicine and Dentistry, University of Rochester, Rochester, New York 14642

Received April 30, 2002; Revised Manuscript Received September 30, 2002

ABSTRACT: The 5′(rGGCAAGCCU)$_2$ duplex contains tandem A·A pairs. The three-dimensional structure of the 5′(rGGCAAGCCU)$_2$ duplex was modeled by molecular dynamics and energy minimization with NMR-derived distance and dihedral angle restraints. Although the 5′(rCAAG)$_2$ loop is thermodynamically destabilizing by 1.1 kcal/mol, the tandem A·A pairs adopt a predominant conformation: a sheared anti−anti (A·A trans Hoogsteen/Sugar-edge) alignment similar to that observed in the crystal structure of the P4−P6 domain of the *Tetrahymena thermophila* intron [Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science 273*, 1678−1685]. The NMR-derived structure of the 5′(rGGCAAGCCU)$_2$ duplex exhibits cross-strand hydrogen bonds from N3 of A4 to an amino hydrogen of A5 and from the 2′ oxygen of the A4 sugar to the other amino hydrogen of A5. An intrastrand hydrogen bond is formed from the 2′ OH hydrogen of A4 to O5′ of A5. The cross-strand A5 bases are stacked. The Watson−Crick G−C regions are essentially A-form. The sheared anti−anti (A·A trans Hoogsteen/Sugar-edge) alignment provides potential contact sites for tertiary interactions and, therefore, is a possible target site for therapeutics. Thus, thermodynamically destabilizing internal loops can be preorganized for tertiary interactions or ligand binding.

Due to the presence of various hydrogen-bonding donors and acceptors, a purine base can form hydrogen bonds with another purine base to form a purine−purine pair (*1*−*3*). Purine−purine pairs have been observed in many natural RNAs, including the yeast initiator tRNA (*4*), the triplet coding for Gly12 in the *K-ras* mRNA (*5*), the P4−P6 domain of the *Tetrahymena thermophila* group I intron (*6*, *7*), and the loop B domain of the hairpin ribozyme (*8*). The G·A pair is the most extensively studied purine−purine pair, in both RNA and DNA, and has been observed in several different geometries. In crystals, the G·A pair has been found as a sheared G_{anti}·A_{anti} pair (*9*, *10*), also classified as A·G trans Hoogsteen/Sugar-edge (*2*, *11*), a G_{anti}·A_{syn} pair (G·A cis Watson−Crick/Hoogsteen) (*12*−*14*), and a G_{syn}·A^+_{anti} pair (A^+·G cis Watson−Crick/Hoogsteen) (*12*, *15*). In solution, the G·A pair has been observed as an imino hydrogen-bonded G_{anti}·A_{anti} pair (G·A cis Watson−Crick/Watson−Crick) (*16*−*19*), a sheared G_{anti}·A_{anti} pair (A·G trans Hoogsteen/Sugar-edge) (*20*−*24*), and a G_{syn}·A^+_{anti} pair (A^+·G cis Watson−Crick/Hoogsteen) (*25*). From the chemical structure of adenosine, it is evident that various types of A·A pairs are also possible (*1*−*3*). Three different configurations for A·A pairs have been reported in solution, namely, the A_{syn}·A_{anti} configuration (A·A trans Watson−Crick/Watson−Crick) (*26*), the face-to-face A_{anti}·A_{anti} configuration (A·A cis Watson−Crick/Watson−Crick) (*5*), and the sheared A_{anti}·A_{anti} configuration (A·A trans Hoogsteen/Sugar-edge) (*27*, *28*). Nagaswamy et al. (*29*, *30*) have compiled a database of noncanonical interactions in known RNA structures, including various types of purine·purine pairs found by X-ray crystallography and NMR spectroscopy.

A recent analysis of RNA secondary structures revealed that nucleotides not involved in Watson−Crick pairs are frequently adenosines (*31*). For example, in known cases of two consecutive mismatches in a database of rRNA, at least three out of the four nucleotides are adenosines in ∼16% of all occurrences. The most frequent motif of consecutive mismatches is a potential tandem A·A pair, occurring in ∼4% of all cases. Potential tandem A·A pairs are also a common motif in other types of RNA. Three rRNA secondary structures, *Campylobacter coli*, *Campylobacter jejuni*, and *Cosmarium botrytis*, contain the (rCAAG)$_2$ motif studied here (*32*). Also, the crystal structure of the large ribosomal subunit contains a tandem $A_{429/430}$·$A_{407/408}$ pair (A·A cis Watson−Crick/Watson−Crick) (*33*).

Several studies have investigated the context dependence of the structure and stability of tandem purine-purine pairs in RNA. For example, NMR studies on tandem G·A pairs have shown that the G·A pairs of the 5′(rGGCGAGCC)$_2$ duplex are in the sheared G_{anti}·A_{anti} conformation (*21*), while those in the 5′(rGGCAGGCC)$_2$ duplex are in the imino

[†] This work was supported by NIH Grants GM 22939 (D.H.T.) and GM 53826 (T.R.K.).

[‡] Protein Data Bank entry 1MUV.

* To whom correspondence should be addressed. Phone: (585) 275-3207. Fax: (585) 506-0205. E-mail: Turner@chem.rochester.edu.

[§] Department of Chemistry.

[||] Department of Pediatrics and Center for Human Genetics and Molecular Pediatric Disease.

10.1021/bi020326f CCC: $22.00
Published on Web 11/21/2002

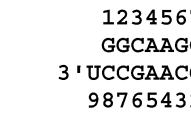

FIGURE 1: Schematic representation of the duplex studied by NMR.

hydrogen-bonded $G_{anti}{\cdot}A_{anti}$ conformation (*18*). Thermodynamic studies have shown that in such duplexes, $\binom{GA}{AG}$ and $\binom{AG}{GA}$ tandem pairs are stabilizing, with ΔG°_{37} of -0.7 kcal/mol. In contrast, $\binom{AA}{AA}$ tandem pairs flanked by two G−C pairs are destabilizing, with an average ΔG°_{37} of 1.3 kcal/mol (*34*, *35*). This is within experimental error of the ΔG°_{37}'s for $\binom{CA}{AC}$, $\binom{CU}{UC}$, $\binom{CC}{CC}$, and $\binom{UC}{CU}$ tandem pairs (*34*, *35*). These destabilizing loops have been assumed to be relatively unstructured. To probe the structural properties of the destabilizing $\binom{AA}{AA}$ motif, the 5′(rGGCAAGCCU)$_2$ duplex (Figure 1) was studied by NMR. The 3′ dangling uracils are present to increase duplex stability. The $\binom{AA}{AA}$ motif is structured, and a model is derived for the predominant structure on the basis of restrained molecular dynamics.

MATERIALS AND METHODS

RNA Synthesis and Purification. The 5′rGGCAAGCCU oligonucleotide was synthesized on CPG[1] support with an Applied Biosystems 392 DNA/RNA synthesizer using the phosphoramidite method (*36*, *37*). The support and phosphoramidites with 2′-*tert*-butyl-dimethyl-silyl ether protecting groups were acquired from Glen Research (Baltimore, MD). Base-labile protecting groups were removed by treatment with a 3:1 (v/v) ammonia/ethanol solution at 55 °C overnight. A disposable filter column was used to separate the oligonucleotide from the support. The silyl protecting group was removed by incubation with anhydrous 1 M triethylamine hydrogen fluoride in pyridine (50 equivalents) at 55 °C for 48 h. The sample was extracted with diethyl ether to remove organic impurities and then lyophilized before redissolving in 5 mM ammonium bicarbonate at pH 7.0. This solution was loaded onto a Waters Sep-Pak C18 chromatography column to remove excess inorganic salts. The oligonucleotide was purified on a large preparative Whatman TLC plate (20 cm × 20 cm, 500 μm thick) with 55:35:10 (v/v/v) 1-propanol/ammonia/water as the solvent. The main product band was identified by UV shadowing and extracted from the silica with distilled water. The Sep-Pak procedure was repeated. The purity of the 5′rGGCAAGCCU oligonucleotide was checked by reverse-phase HPLC and was greater than 95%.

Optical Melting Experiments. The purified oligonucleotide was lyophilized and redissolved in 1 M NaCl, 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA at pH 7.0. Curves of absorbance at 280 nm versus temperature were acquired with a Gilford 250 spectrophotometer and a Gilford 2527 thermoprogrammer, using a heating rate of 1 °C/min. Melting curves from 1.3×10^{-5} to 6.5×10^{-4} M strand concentration were fit to a two-state model, assuming linear sloping baselines and temperature-independent ΔH° and ΔS° (*38*, *39*). Additionally, the T_M values at different concentrations were used to calculate thermodynamic parameters according to Borer et al. (*40*):

$$T_M^{-1} = (2.303\,R/\Delta H^\circ)\log C_T + (\Delta S^\circ/\Delta H^\circ) \quad (1)$$

Sample Preparation for NMR. The sample was dialyzed in a Gibco Life Technologies microdialysis system with a 1000 MW cutoff Spectro-Por #7 dialysis membrane. The microdialysis system was attached to a Rainin Dynamax peristaltic pump. The sample (~600 μL) was dialyzed first against doubly distilled H_2O for 12 h and then against 40 mM NaCl, 1.5 mM KH_2PO_4, 3.5 mM K_2HPO_4, 0.05 mM Na_2EDTA, pH 6.0 for 24 h. The sample was lyophilized and redissolved in 300 μL of D_2O, resulting in final concentrations of 80 mM NaCl, 3 mM KH_2PO_4, 7 mM K_2HPO_4, 0.5 mM Na_2EDTA. For the nonexchangeable proton sample, D_2O exchange was performed with lyophilization from 99.96% D_2O. The sample was redissolved in 300 μL of 99.996% D_2O (Cambridge Isotope Laboratories). The strand concentration of 5′rGGCAAGCCU for this sample was ~1 mM. NMR spectra of the nonexchangeable protons were recorded at 30 °C, pH 6.9.

NMR spectra of the exchangeable protons of ~1 mM 5′rGGCAAGCCU were recorded at 0−45 °C in 80 mM NaCl, 3 mM KH_2PO_4, 7 mM K_2HPO_4, 0.5 mM Na_2EDTA, pH 6.8 in 90:10 H_2O/D_2O.

NMR Spectroscopy. All spectra were acquired on a Varian Inova 500 MHz spectrometer. One-dimensional imino proton spectra were acquired with a binomial pulse sequence. Offset delays were selected to maximize peak intensity at 12.5 ppm. For 1D NOE experiments, selective proton decoupling preceded the acquisition during a 3 s delay between transients. Difference spectra were obtained by subtracting from a reference spectrum, which was acquired with the decoupler set downfield of the imino resonances. One-dimensional nonexchangeable proton spectra were acquired at 20−50 °C with a spectral width of 5000 Hz. A 2D NOESY spectrum of exchangeable protons for the sample dissolved in 90:10 (v:v) H_2O:D_2O was recorded with a 150 ms mixing time using a WATERGATE water suppression sequence. This spectrum was acquired with 4k points and a spectral width of 11 kHz at 0 °C.

The NOESY spectra of samples in D_2O were acquired at 100, 200, and 400 ms mixing times. The FIDs were acquired over 256 t_1 increments, each having k points with 64 scans per FID and a recycle delay of 2.8 s. The DQF−COSY spectrum for sugar assignments was acquired at 30 °C with 256 t_1 increments, each having 4k data points. The spectral width was 4200 Hz and was centered on the HDO resonance. For each FID, 64 scans were collected with a recycle delay of 2.5 s.

A one-dimensional phosphorus spectrum was acquired at 30 °C, with a spectral width of 2450 Hz, 4k data points, and 512 transients. The ^1H-^{31}P HETCOR experiment followed the method of Sklenar et al. (*41*). The spectrum was acquired at 30 °C with a spectral width of 1000 Hz in the direct, ^1H, dimension and 1200 Hz in the indirect, ^{31}P, dimension. FIDs were acquired over 256 t_1 intervals with 192 scans per interval and each having 1k complex points per FID. A TOCSY spectrum was acquired at 30 °C with a 120 ms

[1] Abbreviations: CPG, controlled pore glass; HETCOR, heteronuclear correlation spectroscopy; HMQC, heteronuclear multiple quantum correlation spectroscopy; rMD, restrained molecular dynamics; rmsd, root-mean-squared deviation; T_m, melting temperature; TOCSY, total correlation spectroscopy.

mixing time, 4k data points, and a spectral width of 4000 Hz in both dimensions. A natural abundance ^{13}C HMQC spectrum was acquired with a 5000 Hz spectral width for proton and 15 000 Hz spectral width for carbon. The HMQC spectrum consisted of 512 increments, each having 2k data points with 512 scans per FID. The T_1 relaxation times were measured by inversion recovery.

Two-dimensional spectra were processed with the Felix 2000 software package (Molecular Simulations Inc.). Proton spectra were referenced to H_2O or HDO at 4.74 ppm at 30 °C. Phosphorus resonances were referenced to the phosphate peak at 0.0 ppm at 30 °C (pH 6.9).

Restraint Generation. NOESY cross-peak volumes were determined by measuring the volume of interest with the Felix software package. Distance restraints were generated from 100 and 200 ms mixing time NOE volumes (excluding H5′ and H5″ because assignments are not stereospecific) with pyrimidine H5−H6 cross-peaks as a reference at 2.45 Å. Other proton−proton distances were calculated from the two-spin approximation by $1/r^6$ scaling. Error limits of ±30% of the calculated distances were assigned to allow for spin diffusion, phasing, baseline abnormalities, and noise. Included for structural refinement were a total of 83 NMR-derived interproton distance restraints per strand (49 intranucleotide, 32 internucleotide, and 2 cross-strand), 18 hydrogen bond restraints for six Watson−Crick G−C pairs (limits of 1.8−2.5 Å), and 46 dihedral angle restraints per strand. No hydrogen bonding restraints between A·A pairs were used. Because coupling patterns in the DQF−COSY and HETCOR experiments suggested typical A-form geometry, parameters for the dihedral angles were loosely restrained to those observed in canonical A-form RNA duplexes for all residues except some angles of the A4 and A5 residues to allow for conformational freedom in the loop. Because NMR data indicate that 5′rGGCAAGCCU forms a symmetric self-complementary duplex, identical restraints were applied to both strands.

Simulated Annealing. Models of the 5′(rGGCAAGCCU)$_2$ duplex consistent with NMR data were derived from restrained energy minimization and simulated annealing with the Discover 95 package on a Silicon Graphics computer. Initially, five starting structures were generated with Biosym InsightII software: standard A-form RNA, B-form RNA, and three A-form structures with varying axial rise and twist angle parameters. To check for bias from the starting structures, the same restrained energy mimizimation and simulated annealing protocol was also performed with five random structures as starting structures. Four of the five refined structures had essentially the same structure as that generated from the helical starting stuctures but with a larger RMSD. Calculations used the AMBER 95 force field (*42*) in addition to flat-bottom restraint potentials with force constants of 25 kcal/(mol Å^2) for distance restraints and 50 kcal/(mol rad^2) for torsion angle restraints. Calculations were performed both without solvent and with water as the solvent. For electrostatics, the cell-multipole method was used with a distance-dependent dielectric constant of $\epsilon = 2r$, where r is distance in Ångstroms. For van der Waals interactions, group-based summation was used with an 18 Å cutoff. The simulation (*43*) involved a total of 14 steps: (1) Coulombic interactions were turned off and van der Waals interactions were scaled to 1%, (2) 4000 steps of steepest descent energy minimization were performed to relieve steric clashes or other high energy interactions, (3) 4 ps of rMD with 1 fs time steps was performed at 1000 K, (4) 2 ps rMD at 900 K, (5) 2 ps rMD at 800 K, (6) van der Waals and Coulombic interactions were increased to 33%, (7) 2 ps rMD at 700 K, (8) van der Waals and Coulombic interactions were increased to 67%, (9) 2 ps rMD at 600 K, (10) van der Waals and Coulombic interactions were increased to 100%, (11) 2 ps rMD at 500 K, (12) 2 ps rMD at 400 K, (13) 2 ps rMD at 300 K, and (14) convergence with up to 40 000 steps of conjugate gradient energy minimization.

Table 1: Thermodynamic Parameters of Duplex Formation for 5′(rGGCAAGCCU)$_2$[a]

	$-\Delta G^\circ_{37}$ (kcal/mol)	$-\Delta H^\circ$ (kcal/mol)	$-\Delta S^\circ$ (eu)	T_m[b] (°C)
$1/T_M$ vs log C_T parameters[c]	10.04 ± 0.11	77.2 ± 2.0	216.5 ± 6.0	55.6
curve fit parameters	9.98 ± 0.40	75.3 ± 6.2	210.5 ± 18.8	55.8

[a] Solutions are 1 M NaCl, 20 mM sodium cacodylate, 0.5 mM Na_2EDTA, pH 7.0. [b] Calculated for 10^{-4} M oligonucleotide concentration. [c] Plots of T_M^{-1} versus log C_T resulted in an r-value of 0.998.

Since the strands are symmetric, close inspection is required to distinguish interstrand from intrastrand adenosine-adenosine cross-peaks. First, modeling studies were done without including internucleotide adenosine-adenosine restraints. Interstrand and intrastrand proton−proton distances were measured for all structures that resulted in few violations. From these distances, each adenosine−adenosine cross-peak was determined to be either interstrand or intrastrand. Then, with these adenosine−adenosine restraints added, models of the 5′(rGGCAAGCCU)$_2$ duplex consistent with NMR data were derived with the above protocol. Subsequently, to check the validity of the adenosine restraints, modeling studies with the interstrand restraints changed to intrastrand restraints and with the interstrand restraints removed were performed. Most of these structures had violations in the loop region that were greater than 0.1 Å, adding support to the assignment of the interstrand cross-peaks.

RESULTS

Thermodynamic Parameters. Thermodynamic parameters for the 5′(rGGCAAGCCU)$_2$ duplex are listed in Table 1. The parameters from the fits of melting curves and from T_M^{-1} versus log C_T plots agree within 3%, suggesting that the two-state model is a reasonable approximation for this transition. Peritz et al. (1991) observed an unusual transition at low temperature for concentrations of (CGCAAGCG)$_2$ above 1×10^{-4} M (*44*). This behavior is not observed for 5′(rGGCAAGCCU)$_2$.

Assignments of the Exchangeable Protons. The imino (10−15 ppm) and amino (6−9 ppm) regions of the ^{1}H NMR spectra recorded in H_2O provide information concerning hydrogen bonding between base pairs. As expected for a self-complementary duplex with six Watson−Crick G−C pairs (*25*), three G imino resonances are observed between 12 and 14 ppm (Supporting Information). The G1, G2, and G6 imino resonances were assigned from 1D NOE difference spectra, a 2D NOESY WATERGATE spectrum, and comparison to assignments for similar sequences (*21*, *45*). The imino

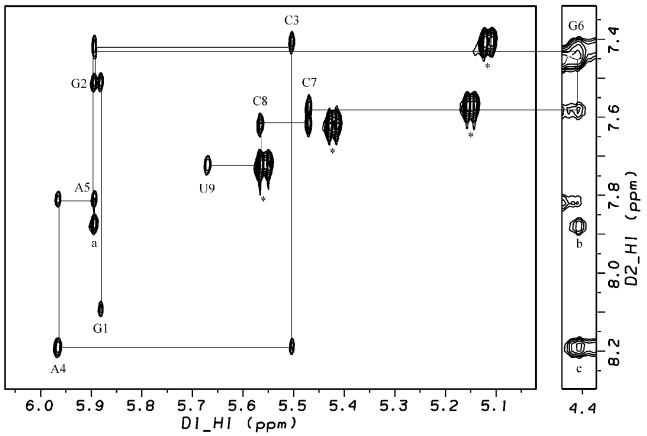

FIGURE 2: (H8/H6/H2)−(H1′/H5) region of the 400 ms NOESY spectrum of the 5′(rGGCAAGCCU)$_2$ duplex at 30 °C. Sequential assignments of base to H1′ protons are connected by lines. The assignments follow standard connectivity pathways. Intranucleotide H1′−H6/H8 cross-peaks are labeled. The four pyrimidine H5−H6 peaks are labeled "*". The A5H1′−A5H2 cross-strand peak is labeled "a". The G6H1′−A5H2 cross-peak is labeled "b" and was instrumental in assigning the G6H1′ resonance. The C3H3′−A4H8 cross-peak is labeled "c".

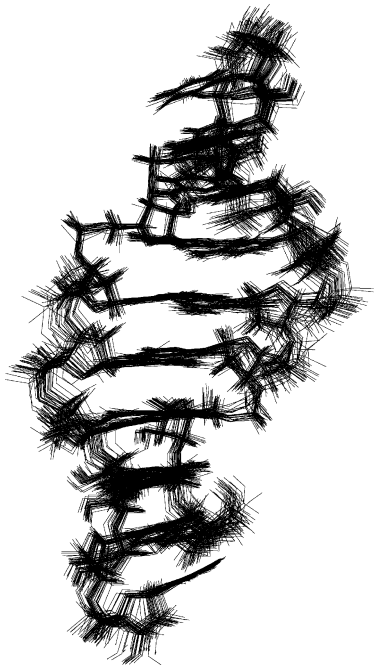

FIGURE 3: Superposition of 32 low-energy structures derived from restrained molecular dynamics.

protons displayed typical NOEs to adjacent imino protons and cross-strand amino protons (Supporting Information). The imino resonance of the 3′ dangling U residue was not observed, as expected due to rapid exchange with solvent. This is consistent with a previous study of (GCGAGCU)$_2$, where a resonance potentially due to the 3′ dangling U was only observed in the presence of 10 mM $MgCl_2$ (*46*).

Assignment of Nonexchangeable Proton and Phosphorus Resonances. NMR resonances of the 5′(rGGCAAGCCU)$_2$ duplex were assigned essentially as described by Varani et al. (*47*, *48*). A table in Supporting Information summarizes chemical shift assignments at 30 °C. The H5′ and H5″ assignments are not stereospecific, although the H5′ protons are expected to resonate downfield from the H5″ protons due to the negative charge on the phosphate (*49*). The (H8/H6/H2)−(H1′/H5) region of the 400 ms NOESY spectrum at 30 °C is shown in Figure 2. The assignments follow standard connectivity pathways from residues G1 to U9. Notice that the G6H1′ resonance is shifted upfield to ∼4.4 ppm. An upfield shift of H1′ resonances, due to the location of H1′ above or below bases with strong ring currents, has been reported previously (*21*, *50*−*52*).

Assignments of H2′ resonances follow from strong cross-peaks to H1′ resonances in the NOESY spectrum (Supporting Information) and from weak cross-peaks in the DQF−COSY spectrum. The H2′ assignments are confirmed by strong H8/H6(n) to H2′(n−1) cross-peaks in the base to sugar region of the NOESY spectra (Supporting Information), as typically observed for A-form conformations. TOCSY and ^{13}C HMQC spectra were also used to confirm proton assignments; e.g., the ^{13}C HMQC spectrum confirmed A4H2 and A4H8 at 8.19 ppm.

Phosphorus assignments were made directly from the strong H3′−P correlations in the ^{1}H-^{31}P HETCOR spectrum (Supporting Information). All eight ^{31}P resonances were observed, and all are in a 1 ppm range (−0.65 to −1.41 ppm), consistent with A-form geometry (*53*). The A4H3′−A5P cross-peak is weak, suggesting structural flexibility in the loop region.

Structure Determination. Distance and dihedral angle restraints for structure determination are listed in Supporting Information. Structures derived in the absence of solvent as well as structures derived with water as the solvent both displayed similar structural features and overall geometry. Since there were no apparent differences, further discussion is based on 50 structures derived in the absence of solvent. Thirty-two of the 50 structures generated from helical starting structures by the rMD protocol described in Materials and Methods have simulated total energies of −416 ± 2 kcal/mol. Of these 32 structures, at least three structures were derived from each of the five helical starting structures. The superposition reveals that the overall structure and local features (Figure 3) are well determined. The average RMSD for the all-atom pairwise superposition of these 32 structures is 0.42 Å. These structures converged to satisfy all NMR derived distance restraints within 0.10 Å. More specifically, in the minimized average structure, there were no violations greater than 0.10 Å. There were four violations less than 0.10 Å, with an average value of 0.06 Å. None of these violations were in the loop region. These structures contain sheared A_{anti}·A_{anti} hydrogen bonded pairs similar to the A·A pairs described by Cate et al. (*6*) and Strobel et al. (*7*) and to the G·A pairs described by SantaLucia and Turner (*21*) (Figure 4). Figure 5 shows a stereoview of the A·A pairs and closing C−G base pairs from the minimized average structure. Molecular dynamic simulations with random structures as the starting structures also contain sheared A_{anti}·A_{anti} hydrogen bonded pairs with no violations in the loop region. Due to greater convergence with A-form and B-form type starting structures, further discussion will focus on the structures derived from these starting structures.

The T_1 relaxation times are consistent with the final structure (Supporting Information). The measured relaxation time of 5.9 s for the A5H2 proton is significantly longer than that for other base protons. H2 protons typically relax slowly because they are not near any other protons. The

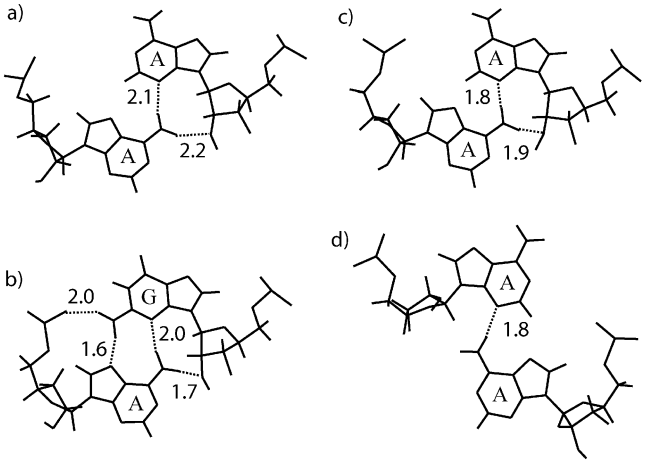

FIGURE 4: Structures and hydrogen bonding patterns of (a) the A·A pair from the lowest-energy structure of 5′(rGGCAAGCCU)$_2$, (b) the G·A pair from the NMR-derived structures of 5′(rGGCGAGCC)$_2$ (*21*), and (c and d) A·A pairs from the tandem A·A mismatch from the crystal structure of the J4/5 loop of the *T. thermophila* group I intron (*6*). Dashed lines indicate functional groups that are close enough to form standard hydrogen bonds. The numbers indicate the approximate distance in angstroms between the hydrogen and the hydrogen bond acceptor. In panel a, there is also a hydrogen bond from the A4 2′ OH to A5O5′ with an average distance of 2.3 Å. In some of the NMR-derived structures of 5′(rGGCAAGCCU)$_2$, the A4 2′ OH hydrogen is close enough to form a bifurcated hydrogen bond with the A5O4′ and A5O5′ atoms of the following residue, with an average distance of 2.7 Å to the A5O4′ atom. Also, a C−H hydrogen bond may be possible between A4H2 and A5N7, with an average distance of 2.6 Å.

A4H2 resonance is overlapped with the A4H8 resonance, so a reliable value for the A4H2 relaxation time was not obtained. Other base protons and sugar H1′ protons had typical relaxation times (*54*, *55*). A theoretical NOESY spectrum calculated with a multiple spin model using the minimized average structure was in agreement with the experimental spectrum.

DISCUSSION

Thermodynamics. The melting temperature of the 5′-(rGGCAAGCCU)$_2$ duplex is dependent on oligonucleotide concentration, showing that a duplex, not a hairpin, is formed. Thermodynamic increments for the internal loop can be calculated from the equation (*56*)

$$\Delta G^\circ_{37,\text{loop}} = \Delta G^\circ_{37}(\text{rGGCAAGCCU})_2 - \Delta G^\circ_{37}(\text{rGGCGCC})_2 + \Delta G^\circ_{37}(\text{CG})_2 - 2\Delta G^\circ_{37}(\text{C}U) \quad (2)$$

where ΔG°_{37}(rGGCAAGCCU)$_2$ is the measured value of the duplex containing the internal loop, ΔG°_{37}(rGGCGCC)$_2$ is the measured value of the duplex without the loop and without the dangling end (*57*), ΔG°_{37}(CG)$_2$ is the free energy increment for the nearest neighbor base pair interaction interrupted by the internal loop (*58*), and ΔG°_{37}(C*U*) is the average free energy increment of −1.1 kcal/mol measured for a dangling U on a C−G base pair (*59*, *60*). Similar calculations were done for $\Delta H^\circ_{\text{loop}}$ and $\Delta S^\circ_{\text{loop}}$, and the thermodynamic parameters of loop formation are listed in Table 2. Values for the internal loops of 5′(rCGCAAGCG)$_2$, 5′(rGGCGAGCC)$_2$, 5′(rCGCAGGCG)$_2$, and 5′(rGGCAGGCC)$_2$ duplexes are shown for comparison. The $\Delta G^\circ_{37,\text{loop}}$ for the 5′(rGGCAAGCCU)$_2$ duplex is 1.1 kcal/mol, similar to the value of 1.3 kcal/mol for the loop in the 5′-(rCGCAAGCG)$_2$ duplex (*35*, *44*). This loop is 2 kcal/mol destabilizing in comparison to the values of GA 2 × 2 internal loops with C−G closing base pairs: −0.7 kcal/mol for both the 5′(rCGAG)$_2$ and the 5′(rCAGG)$_2$ motifs (*34*, *35*, *46*). Although the 5′(rGGCAAGCCU)$_2$ duplex forms a similar sheared loop structure as the 5′(rGGCGAGCC)$_2$ duplex (*21*), the C6 location of the A4 amino group is oriented less well for hydrogen bonding than the C2 location of the G4 amino group (Figure 4a,b). The 5′(rGGCAGGCC)$_2$ duplex forms an imino-hydrogen bonded structure, which also uses the exocyclic amino group of guanosine in a way not available to adenosine (*45*). Other factors such as electrostatic interactions, chelation effects, and backbone distortion may also contribute differently to the stabilities of these three internal loops (*45*). Although it is thermodynamically destabilizing by 1.1 kcal/mol, the (rCAAG)$_2$ loop adopts a predominant conformation stabilized by A4·A5 hydrogen bonds and cross-strand A5-A5 stacking.

Structural Analysis. Exchangeable ^{1}H NMR NOE spectra show that Watson−Crick base pairs are formed between the G1−C8, G2−C7, and C3−G6 bases. The base to H1′ region of the NOESY spectrum contains no evidence of a syn base, indicating that all bases are in the anti conformation. Weak or nonexistent cross-peaks in the DQF−COSY spectrum indicate H1′−H2′ couplings of <2 Hz, which suggests predominantly C3′-endo sugar puckers for all nonterminal Watson−Crick residues. Both A4 and U9 residues have observable H1′−H2′ COSY cross-peaks (Supporting Information), suggesting dynamics between C3′-endo and C2′-endo sugar puckers. The C3′-endo sugar puckers for the nonterminal Watson−Crick residues are confirmed by large H3′−H4′ couplings. Strong cross-peaks are observed in the 100 ms mixing time NOESY spectrum from base H8/H6 protons to 5′ neighboring H2′ protons, consistent with typical A-form structure. All resonances in the 1D ^{31}P spectrum are within a 1 ppm range, suggesting that the A·A pairs do not significantly distort the conformation of the backbone from A-form. Four C8−U9 intrastrand restraints position the dangling U9 to stack on C8. The amino group of C8 closely overlaps the carbonyl-4 of U9, which has been suggested as a favorable electrostatic interaction in similar 3D structures (*60*).

Several unusual NOEs were crucial for structural determination of the sheared pairs. The medium strength A5H2-A5H1′ cross-peak (Figure 2) led to an NMR-derived distance of 3.5 Å. Since the AH2 proton cannot be closer than 4.5 Å to its own H1′ proton (*61*), this is an interstrand cross-peak. This cross-peak suggests that the minor groove is narrower than typically observed for A-form geometry (where a 4.7 Å distance is found for equivalent protons) and, therefore, that the minor groove does not widen to relieve steric clashes between two face-to-face adenosines. The A4H2−A5H8 interstrand cross-peak led to an NMR-derived distance of 4.4 Å. This proximity also indicates that the mismatches are not in a face-to-face geometry, which would result in an interstrand A4H2−A5H8 distance of at least 7 Å.

In the 32 lowest energy structures, the functional groups of opposing adenines are close enough to form hydrogen bonded A·A pairs similar to those observed in the crystal structure of the *T. thermophila* group I intron (*6*, *7*). The

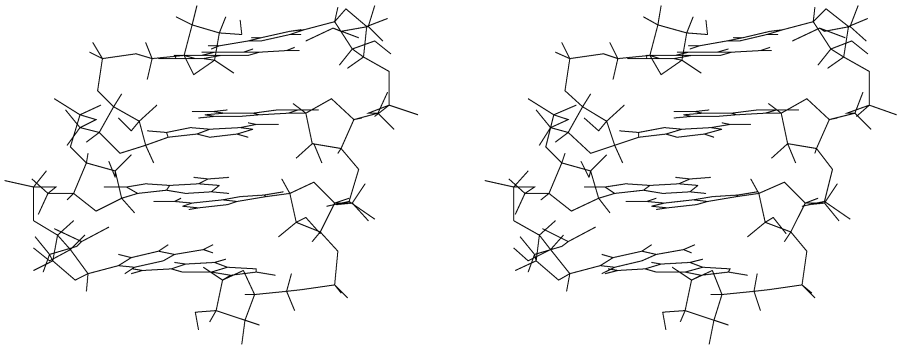

FIGURE 5: Stereoview of the minor groove of the lowest energy structure. Shown are the A·A pairs and the closing C−G base pairs.

Table 2: Thermodynamic Parameters of Loop Formation[a]

	$\Delta G°_{37}$ (kcal/mol)	$\Delta H°$ (kcal/mol)	$\Delta S°$ (eu)
5′(rGGCAAGCCU)$_2$	1.13	−5.0	−20.4
5′(rCGCAAGCG)$_2$[b]	1.31	−4.3	−17.8
5′(rGGCGAGCC)$_2$[c]	−0.70	−8.8	−26.3
5′(rCGCAGGCG)$_2$[d]	−1.00	−16.6	−50.1
5′(rGGCAGGCC)$_2$[e]	−0.43	−12.6	−39.3

[a] Calculated from T_M^{-1} vs log C_T parameters. Values for duplexes other than 5′(rGGCAAGCCU)$_2$ were recalculated from the data in the cited references using updated nearest neighbor parameters from Xia et al. (*58*). [b] Ref *44*. [c] Ref *34*. [d] Ref *46*. [e] Ref *88*.

average hydrogen bond distances in the A·A pairs of the NMR structures of the present work are 2.1 ± 0.3, 2.2 ± 0.3, and 2.3 ± 0.3 Å for the A4N3−A5NH6b, A4O2′−A5NH6a, and A4 2′ OH−A5O5′ atoms, respectively (Figure 4a). In some of the structures, the A4 2′ OH proton is close enough to form a bifurcated hydrogen bond with the A5O4′ and A5O5′ oxygens, with an average distance of 2.7 ± 0.6 Å to the A5O4′ oxygen. Because we did not include hydrogen bonding between opposing adenines, the observation of hydrogen-bonded A·A pairs in the lowest energy structures adds confidence to the proposed structure as the predominant conformer. In addition, there is the possibility of small favorable electrostatic interactions from weak hydrogen bonds (*2*, *62*) involving C−H groups in the loop: A5N7 is ∼2.6 Å from A4H2.

As a result of the sheared geometry of the A·A pairs, there is unusual stacking between adjacent A·A pairs and between A·A pairs and adjacent C−G pairs (Figure 6). There is significant stacking of one A5 base on the A5 base of the other strand, while the A4 bases overlap slightly. Also, there is stacking of the A·A pair on the closing C−G pair. The unusual stacking geometry places the A5 six-membered ring below the G6H1′ proton, which is consistent with an unusual chemical shift (4.45 ppm) of this proton (*63*).

The stacking information can be used in conjunction with the other results to form hypotheses about the dynamics of the loop. Since there is extensive cross-strand stacking of A5 and two hydrogen bonding interactions fixing this base, it is proposed that the A5 residue is in a relatively stable conformation. The well-defined NMR resonances of this region support this hypothesis. On the other hand, the A4 base stacks poorly and has only one apparent hydrogen bond to the base. Therefore, there are few favorable interactions that would be disrupted by dynamic motion of the A4 base. The NMR data also suggest that the A4 residue may be dynamic; the A4H3′−A5P cross-peak in the HETCOR spectra is weak (Supporting Information). Also, an A4H1′−A4H2′ cross-peak is observed in both the COSY and TOCSY spectra (Supporting Information), suggesting dynamics between C2′-endo and C3′-endo sugar puckers. Other internal nucleotides lack this cross-peak. Since there are potential hydrogen bonds between the A4 and A5 bases, it is probable that the A4 residue spends a significant proportion of time within hydrogen-bonding range of the A5 residue. Additionally, the direct hydrogen bonding between the A4 base and the A5 base could be replaced by a water-mediated hydrogen bond with minor structural changes, as seen in GNRA tetraloops (*64*).

Comparison to Other Structures. The structural features of the internal loop in the 5′(rGGCAAGCCU)$_2$ duplex can be compared to those observed in other RNA internal loops. One feature of the 5′(rGGCAAGCCU)$_2$ duplex is the slight backbone narrowing at the internal loop. The typical C1′−C1′ distance in standard A-form geometry is ∼10.9 Å. The corresponding distance (A4C1′ to A5C1′, cross-strand) at the location of the A·A pair is ∼9.4 Å. Similarly, the G·A pair in the 5′(rGGCGAGCC)$_2$ duplex has a slightly drawn in backbone at the loop, with a C1′−C1′ distance of 9.1 Å (*21*). Another atypical feature of the 5′(rGGCAAGCCU)$_2$ duplex is the cross-strand stacking of the A5 base on the A5 base. Typically, bases stack with the sequential base in the same strand. In the loop region of this structure, however, the A5 base exhibits interstrand stacking with the A5 base from the opposite strand. The cross-strand medium intensity NOE from the A5H1′ proton to the A5H2 proton is consistent with this type of stacking. Others (*8*, *21*, *27*, *65*−*67*) have also observed cross-strand stacking of adenosines in internal loops.

Several internal loops with large percentages of adenosines have been determined to adopt adenosine platform conformations (*68*−*71*). The NOE data, however, suggest that the adenosines of 5′(rGGCAAGCCU)$_2$ form a sheared conformation rather than an adenosine platform. In general, Amarasinghe et al. (*69*) observed NOE cross-peak patterns for the adenosines in an A-platform that are not consistent with standard A-form geometry. On the other hand, the adenosines of 5′(rGGCAAGCCU)$_2$ exhibit NOE cross-peak patterns that are consistent with standard A-form geometry. More specifically, several cross-peaks would be expected for adenosines that are involved in an A-platform geometry that are not observed in the NMR spectra of 5′(rGGCAAGCCU)$_2$. For example, due to the coplanarity of the adenosines in A-platform structures, cross-peaks from non-nearest neighbors and cross-peaks from

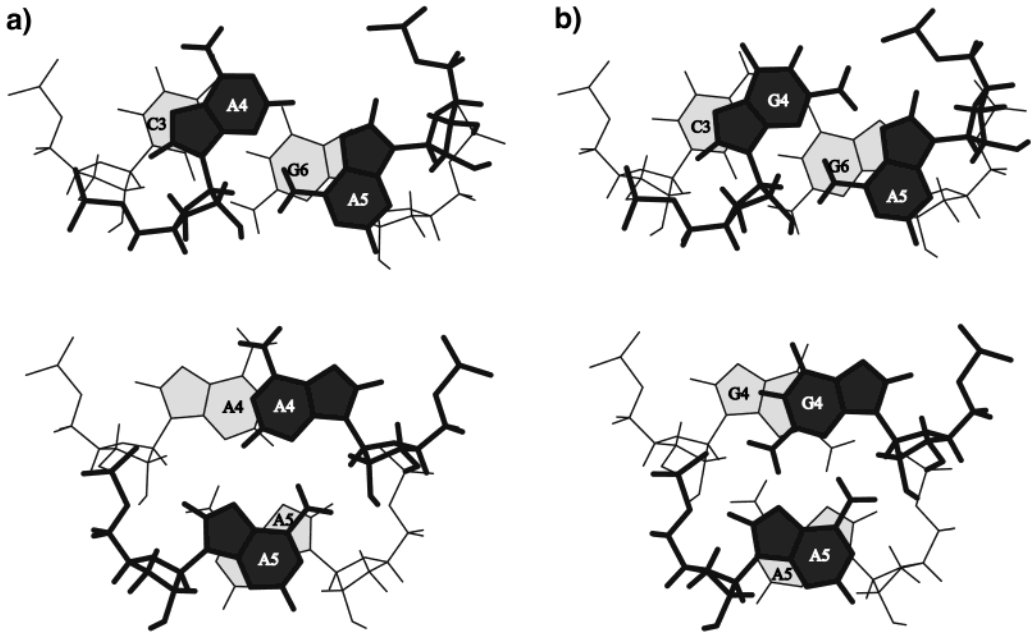

FIGURE 6: Base stacking between (a) adjacent A·A pairs (bottom) and the A·A pair and adjacent C−G pair (top) of 5′(rGGCAAGCCU)$_2$ and (b) adjacent G·A pairs (bottom) and the G·A pair and adjacent C−G pair (top) of 5′(rGGCGAGCC)$_2$ (*21*) shown for comparison. The darker residues are closer to the viewer. Note the extensive stacking of the A5 base on the cross-strand A5 base. In the top figure, the G6H1′ is situated below the ring of the A5 base. Figures were created with 3DNA (*89*).

H4′-(n+1)H8 are expected for the adenosines in the A-platform in NOESY spectra; neither was observed for 5′(rGGCAAGCCU)$_2$. Also, cross-peaks observed in the NOESY spectra of 5′(rGGCAAGCCU)$_2$ are inconsistent with an A-platform conformation. For example, the cross-strand NOE cross-peak from A5H1′-A5H2 and intrastrand A4H2′-A5H8 cross-peak were observed for 5′(rGGCAAGCCU)$_2$ but are not expected for an A-platform conformation. Last, an A-platform conformation results in a kink of the phosphate backbone which may result in a downfield shifted phosphorus resonance, which was not observed for 5′(rGGCAAGCCU)$_2$. The fact that 5′(rGGCAAGCCU)$_2$ does not form an A-platform structure is not surprising on the basis of the observations of Cate et al. (*68*). They observed that A-platform structures usually occur adjacent to non-Watson−Crick pairs. Also, each adenosine platform is involved in tertiary contacts and undergoes a conformational change in the absence of those tertiary interactions (*68*).

Hydrogen bonding between an adenosine N3 and an adenosine amino group has been observed previously, for example, in the group I intron (*6*, *72*), 30S ribosomal subunit (*73*), 50S ribosomal subunit (*33*), yeast tRNA (fmet) (*4*), tRNA aptamer (*74*), and loop B of the hairpin ribozyme (*8*). The hydrogen bonding pattern of the A·A pairs in the tandem mismatch of the 5′(rGGCAAGCCU)$_2$ duplex is strikingly similar in 5′(rCGAG)$_2$ and $\begin{pmatrix} G & A & A & A & G \\ C & A & A & & U \end{pmatrix}$ loops (Figure 4). SantaLucia and Turner (*21*) observed sheared G·A pairs with similar hydrogen bonding of the A amino group in 5′-(rGGCGAGCC)$_2$. Cate et al. (*6*) observed an identical A·A hydrogen bonding pattern in the crystal structure of the P4/P6 domain of the *T. thermophila* group I intron, which has an all adenosine 2 × 3 internal loop in which one adenosine is stacked on a tandem A·A pair. As shown in Figure 4, the only significant difference between the G·A and A·A configurations is that A·A lacks the hydrogen bonds formed by the exocyclic amino group of guanosine. Presumably, this contributes to the 2 kcal/mol difference in stability of internal loops with two sheared G·A or A·A pairs (Table 2). It is interesting to note that the tandem purine·purine pairs are arranged in the same position regardless of whether the loop is (rCGAG)$_2$, (rCAAG)$_2$, or $\begin{pmatrix} G & A & A & A & G \\ C & A & A & & U \end{pmatrix}$. This common structure and the common occurrence of adjacent purine·purine pairs suggest that the sheared geometry is biologically important and determined by local interactions.

Implications of the RNA Structure. Sheared purine·purine pairs are common in biology (*2*, *29*, *30*, *75*). More specifically, tandem A·A pairs are a common feature of the J4/5 loop of group I introns (*76*−*78*), suggesting a functional role. For example, the A113 and A206 residues of the *T. thermophila* group I intron, which are comparable to the A4 residue in the 5′(rGGCAAGCCU)$_2$ duplex, are adenosines in 71% and 83%, respectively, of all group I introns. Similarly, the A114 and A207 residues of the *T. thermophila* group I intron, which are comparable to the A5 residue in the 5′(rGGCAAGCCU)$_2$ duplex, are adenosines in 99% and 100%, respectively, of all group I introns (*76*).

One possible reason for conservation of these nucleotides is the availability of functional groups of the adenosine residues for hydrogen bonding in the major and minor grooves (*21*). In the 5′(rGGCAAGCCU)$_2$ duplex, the N7 and NH6 atoms of the A4 residue extend into the major groove (Figure 7). With the N7 and NH6 atoms available for hydrogen bonding, the A4 residue may form various Hoogsteen or A·purine base pairs, resulting in a base triple. In addition, these functional groups may also form tertiary contacts with amino acids, e.g., asparagine, glutamine, serine, threonine, or arginine. Similarly, the 2′ OH and N3 groups of A5 extend into the minor groove (Figure 7). These groups are also available to form tertiary or binding contacts with proteins or with other nucleotides, such as those reported

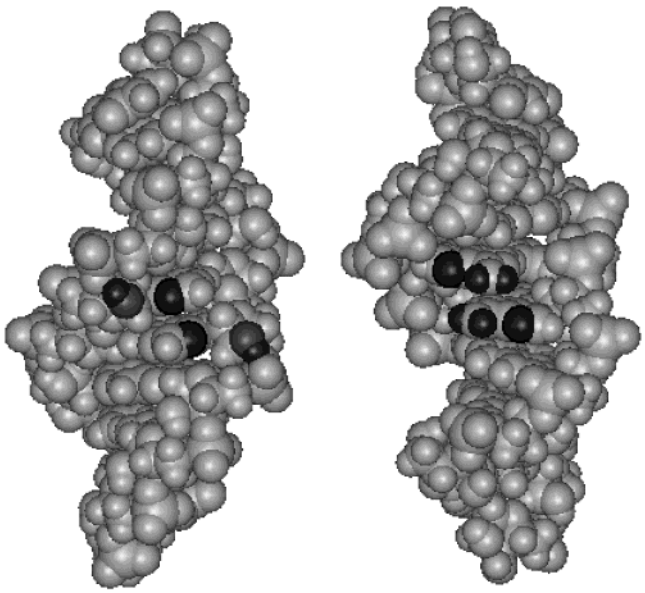

FIGURE 7: Minor (left) and major (right) grooves of 5′-(rGGCAAGCCU)$_2$. Highlighted functional groups on the left are identical to functional groups that form tertiary interactions in *T. thermophila*. Those highlighted on the right are available to form Hoogsteen pairs, A·purine pairs, or tertiary contacts to proteins.

for *T. thermophila*. In *T. thermophila*, it has been shown that the 2′ OH group of the A207 residue forms tertiary contacts with the O2′ group of the U(−1) residue of the P1 helix (*7*, *78*). Also, the A207 O2′, A207 N3, A114 2′ OH, and A114 N3 atoms have tertiary contacts with the NH2, NH2, O2′, and 2′ OH groups of the G22 residue of the P1 helix, respectively (*7*, *78*). As a result, tandem A·A pairs that contain hydrogen bonds and exhibit cross-strand stacking may be prevalent in nature and conserved in group I introns because functional groups in both the major and minor grooves permit formation of base pair triples with nucleic acids or tertiary interactions with proteins.

The above hypothesis is consistent with observations made by Gautheret et al. (*79*) concerning variations on the $\binom{\text{GA}}{\text{AG}}$ motif and their functional implications. They suggest that in sheared base pairs, groups usually involved in Watson−Crick base pairing are exposed to the solvent in both the major and minor grooves. They also report that sequence variations of 2 × 2 internal loops containing G·A oppositions prefer to have an adenosine residue at the 3′ positions of the internal loops, suggesting that the A5 cross-strand stacking and arrangement of the functional groups in the minor groove may be a conserved functional motif.

Many RNAs are being considered as targets for therapeutics (*80*−*85*). Designing such therapeutics requires an understanding of the interactions that are important for determining RNA structure, stability, and tertiary contacts. This study has shown that an isolated 2 × 2 loop in solution and a 3 × 2 loop in the crystal structure of a large RNA (*6*, *7*) form an almost identical conformation of a tandem A·A pair. Evidently, local hydrogen bonding and stacking interactions shape this structure. Thus, it may be possible to predict this shape on the basis of local secondary structure. The tandem sheared A_{anti}·A_{anti} pairs (A·A trans Hoogsteen/Sugar-edge) may be an attractive motif to target with therapeutics. The J4/5 loop in group I introns is highly conserved due to long-range tertiary interactions between the loop and the P1 substrate helix for 5′ splice site selection (*7*, *78*, *86*). It is feasible that a therapeutic could selectively target functional groups in such loops and interrupt tertiary interactions, resulting in an inactive RNA. Both the (rCAAG)$_2$ loop studied here (Table 2) and the J4/5 loop from *T. thermophila* (*87*), $\binom{\text{G A A A G}}{\text{C A A \quad U}}$, are thermodynamically destabilizing. Nevertheless, both appear sufficiently structured to provide reasonable targets for rational drug design. The results suggest that other internal loops that are thermodynamically destabilizing may also be structured.

ACKNOWLEDGMENT

We thank T. W. Barnes, III, and Dr. M. D. Disney for helpful discussions. We also thank H. Huang and Dr. S. Kennedy for NMR suggestions.

SUPPORTING INFORMATION AVAILABLE

Four tables listing proton and phosphorus chemical shifts, torsion angle restraints, distance restraints, and T_1 relaxation times are available. Eight figures show a 1D imino spectrum, the imino−amino region of the NOESY WATERGATE spectrum, the H1′-sugar region of the NOESY spectrum, the base-sugar region of the NOESY spectrum, a HETCOR spectrum, the sugar−sugar region of the DQF−COSY spectrum, the H1′/H5 region of the TOCSY spectrum, and the H1′−H2′ region of the DQF−COSY spectrum. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Saenger, W. (1989) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
2. Leontis, N. B., and Westhof, E. (1998) *Q. Rev. Biophys. 31*, 399−455.
3. Burkard, M. E., Turner, D. H., and Tinoco, I., Jr. (1999) in *The RNA World*, 2nd ed. (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) pp 675−680, Cold Spring Harbor Laboratory Press, Plainview, NY.
4. Basavappa, R., and Sigler, P. B. (1991) *EMBO J. 10*, 3105−3111.
5. Gervais, V., Cognet, J. A., Le Bret, M., Sowers, L. C., and Fazakerley, G. V. (1995) *Eur. J. Biochem. 228*, 279−290.
6. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science 273*, 1678−1685.
7. Strobel, S. A., Ortoleva-Donnelly, L., Ryder, S. P., Cate, J. H., and Moncoeur, E. (1998) *Nat. Struct. Biol. 5*, 60−66.
8. Butcher, S. E., Allain, F. H., and Feigon, J. (1999) *Nat. Struct. Biol. 6*, 212−216.
9. Prive, G. G., Heinemann, U., Chandrasegaran, S., Kan, L. S., Kopka, M. L., and Dickerson, R. E. (1987) *Science 238*, 498−504.
10. Prive, G. G., Heinemann, U., Chandrasegaran, S., Kan, L. S., Kopka, M. L., and Dickerson, R. E. (1988) in *Structure and Expression* (Sarma, R. H., and Sarma, M. H., Eds.) Vol. 2, pp 27−47, Adenine Press, Schenectady, NY.
11. Leontis, N. B., Stombaugh, J., and Westhof, E. (2002) *Nucleic Acids Res. 30*, 3497−3531.
12. Brown, T., Hunter, W. N., Kneale, G., and Kennard, O. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2402−2406.
13. Hunter, W. N., Brown, T., and Kennard, O. (1986) *J. Biomol. Struct. Dynam. 4*, 173−191.
14. Webster, G. D., Sanderson, M. R., Skelly, J. V., Neidle, S., Swann, P. F., Li, B. F., and Tickle, I. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6693−6697.
15. Leonard, G. A., Booth, E. D., and Brown, T. (1990) *Nucleic Acids Res. 18*, 5617−5623.
16. Kan, L. S., Chandrasegaran, S., Pulford, S. M., and Miller, P. S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4263−4265.

17. Nikonowicz, E. P., and Gorenstein, D. G. (1990) *Biochemistry 29*, 8845−8858.
18. Wu, M., SantaLucia, J., Jr., and Turner, D. H. (1996) *Biochemistry 36*, 4449−4460.
19. Jiang, L. C., and Patel, D. J. (1998) *Nat. Struct. Biol. 5*, 769−774.
20. Heus, H. A., and Pardi, A. (1991) *Science 253*, 191−194.
21. SantaLucia, J., Jr., and Turner, D. H. (1993) *Biochemistry 32*, 12612−12623.
22. Cai, Z. P., Gorin, A., Frederick, R., Ye, X. M., Hu, W. D., Majumdar, A., Kettani, A., and Patel, D. J. (1998) *Nat. Struct. Biol. 5*, 203−212.
23. Ye, X. M., Gorin, A., Frederick, R., Hu, W. D., Majumdar, A., Xu, W. J., McLendon, G., Ellington, A., and Patel, D. J. (1999) *Chem. Biol. 6*, 657−669.
24. Li, Y., Zon, G., and Wilson, W. D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 26−30.
25. Carbonnaux, C., van der Marel, G. A., van Boom, J. H., Guschlbauer, W., and Fazakerley, G. V. (1991) *Biochemistry 30*, 5449−5458.
26. Bartel, D. P., Zapp, M. L., Green, M. R., and Szostak, J. W. (1991) *Cell 67*, 529−536.
27. Maskos, K., Gunn, B. M., LeBlanc, D. A., and Morden, K. M. (1993) *Biochemistry 32*, 3583−3595.
28. Fourmy, D., Yoshizawa, S., and Puglisi, J. D. (1998) *J. Mol. Biol. 277*, 333−345.
29. Nagaswamy, U., Voss, N., Zhang, Z. D., and Fox, G. E. (2000) *Nucleic Acids Res. 28*, 375−376.
30. Nagaswamy, U., Larios-Sanz, M., Hury, J., Collins, S., Zhang, Z. D., Zhao, Q., and Fox, G. E. (2002) *Nucleic Acids Res. 30*, 395−397.
31. Gutell, R. R., Cannone, J. J., Shang, Z., Du, Y., and Serra, M. J. (2000) *J. Mol. Biol. 304*, 335−354.
32. Cannone J. J., Subramanian S., Schnare M. N., Collett J. R., D'Souza L. M., Du Y., Feng B., Lin N., Madabusi L. V., Muller K. M., Pande N., Shang Z., Yu N., and Gutell R. R. (2002) *BioMed Central Bioinformatics 3*, 2.
33. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) *Science 289*, 905−920.
34. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1991) *Biochemistry 30*, 8242−8251.
35. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) *J. Mol. Biol. 288*, 911−940.
36. Usman, N., Ogilvie, K. K., Jiang, M. Y., and Cedergren, R. J. (1987) *J. Am. Chem. Soc. 109*, 7845−7854.
37. Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Tracz, D., Workman, C., Sweedler, D., Gonzalez, C., Scaringe, S., and Usman, N. (1995) *Nucleic Acids Res. 23*, 2677−2684.
38. Petersheim, M., and Turner, D. H. (1983) *Biochemistry 22*, 256−263.
39. McDowell, J. A., and Turner, D. H. (1996) *Biochemistry 35*, 14077−14089.
40. Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843−853.
41. Sklenar, V., Miyashiro, H., Zon, G., Miles, H. T., and Bax, A. (1986) *FEBS Lett. 208*, 94−98.
42. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc. 117*, 5179−5197.
43. Burkard, M. E., and Turner, D. H. (2000) *Biochemistry 39*, 11748−11762.
44. Peritz, A. E., Kierzek, R., Sugimoto, N., and Turner, D. H. (1991) *Biochemistry 30*, 6428−6436.
45. Wu, M., SantaLucia, J., Jr., and Turner, D. H. (1997) *Biochemistry 36*, 4449−4460.
46. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1990) *Biochemistry 29*, 8813−8819.
47. Varani, G., and Tinoco, I. (1991) *Q. Rev. Biophys. 24*, 479−532.
48. Varani, G., Aboulela, F., and Allain, F. H. T. (1996) *Prog. Nucl. Magn. Reson. Spectrosc. 29*, 51−127.
49. Varani, G., Cheong, C., and Tinoco, I., Jr. (1991) *Biochemistry 30*, 3280−3289.
50. Chou, S. H., Zhu, L. M., and Reid, B. R. (1994) *J. Mol. Biol. 244*, 259−268.
51. Fountain, M. A., Serra, M. J., Krugh, T. R., and Turner, D. H. (1996) *Biochemistry 35*, 6539−6548.
52. Chou, S. H., and Chin, K. H. (2001) *J. Mol. Biol. 312*, 753−768.
53. Gorenstein, D. (1984) *^{31}P NMR, Principles and Applications*, Academic Press, New York.
54. Petersheim, M., and Turner, D. H. (1983) *Biochemistry 22*, 264−268.
55. Wang, A. C., Kim, S. G., Flynn, P. F., Chou, S. H., Orban, J., and Reid, B. R. (1992) *Biochemistry 31*, 3940−3946.
56. Gralla, J., and Crothers, D. M. (1973) *J. Mol. Biol. 78*, 301−319.
57. Freier, S. M., Sugimoto, N., Sinclair, A., Alkema, D., Neilson, T., Kierzek, R., Caruthers, M. H., and Turner, D. H. (1986) *Biochemistry 25*, 3214−3219.
58. Xia, T., SantaLucia, J., Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) *Biochemistry 37*, 14719−14735.
59. Turner, D. H., Sugimoto, N., and Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.
60. Burkard, M. E., Kierzek, R., and Turner, D. H. (1999) *J. Mol. Biol. 290*, 967−982.
61. Wutrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York.
62. Desiraju, G. R., and Steiner, T. (1999) *The Weak Hydrogen Bond in Structural Chemistry and Biology*, Oxford University Press, New York.
63. Giessner-Prettre, C., Pullman, B., Borer, P. N., Kan, L. S., and Ts'o, P. O. (1976) *Biopolymers 15*, 2277−2286.
64. Jucker, F. M., Heus, H. A., Yip, P. F., Moors, E. H., and Pardi, A. (1996) *J. Mol. Biol. 264*, 968−980.
65. Katahira, M., Kanagawa, M., Sato, H., Uesugi, S., Fujii, S., Kohno, T., and Maeda, T. (1994) *Nucleic Acids Res. 22*, 2752−2759.
66. Baeyens, K. J., De Bondt, H. L., Pardi, A., and Holbrook, S. R. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 12851−12855.
67. Dallas, A., and Moore, P. B. (1997) *Structure 5*, 1639−1653.
68. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science 273*, 1696−1699.
69. Amarasinghe, G. K., De Guzman, R. N., Turner, R. B., and Summers, M. F. (2000) *J. Mol. Biol. 299*, 145−156.
70. Conn, G. L., Draper, D. E., Lattman, E. E., and Gittis, A. G. (1999) *Science 284*, 1171−1174.
71. Wimberly, B. T., Guymon, R., McCutcheon, J. P., White, S. W., and Ramakrishnan, V. (1999) *Cell 97*, 491−502.
72. Juneau, K., Podell, E., Harrington, D. J., and Cech, T. R. (2001) *Structure 9*, 221−231.
73. Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Wimberly, B. T., and Ramakrishnan, V. (2000) *Nature 407*, 340−348.
74. Bullock, T. L., Sherlin, L. D., and Perona, J. J. (2000) *Nat. Struct. Biol. 7*, 497−504.
75. Chou, S. H., Zhu, L., and Reid, B. R. (1997) *J. Mol. Biol. 267*, 1055−1067.
76. Michel, F., and Westhof, E. (1990) *J. Mol. Biol. 216*, 585−610.
77. Ortoleva-Donnelly, L., Szewczak, A. A., Gutell, R. R., and Strobel, S. A. (1998) *RNA 4*, 498−519.
78. Strauss-Soukup, J. K., and Strobel, S. A. (2000) *J. Mol. Biol. 302*, 339−358.
79. Gautheret, D., Konings, D., and Gutell, R. R. (1994) *J. Mol. Biol. 242*, 1−8.
80. Mol, J. N. M., and van der Krol, A. R. (1991) *Antisense Nucleic Acids and Proteins, Fundamentals and Applications*, Marcel Dekker, Inc., New York.
81. Agrawal, S. (1995) *Methods in Molecular Medicine: Antisense Therapeutics*, Humana Press, Totowa, NJ.
82. Testa, S. M., Disney, M. D., Turner, D. H., and Kierzek, R. (1999) *Biochemistry 38*, 16655−16662.
83. Disney, M. D., Testa, S. M., and Turner, D. H. (2000) *Biochemistry 39*, 6991−7000.
84. Lukavsky, P. J., Otto, G. A., Lancaster, A. M., Sarnow, P., and Puglisi, J. D. (2000) *Nat. Struct. Biol. 7*, 1105−1110.
85. Gallego, J., and Varani, G. (2001) *Acc. Chem. Res. 34*, 836−843.
86. Lehnert, V., Jaeger, L., Michel, F., and Westhof, E. (1996) *Chem. Biol. 3*, 993−1009.
87. Schroeder, S. J., and Turner, D. H. (2001) *Biochemistry 40*, 11509−11517.
88. Xia, T. B., McDowell, J. A., and Turner, D. H. (1997) *Biochemistry 36*, 12486−12497.
89. Lu, X. J., Shakked, Z., and Olson, W. K. (2000) *J. Mol. Biol. 300*, 819−840.

BI020326F