Structural Features and Thermodynamics of the J4/5 Loop from the *Candida albicans* and *Candida dubliniensis* Group I Introns[†,‡]

Brent M. Znosko,[§,||] Scott D. Kennedy,[⊥] Pamela C. Wille,[§] Thomas R. Krugh,[§] and Douglas H. Turner*[,§,#]

Department of Chemistry, RC Box 270216, Department of Biochemistry and Biophysics, School of Medicine and Dentistry, 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 14, 2004; Revised Manuscript Received August 12, 2004

ABSTRACT: The J4/5 loop of group I introns has tertiary interactions with the P1 helix that position the P1 substrate for the self-splicing reaction. The J4/5 loop of *Candida albicans* and *Candida dubliniensis*, 5′GAAGG3′/3′UAAUU5′, potentially contains two A·A pairs flanked by one G·U pair on one side and two G·U pairs on the other side. Results from optical melting, nuclear magnetic resonance spectroscopy, and functional group substitution experiments with a mimic of the *C. albicans* and *C. dubliniensis* J4/5 loop are consistent with the adenosines forming tandem sheared A·A pairs with a cross-strand stack and only the G·U pair not adjacent to an A·A pair forming a static wobble G·U pair. The two G·U pairs adjacent to the tandem A·A pairs are likely in a dynamic equilibrium between multiple conformations. Although $Co(NH_3)_6^{3+}$ stabilizes the loop by several kilocalories per mole at 37 °C, addition of Mg^{2+} or $Co(NH_3)_6^{3+}$ has no effect on the structure of the loop. The tandem G·U pairs provide a pocket of negative charge for $Co(NH_3)_6^{3+}$ to bind. The results contribute to understanding the structure and dynamics of purine-rich internal loops and potential G·U pairs adjacent to internal loops.

Internal loops are often important for folding and function of RNA and therefore represent potential therapeutic targets (*1*−*5*). Little is known, however, about the sequence dependence of internal loop stabilities, structures, and dynamics (*6*). An understanding of the elements of molecular recognition that determine these properties would facilitate modeling of RNA structure and potentially aid the design of therapeutics targeting RNA.

A particularly interesting internal loop is the J4/5 loop in group I self-splicing introns. Group I self-splicing introns are found in many genes of both prokaryotic and eukaryotic organisms (*7*−*9*), and more than 1500 are known (*10*, *11*). The J4/5 loop of group I introns has tertiary interactions with the P1 helix (*9*, *12*) that position P1 for the self-splicing reaction. This loop has several phylogenetically conserved features (*9*, *10*, *13*). In an analysis of 71 group I introns, 100% contain purine-rich J4/5 internal loops (*14*), 97% contain adenines at J4/5 loop positions 114 and 207 (*Tetrahymena thermophila* numbering) (*14*), and 80% contain at least one potential G·U pair adjacent to the J4/5 loop (*10*).

Candida albicans and *Candida dubliniensis* are fungal pathogens that pose a serious health risk to immunocompromised individuals, including AIDS patients, cancer patients, diabetics, newborns, and the elderly. *C. albicans* is responsible for ∼8% of all hospital-acquired infections (*15*), and *C. dubliniensis* has been isolated from the oral cavity of ∼30% of patients with AIDS and oral candidiasis (*16*). About 40% of *C. albicans* and all *C. dubliniensis* strains contain a group I intron in their LSU[1] rRNA precursor (*17*). Because self-splicing is necessary for maturation of ribosomes (*18*), group I introns are a potential RNA drug target. Therefore, a better understanding of the thermodynamics and structure of *Candida* internal loops may aid the rational design of drugs.

The J4/5 loop in both *C. albicans* and *C. dubliniensis* is the same, 5′GAAGG3′/3′UAAUU5′ (Figure 1). This J4/5 loop potentially contains two A·A pairs flanked by one G·U pair on one side and two consecutive G·U pairs on the other side. To investigate the thermodynamics and structure of an isolated J4/5 loop, duplex mimics were designed that contain the J4/5 loop from *C. albicans* and *C. dubliniensis* (Figure 1). Insight into elements determining stability and structure was obtained by substituting purines for adenines in the loop, thereby replacing the 6-amino group with a hydrogen, and by substituting inosines for guanosines

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

[‡] Protein Data Bank entry 1TUT.

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

[§] Department of Chemistry.

[||] Current address: Department of Chemistry, St. Louis University, St. Louis, MO 63103-2010.

[⊥] Department of Biochemistry and Biophysics.

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

[1] Abbreviations: 4I, modified J4/5 mimic with guanosine to inosine substitution at position 4; 5/16P, modified J4/5 mimic with adenine to purine substitutions at positions 5 and 16; 6/17P, modified J4/5 mimic with adenine to purine substitutions at positions 6 and 17; 7I, modified J4/5 mimic with guanosine to inosine substitution at position 7; *Ca/Cd* mimic, the *Candida albicans* and *Candida dubliniensis* J4/5 mimic studied by optical melting and NMR; nt, nucleotide; s-*Ca/Cd* mimic, short, unstable *C. albicans* and *C. dubliniensis* J4/5 mimic studied by optical melting and NMR; LSU, large subunit; T_M, melting temperature in kelvin; T_m, melting temperature in degrees Celsius.

10.1021/bi049256y CCC: $27.50
Published on Web 11/23/2004

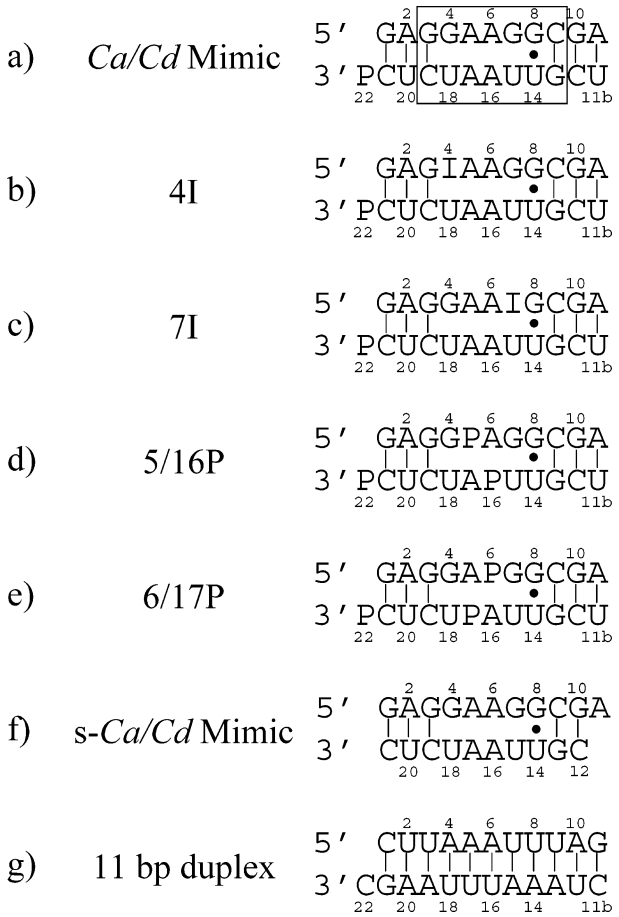

FIGURE 1: Schematic representation and numbering of the (top to bottom) (a) *Ca/Cd* J4/5 loop mimic studied by NMR, (b) 4I duplex, (c) 7I duplex, (d) 5/16P duplex, (e) 6/17P duplex, (f) s-*Ca/Cd* J4/5 loop mimic, and (g) 11 bp duplex. The box in (a) illustrates the sequence identical in the NMR mimic and the natural *Ca/Cd* sequence.

adjacent to the loop, thereby replacing the 2-amino group with a hydrogen (Figure 1).

Divalent metal cations can bind in the deep negative electrostatic potential of the major groove of G·U pairs (*19*−*22*). Crystal structures of the wild-type and mutant P4−P6 domain of the *T. thermophila* group I intron suggest that hexahydrated Mg^{2+} may aid the compaction of the J4/5 loop into a helical fold (*23*, *24*). This suggests that cobalt(III) hexaammine may also be stabilizing because it is similar in size to hydrated Mg^{2+}, prefers the major groove, and is octahedral in geometry (*25*). Due to the potential of adding structural stability to the J4/5 loop, optical melting and NMR data were collected in the presence of $MgCl_2$ and $[Co(NH_3)_6]Cl_3$. To determine if cationic effects were due to specific or nonspecific interactions, a duplex of equivalent length but different sequence, 5′CUUAAAUUUAG3′:3′CGAAUUUAAAUC5′ (referred to as the 11 bp duplex; see Figure 1), was also studied by optical melting.

The results from optical melting, NMR spectroscopy, and functional group substitution experiments are consistent with the adenosines forming tandem sheared A·A pairs with a cross-strand stack and only the G·U pair not adjacent to the tandem A·A pairs forming a static wobble G·U pair. The two G·U pairs adjacent to the tandem A·A pairs are likely in a dynamic equilibrium between multiple conformations. The tandem G·U pairs provide a pocket of negative charge for binding $Co(NH_3)_6^{3+}$.

MATERIALS AND METHODS

RNA Synthesis and Purification. Oligonucleotides were synthesized on CPG support with an Applied Biosystems 392 DNA/RNA synthesizer using the phosphoramidite method (*26*, *27*). Standard support, standard phosphoramidites, and inosine phosphoramidite were acquired from Glen Research (Baltimore, MD). Purine support and purine phosphoramidite were acquired from ChemGenes (Ashland, MD). Base-labile protecting groups were removed by treatment with a 3:1 (v/v) ammonia/ethanol solution at 55 °C overnight (*28*). Quik-Sep disposable filter columns from Perkin-Elmer (Norton, OH) were used to separate the oligonucleotides from support. Removal of the silyl protecting groups was achieved with anhydrous 1 M triethylamine/hydrogen fluoride in pyridine (50 equiv) at 55 °C for 48 h. Samples were extracted with diethyl ether to remove organic impurities and then lyophilized before redissolving in 5 mM ammonium bicarbonate at pH = 7.0. These solutions were loaded onto a Waters Sep-Pak C18 chromatography column to remove excess inorganic salts. Oligonucleotides were 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 (*29*). Main product bands were identified by UV shadowing and extracted from the silica with distilled water. The Sep-Pak procedure was repeated. Purity of the oligomers was checked by reverse-phase HPLC and was greater than 95%. Extinction coefficients of purine and inosine were assumed to be the same as adenine and guanine, respectively.

Optical Melting Experiments. Purified oligonucleotides were lyophilized and redissolved in melt buffer. Melt buffer was either 1 M NaCl, 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH = 7.0, or 80 mM NaCl, 3 mM KH_2PO_4, 7 mM K_2HPO_4, and 0.5 mM Na_2EDTA, pH = 6.1, in the presence of 0−50 mM $MgCl_2$ or 0−5 mM $[Co(NH_3)_6]Cl_3$. Curves of absorbance at 280 nm (and 260 nm for the 11 bp duplex) versus temperature for samples containing ~1−100 μM strand concentrations were acquired at a heating rate of 1 °C/min with a Beckman Coulter DU-640 spectrophotometer equipped with a temperature controller and cooling bath. Melting curves were fit to a two-state model, assuming linear sloping baselines and temperature-independent $\Delta H°$ and $\Delta S°$ (*30*, *31*). Additionally, melting temperatures in kelvin, T_M, at different concentrations were used to calculate thermodynamic parameters according to Borer et al. (*32*):

$$T_M^{-1} = 2.303R/\Delta H° \log(C_T/4) + \Delta S°/\Delta H° \quad (1)$$

For the 11 bp duplex, thermodynamic values measured at 260 and 280 nm were within experimental error, and reported values were obtained at 260 nm where the signal-to-noise ratio was higher.

Sample Preparation for NMR. Samples were lyophilized and redissolved in 200−250 μL of 80 mM NaCl, 3 mM KH_2PO_4, 7 mM K_2HPO_4, and 0.5 mM Na_2EDTA, pH = 6.1, and appropriate amounts of $MgCl_2$ or $[Co(NH_3)_6]Cl_3$. For nonexchangeable proton samples, D_2O exchange was performed with lyophilization from 99.96% D_2O. Samples were redissolved in 200−250 μL of 99.996% D_2O (Cambridge

Isotope Laboratories). Duplex concentrations for these samples were ~2 mM, giving total strand concentrations of ~4 mM.

NMR Spectroscopy. Spectra were acquired on Varian Inova 500 and 600 MHz spectrometers. One-dimensional imino proton spectra and two-dimensional 150 ms mixing time NOESY spectra of exchangeable protons for samples dissolved in 90:10 (v/v) H_2O/D_2O were recorded with an S-shaped pulse (*33*). NOESY spectra of samples in D_2O were acquired at 200, 400, and 600 ms mixing times. FIDs were acquired over 256 t_1 increments, each having 4k points with 64 scans per FID and a recycle delay of 2.8 s. One-dimensional phosphorus spectra were acquired at 25 °C with a spectral width of 2450 Hz, 4k data points, and 512 transients. TOCSY spectra were acquired at 25 °C with a 120 ms mixing time, 4k data points, and a spectral width of 4000 Hz in both dimensions.

Two-dimensional spectra were processed with the Felix 2000 software package (Molecular Simulations Inc.). Proton spectra were referenced to H_2O or HDO. Phosphorus resonances were referenced to the phosphate peak at 0.0 ppm at 25 °C (pH = 6.5).

Spectra were collected of the s-*Ca/Cd* and *Ca/Cd* mimics in the absence of $MgCl_2$ and $[Co(NH_3)_6]Cl_3$ and of the *Ca/Cd* mimic in the presence of 10−30 mM $MgCl_2$ or 1−3 mM $[Co(NH_3)_6]Cl_3$. Spectra of the 4I, 7I, 5/16P, and 6/17P duplexes were collected in 3 mM $[Co(NH_3)_6]Cl_3$.

Restraint Generation. Distance restraints were generated for the *Ca/Cd* mimic in 3 mM $Co(NH_3)_6^{3+}$ at 25 °C. NOESY cross-peak volumes were determined by measuring the volume of interest with the Felix software package. Distance restraints were generated from 200, 400, and 600 ms mixing time NOE volumes (excluding H5′ and H5″ because assignments are not stereospecific) with certain pyrimidine H5−H6 cross-peaks as references at 2.45 Å. Other proton−proton distances were calculated from the two-spin approximation by $1/r^6$ scaling. If a cross-peak was first observed in the 200, 400, or 600 ms spectrum, error limits of ±30%, 40%, or 50% of the calculated distances were assigned to allow for spin diffusion, phasing, baseline abnormalities, and noise. Larger error limits were used for longer mixing times to ensure that restraints were not too tight as a result of spin diffusion. Three restraints were used from the 600 ms spectrum (G7H1′−G7H8, 2.19−6.56 Å; A11H1′−A11H2, 2.88−8.07 Å; and C21H5−U20H6, 2.19−6.57 Å), resulting in restraints that were loose enough to include standard A-form distances. Specific restraints were loosened if extreme overlap was observed. Included for structural refinement were a total of 217 NMR-derived interproton distance restraints (108 intranucleotide and 109 internucleotide). There is a small population of minor conformations (<5−10%) as evidenced by exchange peaks near the diagonal for some resonances. No NOE cross-peaks were observed from these minor cross-peaks.

Hydrogen bond restraints (1.6−2.1 Å) were used for the six Watson−Crick pairs and for the G8·U14 pair. On the basis of NMR data that is consistent with the formation of tandem sheared A·A pairs (discussed below), two artificial hydrogen-bonding restraints (amino proton−2′OH, 1.5−2.8 Å, and amino proton−N3, 1.4−2.7 Å) were used within each A·A pair. These restraints were generated from ±20% of the average hydrogen bond distances from the NMR-derived structure of tandem sheared A·A pairs studied previously (*34*). No hydrogen-bonding restraints were used for the G4·U18 and G7·U15 pairs.

For modeling with $Co(NH_3)_6^{3+}$ present, the $Co(NH_3)_6^{3+}$ was placed in the major groove surface of the tandem G·U base pairs. Since all 18 protons of $Co(NH_3)_6^{3+}$ are equivalent, intermolecular RNA−$Co(NH_3)_6^{3+}$ NOE distance restraints were defined between the cobalt atom and RNA protons. Co−G8H1, Co−G13H1, and Co−U14H3 were given distance restraints of 1.8−7.5 Å to account for the RNA proton−hexaammine proton NOE and hexaammine proton−cobalt distance.

Sugar conformations were characterized by H1′−H2′ couplings evident in TOCSY spectra. Sugars for which no H1′−H2′ coupling is observed in the TOCSY spectrum but typical H1′ cross-peaks are observed in the NOESY spectra (G1−A6, G8−G10, C12−U15, and C19−C21) were restrained to the C3′-*endo* conformation with the torsion angles ν_0, ν_1, ν_2, ν_3, and ν_4. Other sugars were not restrained.

The glycosidic torsion angle was restrained to the *anti* conformation for each stem residue. It was also restrained to the *anti* conformation for the loop residues in which the intranucleotide NOE cross-peak between H1′ and aromatic resonances was similar in intensity to the intranucleotide NOE cross-peak between H1′ and aromatic resonances in the stem. This condition is not met for G7 and U18, and glycosidic torsion angles of these residues were not restrained. Also, G4 and U15 glycosidic torsion angles were left unrestrained to allow for flexibility. Parameters for the dihedral angles of the stem nucleotides were loosely restrained to those observed in canonical A-form RNA duplexes. Some backbone torsion angles of some of the loop nucleotides were also restrained to A-form values. A listing of all distance, dihedral angle, and endocyclic sugar torsion angle restraints is available in Supporting Information.

Simulated Annealing. Models of the *Ca/Cd* mimic consistent with NMR data were derived from restrained energy minimization and simulated annealing with the Discover 95 package on a Silicon Graphics computer. A standard A-form starting structure was generated with Biosym Insight II software. The P22 nucleotide was excluded from the calculation. Simulations in the absence of $Co(NH_3)_6^{3+}$ used the AMBER 95 force field (*35*) 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. For simulations in the presence of $Co(NH_3)_6^{3+}$, the ESFF force field was used (*36*). Simulations were performed without solvent and with the NMR restraints always on. For electrostatics, the cell-multipole method was used with a distance-dependent dielectric constant of $\epsilon = 2r$, where r is distance in angstroms. For van der Waals interactions, group-based summation was used with an 18 Å cutoff. The simulations involved a total of 14 steps (*34*, *37*): (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 of rMD was performed at 900 K, (5) 2 ps of rMD was performed at 800 K, (6) van der Waals and Coulombic interactions were increased to 33%, (7) 2 ps of rMD was performed at 700 K, (8) van der Waals and Coulombic

Table 1: Thermodynamic Parameters of Duplex Formation at 1 M NaCl

		T_M^{-1} vs log (C_T/4) plots				average of curve fits			
oligomers[a]	name	-ΔH° (kcal/mol)	-ΔS° (eu)	-ΔG°$_{37}$ (kcal/mol)	T_m[b] (°C)	-ΔH° (kcal/mol)	-ΔS° (eu)	-ΔG°$_{37}$ (kcal/mol)	T_m[b] (°C)
GAGGAAGGCGA PCUCUAAUUGCU	*Ca/Cd* mimic	86.8 ± 5.2	256.7 ± 16.7	7.21 ± 0.08	39.5	82.0 ± 6.0	241.1 ± 19.2	7.24 ± 0.14	39.7
GAGIAAGGCGA PCUCUAAUUGCU	4I	85.3 ± 7.7	252.7 ± 25.0	6.98 ± 0.16	38.6	93.3 ± 16.4	278.2 ± 52.5	6.98 ± 0.15	38.5
GAGGAAIGCGA PCUCUAAUUGCU	7I	79.4 ± 4.9	235.0 ± 16.0	6.48 ± 0.09	36.8	82.2 ± 6.7	244.0 ± 21.5	6.48 ± 0.14	36.8
GAGGPAGGCGA PCUCUAPUUGCU	5/16P	116.3 ± 8.8	353.2 ± 28.5	6.72 ± 0.09	37.5	108.0 ± 5.6	326.6 ± 18.6	6.75 ± 0.18	37.6
GAGGAPGGCGA PCUCUPAUUGCU	6/17P	96.7 ± 7.2	291.2 ± 23.4	6.41 ± 0.12	36.6	96.0 ± 16.2	288.9 ± 51.8	6.39 ± 0.19	36.6
GAGGAAGGCGA CUCUAAUUGC	s-*Ca/Cd* mimic	73.5 ± 2.7	222.0 ± 9.1	4.67 ± 0.10	29.4	69.1 ± 4.9	207.2 ± 16.3	4.84 ± 0.20	29.6
CUUAAAUUUAG CGAAUUUAAAUC	11 bp duplex	86.5 ± 5.4	250.0 ± 17.2	8.99 ± 0.13	46.1	84.1 ± 5.1	242.3 ± 16.5	8.97 ± 0.09	46.3
		(-79.0)[c]	(226.3)[c]	(8.88)[c]					

[a] Solutions are 1 M NaCl, 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH = 7.0. [b] Calculated for 10^{-4} M oligonucleotide concentration. [c] Values predicted for the 11 bp duplex by the nearest neighbor model (*39*, *40*).

interactions were increased to 67%, (9) 2 ps of rMD was performed at 600 K, (10) van der Waals and Coulombic interactions were increased to 100%, (11) 2 ps of rMD was performed at 500 K, (12) 2 ps of rMD was performed at 400 K, (13) 2 ps of rMD was performed at 300 K, and (14) convergence was performed with up to 40000 steps of conjugate gradient energy minimization. Steps 3−5 effectively randomized the starting structure. A total of 50 and 10 structures were generated in the absence and presence of $Co(NH_3)_6^{3+}$, respectively.

RESULTS

Thermodynamic Parameters. Thermodynamic parameters of duplex formation in 1 M NaCl for the duplexes shown in Figure 1 are listed in Table 1. Parameters from fits of melting curves and from T_M^{-1} versus log(C_T/4) plots agree within 15%, suggesting that the two-state model is a reasonable approximation for these transitions. The short *Ca/Cd* mimic, s-*Ca/Cd*, was not stable enough for further studies. Adding a 5′U and 3′ purine, P, to the shorter strand to give the *Ca/Cd* mimic, however, enhanced stability by 2.5 kcal/mol at 37 °C. On the basis of the expected enhancement due to base pairing of the 5′U (1.2 kcal/mol) (*38*), the 3′ dangling end purine stabilizes the duplex by 1.3 kcal/mol at 37 °C. When the amino group of G4 or G7 in the *Ca/Cd* mimic is replaced by hydrogen to give 4I and 7I, the free energy change of duplex formation at 37 °C becomes less favorable by 0.2 and 0.7 kcal/mol, respectively. When the amino groups of A5/A16 or A6/A17 are replaced by hydrogens to give 5/16P and 6/17P, the $\Delta G°_{37}$ becomes less favorable by 0.5 and 0.8 kcal/mol, respectively.

The free energy change for formation of the internal loop is calculated according to the equation:

$$\Delta G°_{37\ \mathrm{GAAG}/\mathrm{UAAU}} = \Delta G°_{37\ \mathrm{GAGGAAGGCGA}/\mathrm{PCUCUAAUUGCU}} - \Delta G°_{37,\mathrm{initiation}} - \Delta G°_{37\ \mathrm{G}/\mathrm{PC}} - \Delta G°_{37\ \mathrm{GAGG}/\mathrm{CUCU}} - \Delta G°_{37\ \mathrm{GGCGA}/\mathrm{UUGCU}} \quad (2)$$

Here, $\Delta G°_{37\ \mathrm{GAGGAAGGCGA}/\mathrm{PCUCUAAUUGCU}}$ is the $\Delta G°_{37}$ measured by optical melting, $\Delta G°_{37,\mathrm{initiation}}$ is 4.09 kcal/mol and accounts for initiation of the duplex (*39*), $\Delta G°_{37\ \mathrm{G}/\mathrm{PC}}$ is −1.3 kcal/mol and accounts for the dangling purine, and $\Delta G°_{37\ \mathrm{GAGG}/\mathrm{CUCU}}$ and $\Delta G°_{37\ \mathrm{GGCGA}/\mathrm{UUGCU}}$ are predicted from the nearest neighbor parameters for the stems adjacent to the loop where values of −0.5 or +0.5 kcal/mol were used for GG/UU (*39*, *40*). The calculated $\Delta G°_{37,\mathrm{loop}}$ for the *Ca/Cd* mimic is 3.6 or 2.6 kcal/mol. The values are similar to the 2.8 kcal/mol predicted from nearest neighbor parameters for tandem mismatches (*40*, *41*) and to the 3.6 kcal/mol measured for a $\frac{\mathrm{GAAG}}{\mathrm{UAAU}}$ loop in a different context (*42*).

Thermodynamic parameters for the *Ca/Cd* mimic and the 11 bp duplex in the presence of Mg^{2+} or $Co(NH_3)_6^{3+}$ added to a solution of 80 mM NaCl are listed in Table 2. The stability of the *Ca/Cd* mimic plateaus near 10 mM Mg^{2+} and 1 mM $Co(NH_3)_6^{3+}$. With addition of 10 mM Mg^{2+}, the $\Delta G°_{37}$ of duplex formation becomes more favorable by 3.0 and 2.2 kcal/mol for the *Ca/Cd* mimic and 11 bp duplex, respectively. With addition of 1 mM $Co(NH_3)_6^{3+}$, the $\Delta G°_{37}$ becomes more favorable by 7.7 and 3.5 kcal/mol for the *Ca/Cd* mimic and 11 bp duplex, respectively. For comparison, increasing [Na^+] from 80 mM to 1 M in the absence of any multivalent enhances the stabilities of the *Ca/Cd* mimic and 11 bp duplex by 3.8 and 3.1 kcal/mol, respectively (Tables 1 and 2). The results suggest that the *Ca/Cd* mimic has a binding site for $Co(NH_3)_6^{3+}$.

Another construct was created in which the sequence of the bottom strand of the *Ca/Cd* mimic was changed to 5′UCGCCUUCCUCP so that all base pairs within the duplex are Watson−Crick pairs. This duplex, however, did not melt in a two-state manner, and the melting temperature in the presence of Mg^{2+} or $Co(NH_3)_6^{3+}$ was at the limits of the spectrometer, thus prohibiting determination of the thermodynamics.

Effect of Magnesium and Cobalt Hexaammine on Chemical Shifts of the Ca/Cd Mimic. NMR spectra were measured

Table 2: Thermodynamic Parameters of Duplex Formation at 80 mM NaCl in the Presence and Absence of Mg^{2+} or $Co(HN_3)_6^{3+}$

	T_M^{-1} vs log C_T plots				average of curve fits			
GAGGAAGGCGA[a] PCUCUAAUUGCU	-ΔH° (kcal/mol)	-ΔS° (eu)	-ΔG°$_{37}$ (kcal/mol)	T_m[b] (°C)	-ΔH° (kcal/mol)	-ΔS° (eu)	-ΔG°$_{37}$ (kcal/mol)	T_m[b] (°C)
no multivalent	79.9 ± 4.5	246.8 ± 14.9	3.41 ± 0.16	25.3	80.3 ± 3.4	247.9 ± 11.1	3.39 ± 0.13	25.3
[$MgCl_2$], mM								
1.0	83.6 ± 5.3	256.7 ± 17.7	4.04 ± 0.21	28.0	83.4 ± 5.7	255.9 ± 18.8	4.03 ± 0.21	28.0
5.0	74.1 ± 5.3	220.5 ± 17.5	5.69 ± 0.16	33.5	75.6 ± 6.2	225.4 ± 20.6	5.67 ± 0.18	33.5
10.0	87.6 ± 1.1	261.9 ± 3.5	6.40 ± 0.01	36.6	85.1 ± 7.2	253.7 ± 23.5	6.44 ± 0.07	36.7
20.0	78.4 ± 4.8	231.7 ± 15.5	6.56 ± 0.09	37.1	76.6 ± 6.5	225.7 ± 21.2	6.57 ± 0.15	37.1
50.0	83.3 ± 5.2	245.7 ± 16.8	7.09 ± 0.08	39.1	83.7 ± 9.1	247.0 ± 29.7	7.09 ± 0.20	39.1
[$Co(NH_3)_6Cl_3$], mM								
1.0	116.7 ± 18.2	340.4 ± 56.5	11.13 ± 0.77	49.7	109.8 ± 6.2	318.8 ± 18.9	10.92 ± 0.41	49.9
5.0	117.5 ± 16.0	337.6 ± 48.8	12.79 ± 0.89	54.4	116.9 ± 13.8	335.8 ± 41.8	12.78 ± 0.86	54.5
10.0	116.2 ± 10.5	332.0 ± 32.3	13.19 ± 0.54	55.9	114.9 ± 5.4	328.2 ± 16.4	13.15 ± 0.35	56.0
CUUAAAUUUAG[a] CGAAUUUAAAUC								
no multivalent	89.9 ± 2.8	271.2 ± 9.2	5.85 ± 0.06	34.7	89.9 ± 10.2	271.1 ± 33.3	5.85 ± 0.20	34.7
[$MgCl_2$], mM								
1.0	79.3 ± 7.8	236.0 ± 25.4	6.13 ± 0.20	35.4	80.7 ± 10.5	240.3 ± 34.3	6.17 ± 0.15	35.6
10.0	89.6 ± 4.5	262.8 ± 14.5	8.07 ± 0.06	42.4	85.8 ± 8.1	250.7 ± 25.9	8.09 ± 0.17	42.8
50.0	93.1 ± 6.8	271.5 ± 21.6	8.86 ± 0.15	45.0	85.6 ± 7.8	247.4 ± 24.9	8.82 ± 0.17	45.5
[$Co(NH_3)_6Cl_3$], mM								
1.0	100.7 ± 21.4	294.6 ± 67.1	9.34 ± 0.82	45.9	87.3 ± 9.2	252.3 ± 29.3	9.10 ± 0.28	46.4

[a] Solutions are 80 mM NaCl, 3 mM KH_2PO_4, 7 mM K_2HPO_4, and 0.5 mM Na_2EDTA, pH = 6.1. [b] Calculated for 10^{-4} M oligonucleotide concentration.

in the absence of multivalent counterions and in the presence of 10, 20, and 30 mM Mg^{2+} and 1, 2, and 3 mM $Co(NH_3)_6^{3+}$. Chemical shifts are relatively unaffected by the presence of Mg^{2+} or $Co(NH_3)_6^{3+}$, and sequential NOE connectivities in NOESY spectra are also relatively unchanged. Evidently, there are no significant structural changes of the *Ca/Cd* mimic upon addition of Mg^{2+} or $Co(NH_3)_6^{3+}$. In the presence of $Co(NH_3)_6^{3+}$, the T_M of the duplex increases significantly, however, allowing collection of NMR data at higher temperatures. For example, at 1 mM $Co(NH_3)_6^{3+}$ and 1 mM duplex, i.e., 2 mM total strand concentration, the predicted melting temperature of the duplex is 55 °C and the predicted K_d for duplex formation at 30 °C is 0.2 nM on the basis of the thermodynamics reported in Table 2. This results in sharper peaks in general and reveals new, weak, broad cross-peaks. In addition to the higher T_M, the presence of $Co(NH_3)_6^{3+}$ offers additional protons that generate cross-peaks in SNOESY spectra. Therefore, most NMR spectra were collected with $Co(NH_3)_6^{3+}$ present.

Assignment of Exchangeable Proton Resonances of the Ca/Cd Mimic in a 3 mM Cobalt Hexaammine Solution. Imino (10−15 ppm) (Figure 2) and amino (6−9 ppm) regions of 1D NMR and 150 ms SNOESY (Supporting Information) spectra recorded in H_2O provide information concerning hydrogen bonding between base pairs. Only one resonance is observed between 13.5 and 14.5 ppm, the typical range for U imino protons in A-U pairs. This resonance at 14.0 ppm is therefore assigned to the U20 imino proton of the A2-U20 Watson−Crick pair. A cross-peak with A2H2 confirms this assignment. The imino proton of U11b in the terminal A11-U11b pair is not observed presumably due to exchange with water.

For a duplex with four Watson−Crick G-C pairs, four G imino resonances are expected between 11.5 and 13.5 ppm

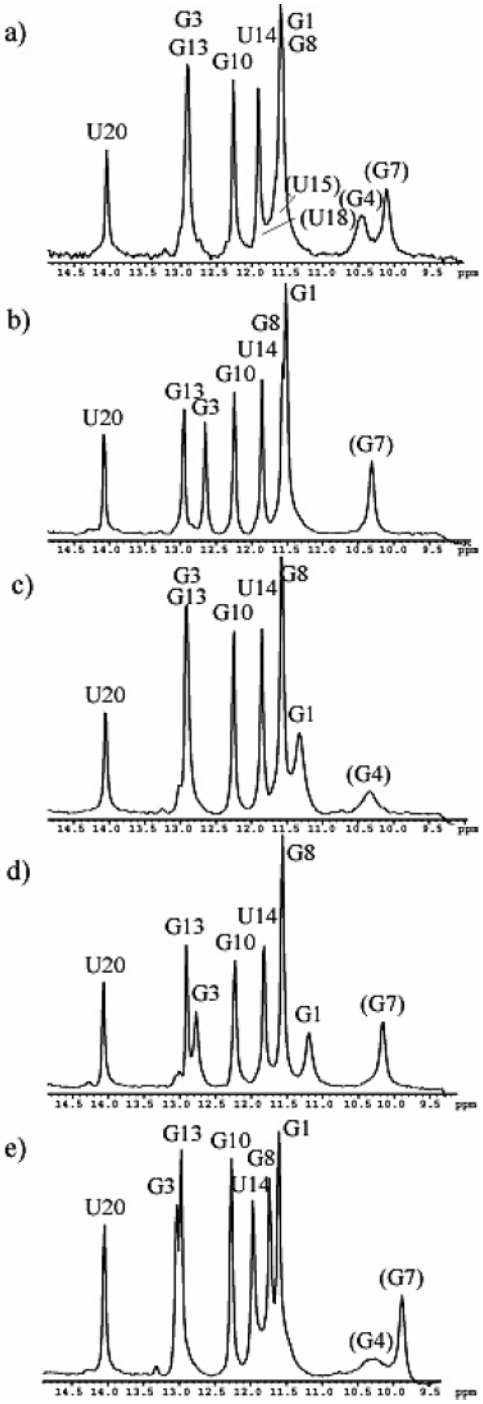

FIGURE 2: One-dimensional imino ^{1}H NMR spectra (9−15 ppm) of the (a) *Ca/Cd* mimic, (b) 4I duplex, (c) 7I duplex, (d) 5/16P duplex, and (e) 6/17P duplex in 80 mM NaCl, 10 mM phosphate, 0.5 mM Na_2EDTA, and 3 mM $Co(NH_3)_6^{3+}$, pH = 6.1 at 10 °C. Base assignment and nucleotide number are given. Assignments in parentheses are tentative.

(*43*) and four are observed (11.6, 12.3, 12.9, and 13.0 ppm) at 10 °C. Two strong cross-peaks to C amino protons are observed for each of these resonances, confirming the assignments to G-C pairs. Residue-specific assignments are made on the basis of additional imino−imino or imino−aromatic cross-peaks. The G1 imino proton (11.6 ppm) is assigned on the basis of cross-peaks to the P22H6 proton. The G3 imino proton (12.9 ppm) is assigned on the basis of its cross-peak with the U20 imino proton. Imino resonances at 11.6 and 11.9 ppm have a very strong cross-peak to each other. Chemical shifts of these resonances and a strong imino−imino cross-peak are indicative of a G·U wobble pair. Additionally, both imino resonances have a cross-peak to the G imino proton at 13.0 ppm. The only potential G·U wobble pair adjacent to a yet unassigned G-C pair is the G8·U14 pair. Therefore, the G imino proton resonance at 13.0 ppm, which has a cross-peak to both the G8 and U14 imino protons, is assigned to G13. The remaining unassigned G imino resonance from a G-C pair (12.3 ppm) is assigned to G10. From this construct alone, imino resonances of G4, G7, U15, and U18 cannot be assigned. There are two broad imino peaks at 10.1 and 10.5 ppm, and integration suggests additional resonances at ∼11.6 and ∼11.8 ppm which are overlapped with other resonances. These resonances can be tentatively assigned on the basis of inosine substitution experiments and are further discussed below. Resonance assignments for all four sets of cytosine amino protons are confirmed by amino−H5 NOEs.

Intermolecular NOEs between $Co(NH_3)_6^{3+}$ protons and RNA imino protons are detected in SNOESY spectra (Supporting Information). At 10 °C, all 18 $Co(NH_3)_6^{3+}$ protons resonate at a single frequency (3.67 ppm), the same chemical shift as free $Co(NH_3)_6^{3+}$ in solution, suggesting that $Co(NH_3)_6^{3+}$ is in fast exchange between the bound and unbound states. The $Co(NH_3)_6^{3+}$ protons have a strong cross-peak to the imino proton of U14 and a weak cross-peak to the imino proton of G8. An additional cross-peak from the $Co(NH_3)_6^{3+}$ protons to an imino proton is observed in the spectra of the 4I construct and is discussed below.

Assignment of Nonexchangeable Resonances of the Ca/Cd Mimic in a 3 mM Cobalt Solution. NMR resonances of the *Ca/Cd* mimic are assigned essentially as described by Varani et al. (*44*, *45*). A table in Supporting Information summarizes chemical shift assignments at 25 °C. H5′ and H5″ assignments are not stereospecific, although H5′ protons are expected to resonate downfield from H5″ protons due to the negative charge on the phosphate (*46*). The (H8/H6/H2)−(H1′/H5) region of the 600 ms NOESY spectrum at 30 °C is shown in Figure 3. Assignments follow standard connectivity pathways from nucleotides G1 to A5. Since the A5H8 and A6H8 resonances are overlapped, the A5−A6 connectivity cannot be confirmed. In addition, the overlap of the A5H8 and A6H8 resonances prevents the observation of distinct sugar−base cross-peaks for this region of the sequence. Due to this overlap, some standard internucleotide restraints were not used in the modeling studies. A standard connectivity pathway continues from nucleotides A6 to A11, with the G7H1′ resonance shifted upfield to ∼4.99 ppm. On the opposite strand, assignments follow standard connectivity pathways from U11b to A17. All U18 resonances are broad and U18 cross-peaks are very weak, including the U18H5−H6 cross-peak. Consequently, there is a break in the NOESY walk at U18. A standard connectivity pathway continues from C19−P22.

Assignments of H2′ resonances follow from strong cross-peaks to H1′ resonances in the 200 ms mixing time NOESY spectrum. H2′ assignments are confirmed by strong H8/H6-(n) to H2′(n − 1) cross-peaks in the base to sugar region of NOESY spectra, as typically observed for A-form conformations. Similar to the overlap in the aromatic region, significant overlap in some portions of the sugar regions prevents the generation of standard internucleotide restraints for the

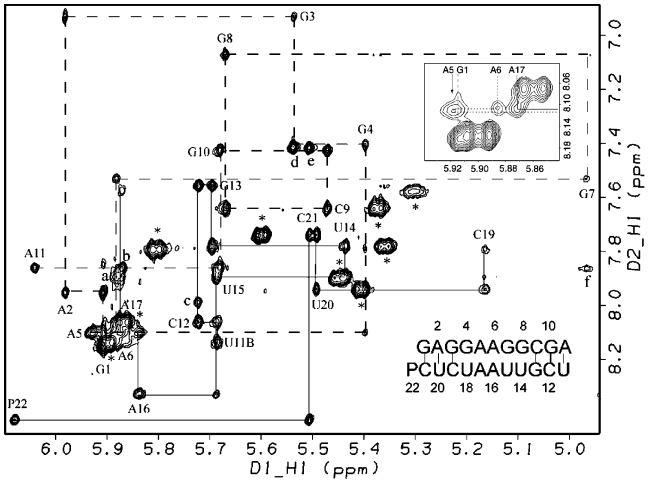

FIGURE 3: (H8/H6/H2)−(H1′/H5) region of the 600 ms mixing time NOESY spectrum of the *Ca/Cd* mimic in solution with 3 mM $Co(NH_3)_6^{3+}$ at 30 °C. Sequential assignments of base to H1′ protons are connected by dotted and solid lines for the upper and lower strands, respectively. Intranucleotide H1′−H6/H8 cross-peaks are labeled. The nine H5−H6 peaks are labeled with an asterisk. The A6H1′−A17H2, A17H1′−A6H2, C12H1′−A11H2, G3H1′−A2H2, C21H1′−A2H2, and G7H1′−A6H2 cross-peaks are labeled a−f, respectively. The inset graph is an expansion of the region near 5.9 and 8.1 ppm drawn at a higher contour level to better delineate the walk through this crowded region.

modeling studies. Observable H1′−H2′ cross-peaks in the TOCSY spectrum (Supporting Information) indicate C2′-*endo* or a mixture of C2′-*endo* and C3′-*endo* sugar conformations for the G7, A11, U11b, A16, A17, and P22 sugars. All ^{31}P resonances are observed in a 1.5 ppm range, consistent with A-form geometry (*47*).

A natural abundance ^{13}C HSQC spectrum was used to identify and/or confirm H2 resonances (Supporting Information). Assignments of AH2 NOESY cross-peaks help to confirm assignment of the upfield-shifted G7H1′ resonance and provide important insight concerning the conformation of the tandem A·A pairs. The assignment of A5H2 is based on the A5H2−A5H1′ and A5H2−A17H8 cross-peaks. The assignment of A6H2 is based on the A6H2−A17H1′ and A6H2−G7H1′ cross-peaks. The A6H2−G7H1′ cross-peak helps to confirm the assignment of the upfield-shifted G7H1′ resonance. The assignment of A16H2 is based on the A16H2−A6H8 cross-peak. Lastly, the assignment of A17H2 is based on the A17H2−A6H1′ cross-peak. The A6H2−A17H1′ and A17H2−A6H1′ cross-peaks lead to important NMR-derived distances of 3.4 and 3.8 Å, respectively. The remaining cross-peaks were too weak to generate significant restraints but are used when considering the validity of the modeled structures.

NMR Spectra with Inosine Substitutions. To gain additional information about the conformation of the G4·U18 and G7·U15 pairs, inosine was substituted for either G4 or G7 to give duplexes 4I and 7I (Figure 1). Two significant changes are observed between the imino spectra of the *Ca/Cd* mimic (Figure 2a) and the 4I duplex (Figure 2b). The resonance of the G3 imino proton is shifted upfield by 0.25 ppm in 4I. This is reasonable since G3 is adjacent to the substitution site and may experience a different shielding environment due to lack of an amino group at I4. In addition, one of the two resonances between 10 and 10.5 ppm in the spectrum of the *Ca/Cd* mimic disappears. The resonance observed at 10.3 ppm in the 4I spectrum has the same line shape and same cross-peak to an amino proton as does the resonance at 10.1 ppm in the *Ca/Cd* spectrum. Therefore, it appears the resonance at 10.5 ppm in the *Ca/Cd* spectrum disappears in the 4I spectrum. There is also a major change in the NOESY spectra (Supporting Information). In the spectra of the *Ca/Cd* mimic, typical NOE connectivities are observed from the 5′ end into the loop region of the top strand. Substitution of inosine at position 4 results in a discontinuity in the NOE connectivities between G3 and A6. These changes in the imino and NOESY spectra suggest that the amino group of G4 may be involved in hydrogen bonding that stabilizes the structure. When the amino group is removed, the region near I4 becomes more dynamic than the region near G4 in the *Ca/Cd* mimic. The imino resonance at 10.5 ppm in the *Ca/Cd* spectrum that is not observed in the 4I spectrum is, therefore, tentatively assigned to the G4 imino proton. As a result, only one G imino proton, G7, remains unassigned. Since the resonance at 10.1 ppm in the spectrum of the *Ca/Cd* mimic has a medium intensity NOE to an amino proton, this resonance is assigned to G7.

Additional information can also be obtained from the SNOESY spectrum of the 4I duplex. The 4I duplex sample was slightly more concentrated than the *Ca/Cd* mimic sample, and two additional weak cross-peaks are observed in the 4I spectrum (Supporting Information). One is an additional cross-peak from $Co(NH_3)_6^{3+}$ protons to an imino proton. This cross-peak does not line up with previously assigned imino protons but is in the region where integration suggests an overlapped imino proton may resonate, 11.6 ppm. Assuming only one $Co(NH_3)_6^{3+}$ binding site, this resonance was assigned to the U15 imino proton, because U15 is adjacent to G8 and U14, whose imino protons exhibit an NOE to $Co(NH_3)_6^{3+}$. This assignment is further confirmed by a very weak NOE between this newly assigned U15 imino proton and the tentatively assigned G7 imino proton, an NOE that would be expected if these bases formed a G·U pair in which the imino protons of G and U are close in space, such as a wobble G·U pair. The only remaining unassigned imino proton is U18. Integration suggests there may be another resonance at ∼11.8 ppm, and this is therefore assigned to U18.

NMR spectra of the 7I duplex (Figure 2c) also differ from those of the *Ca/Cd* mimic (Figure 2a). Consistent with the disappearance of the G4 imino resonance in the I4 imino proton spectrum, the resonance of the G7 imino proton disappears in the imino spectrum of the I7 duplex. Also, in the NOESY spectrum (Supporting Information), the assignment of the I7H1′ resonance cannot be confirmed. These changes in the imino and NOESY spectra suggest that the amino group of G7 may also be involved in hydrogen bonding that stabilizes the structure. When the amino group is removed, the region near I7 may be more dynamic than the region near G7 in the *Ca/Cd* mimic.

NMR Spectra with Purine Substitutions. To gain additional information about the conformation of the A5·A16 and A6·A17 pairs, purines were substituted for adenines, resulting in two duplexes, 5/16P and 6/17P (Figure 1). For the 5/16P duplex, one minor change is observed in the imino proton spectrum (Figure 2d); the resonance of G4 disappears. Since it is broad and weak in the *Ca/Cd* mimic, the replacement of purine for adenine adjacent to G4 may have shifted the

G4 imino resonance such that it is overlapped with the G7 imino resonance or may have caused enough dynamical changes to broaden the resonance so that it is no longer observed. There are also minor changes in the NOESY spectrum (Supporting Information); the G4H1′−P5H8 cross-peak is missing and the G7H1′ resonance is not observed. These are not significant changes since in the spectrum of the *Ca/Cd* mimic the G4H1′−A5H8 cross-peak is weak and the G7H1′ resonance is broad. Evidently, removal of the A5 and A16 amino groups does not substantially alter the loop conformation.

For the 6/17P duplex, two changes are observed in the imino spectrum (Figure 2e). Resonances of the G4 and G7 imino protons shift upfield by 0.20 and 0.25 ppm, respectively, suggesting that these two protons may be in a different local environment. There are many significant changes in the NOESY spectrum of 6/17P (Supporting Information), suggesting a major structural change. For example, the chemical shift of the G7H1′ resonance appears at 5.60 ppm, which is 0.65 ppm downfield in comparison to the G7H1′ resonance in the *Ca/Cd* mimic at the same temperature. Changes are also observed for the U18 base. The U18H5−H6 cross-peak in the 6/17P duplex appears to be typical, unlike the weak, broad cross-peak observed in other constructs. In addition, A5H1′ and A16H1′ appear as doublets, suggesting that these sugars spend a percentage of the time in a C2′-*endo* conformation. Lastly, NOESY connectivities from nucleotides A5−G7 and U15−P17 are weak. Evidently, removal of the A6 and A17 amino groups makes a significant change in the conformation of the loop. Changes described above are consistent with the 6/17P loop having a face-to-face rather than sheared conformation for the A·P pairs.

Structural Modeling of the Ca/Cd Mimic in a 3 mM Cobalt Hexaammine Solution. Because the NMR spectra show no changes upon the addition of $Co(NH_3)_6^{3+}$, most simulations were done in the absence of $Co(NH_3)_6^{3+}$. Simulations with $Co(NH_3)_6^{3+}$ are considered only when discussing the binding site and local interactions with the metal cation.

The NMR data are consistent with tandem A·A pairs in a sheared conformation. For example, the AH1′−AH2 cross-strand NOE cross-peaks and the upfield-shifted G7H1′ resonance (4.99 ppm at 35 °C) are consistent with sheared A·A pairs. In addition, the structural and thermodynamic results of the 5/16P and 6/17P studies are also consistent with sheared A·A pairs in the *Ca/Cd* mimic (discussed below). For these reasons, artificial restraints (amino proton−2′OH, 1.5−2.8 Å, and amino proton−N3, 1.4−2.7 Å) that force the tandem A·A pairs of the *Ca/Cd* mimic into a sheared conformation were added during the molecular dynamics simulation.

Forty of 50 structures generated from helical starting structures by the rMD protocol described in Materials and Methods converge and have simulated total energies of −425 ± 6 kcal/mol. Superposition reveals that the overall structure and local features are consistently reproduced (Figure 4). The average RMSD for the all-atom pairwise superposition of these 40 structures is 0.69 Å. The structure closest to the average satisfies all NMR-derived distance restraints within 0.15 Å and all angle restraints within 0.3°. Because artificial restraints were used to force the A·A pairs into a sheared conformation, the resulting structures were checked to confirm that forcing these pairs into a sheared conformation did not create violations of the NMR-derived distance restraints. There were no violations greater than 0.10 Å for inter- or intrastrand restraints involving the adenosine residues in the loop.

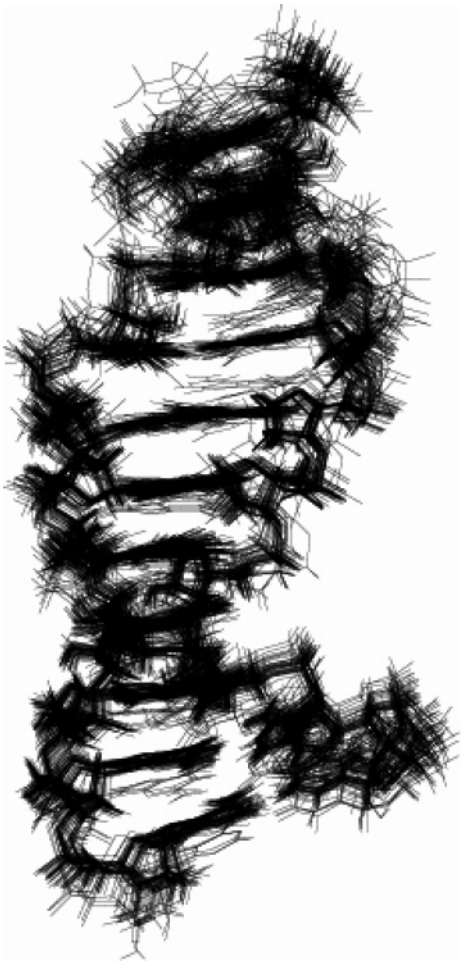

FIGURE 4: Superposition of 40 low-energy structures derived from restrained molecular dynamics. P22 was omitted.

As expected from the artificial hydrogen-bonding restraints, these structures contain *trans* Hoogsteen/sugar edge (sheared) A_{anti}·A_{anti} pairs (Figure 5a) similar to the sheared A·A pairs described by Cate et al. (*23*), Strobel et al. (*12*), and Znosko et al. (*34*). Average hydrogen bond distances are 2.08, 2.08, 2.24, and 2.22 Å for A17H62−A5N3, A6H61−A16N3, A17H61−A5O2′, and A6H62−A16O2′, respectively. The conformation of the sheared A·A pairs results in a cross-strand stack of A6 and A17 (Figure 6e). This is consistent with the observation of the A6H1′−A17H2 and A17H1′−A6H2 cross-peaks in the NOESY spectrum. The G7H1′ proton is situated below the ring of the A6 base, which is consistent with the upfield shift (∼4.99 ppm) of the G7H1′ resonance.

As expected due to the G8·U14 hydrogen bond restraints, G8·U14 is in a *cis* Watson−Crick/Watson−Crick (wobble) conformation (Figure 5b). The G8·U14 wobble pair has extensive stacking with the adjacent C9-G13 pair (Figure 6j). The conformations of the G4·U18 and G7·U15 pairs are ambiguous. In the 40 models, the G4 and U18 residues always form a *cis* Watson−Crick/Watson−Crick standard wobble pair. Because a G4−U18 imino−imino cross-peak was not observed and U18 appears to be dynamic, modeling trials were also done forcing the G4 and U18 imino protons

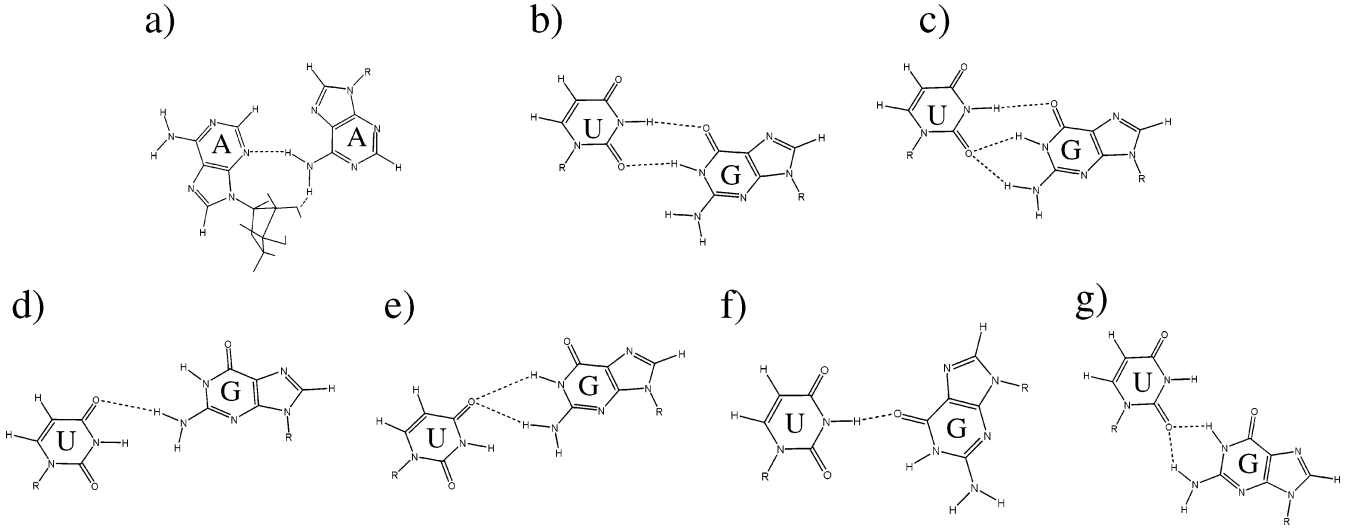

FIGURE 5: (a) *trans* Hoogsteen/sugar edge (sheared) A·A pair, (b) *cis* Watson−Crick/Watson−Crick standard wobble G·U pair, (c) *cis* Watson−Crick/Watson−Crick (bifurcated wobble) G·U pair, (d) *cis* Watson−Crick/Watson−Crick (amino−O4) G·U pair (*65*, *78*), (e) *cis* Watson−Crick/Watson−Crick (bifurcated amino:imino−O4) G·U pair (*79*−*82*), (f) *trans* Hoogsteen/Watson−Crick (O6−imino) G·U pair (*24*, *83*), and (g) *cis* Watson−Crick/Watson−Crick (bifurcated amino:imino−O2) G·U pair (*23*). Nomenclature is that of Leontis and Westhof (*84*).

to be separated by at least 4.0 Å. Slight shifts in the bases due to this restraint result in different hydrogen-bonding patterns in the G4·U18 pair, including bifurcated wobble (Figure 5c), amino−O4 (Figure 5d), and bifurcated imino: amino−O4 pairs (Figure 5e). NMR data suggest dynamics at this pair. Thus, we speculate that an equilibrium between these structures is probable.

In the 40 lowest energy models generated in the absence of $Co(NH_3)_6^{3+}$, the G7 and U15 nucleotides always form a *cis* Watson−Crick/Watson−Crick standard wobble conformation (Figure 5b). NMR data show no direct evidence that the G7·U15 wobble pair is forming, however, and there is evidence to suggest dynamics at this base pair; the G7 resonances are broad and cross-peaks are weak. Because a G7−U15 imino−imino cross-peak was not observed, modeling trials were also done, forcing the G7 and U15 imino protons to be separated by at least 4.0 Å. These trials resulted in conformations that were inconsistent with the NMR data. The simulations done in the presence of $Co(NH_3)_6^{3+}$, however, suggest that the G7·U15 pair is in a *trans* Hoogsteen/Watson−Crick (O6−imino) conformation (Figure 5f). Thus, we speculate that G7·U15 is in a dynamic equilibrium between the *trans* and *cis* conformations.

Further evidence that G7 and U18 are flexible comes from weak A5C2−H2 and A16C2−H2 resonances in the proton−carbon HSQC spectrum and that A5H2 and A16H2 proton line widths (taken from NOE cross-peaks and the carbon HSQC) are broad relative to other AH2 signals. A5 and A16 are the hydrogen bond acceptors in the sheared A·A pairs, and their H2 protons are stacked underneath the G7 and U18 bases while A6H2 and A17H2 are near the surface of the helix. Thus, A5H2 and A16H2 are expected to be influenced more by G7 and U18 dynamics.

Modeling studies done in the presence of $Co(NH_3)_6^{3+}$ confirm that the tandem G·U pairs are the binding site of the cation. The $Co(NH_3)_6^{3+}$ is positioned in a binding pocket defined by the major groove surface of the G·U base pairs. Only hydrogen bond acceptors line this site; therefore, the electrostatic potential is highly negative. The NMR data and analysis of the different cobalt−RNA structures suggest that although the position of the hexaammine is localized at the center of the G·U pairs, multiple orientations and small displacements of the cobalt ion are possible. These different orientations result in different patterns of hydrogen bonding of the hexaammine protons with the RNA acceptor groups. Most conformations of $Co(NH_3)_6^{3+}$ involved hydrogen bonds between the hexaammine protons and the following acceptor groups: the two nonbridging phosphate oxygens of A6 and the one nonbridging phosphate oxygen of G7, G7N7, G8O6, G8N7, U14O4, and U15O4.

DISCUSSION

Effects of Na^+, Mg^{2+}, and $Co(NH_3)_6^{3+}$ on the Ca/Cd Mimic. As expected, addition of Na^+, Mg^{2+}, or $Co(NH_3)_6^{3+}$ stabilizes the *Ca/Cd* mimic (Tables 1 and 2). Addition of Mg^{2+} or $Co(NH_3)_6^{3+}$ produces no significant chemical shift changes and no changes in the NOESY connectivities. Evidently, Mg^{2+} and $Co(NH_3)_6^{3+}$ do not alter the structure of the *Ca/Cd* mimic.

The mimic in 1 M Na^+ is 3.8 kcal/mol more stable at 37 °C than the mimic in 80 mM Na^+ and has a 14.5 °C higher melting temperature. This is similar to the 3.1 kcal/mol stabilization and 11.4 °C higher melting temperature observed for the Watson−Crick paired 11 bp duplex (Tables 1 and 2). Serra et al. (*48*) observed a 2.4 kcal/mol stabilization and 15.1 °C increase in melting temperature in going from 100 mM to 1 M Na^+ for a 14-mer Watson−Crick RNA duplex (*48*). Evidently, the effects of Na^+ on stability are not specific for the J4/5 internal loop.

Addition of 50 mM Mg^{2+} to an 80 mM Na^+ solution stabilized the mimic by 3.7 kcal/mol at 37 °C and increased the melting temperature by 13.8 °C. In comparison, addition of 50 mM Mg^{2+} stabilized the 11 bp duplex by 3.0 kcal/mol and increased the melting temperature by 10.3 °C. Evidently, the effects of Mg^{2+} on stability are also not specific to the J4/5 loop. Interestingly, in a background of 80 mM Na^+, about 50 mM Mg^{2+} is required to provide roughly the same

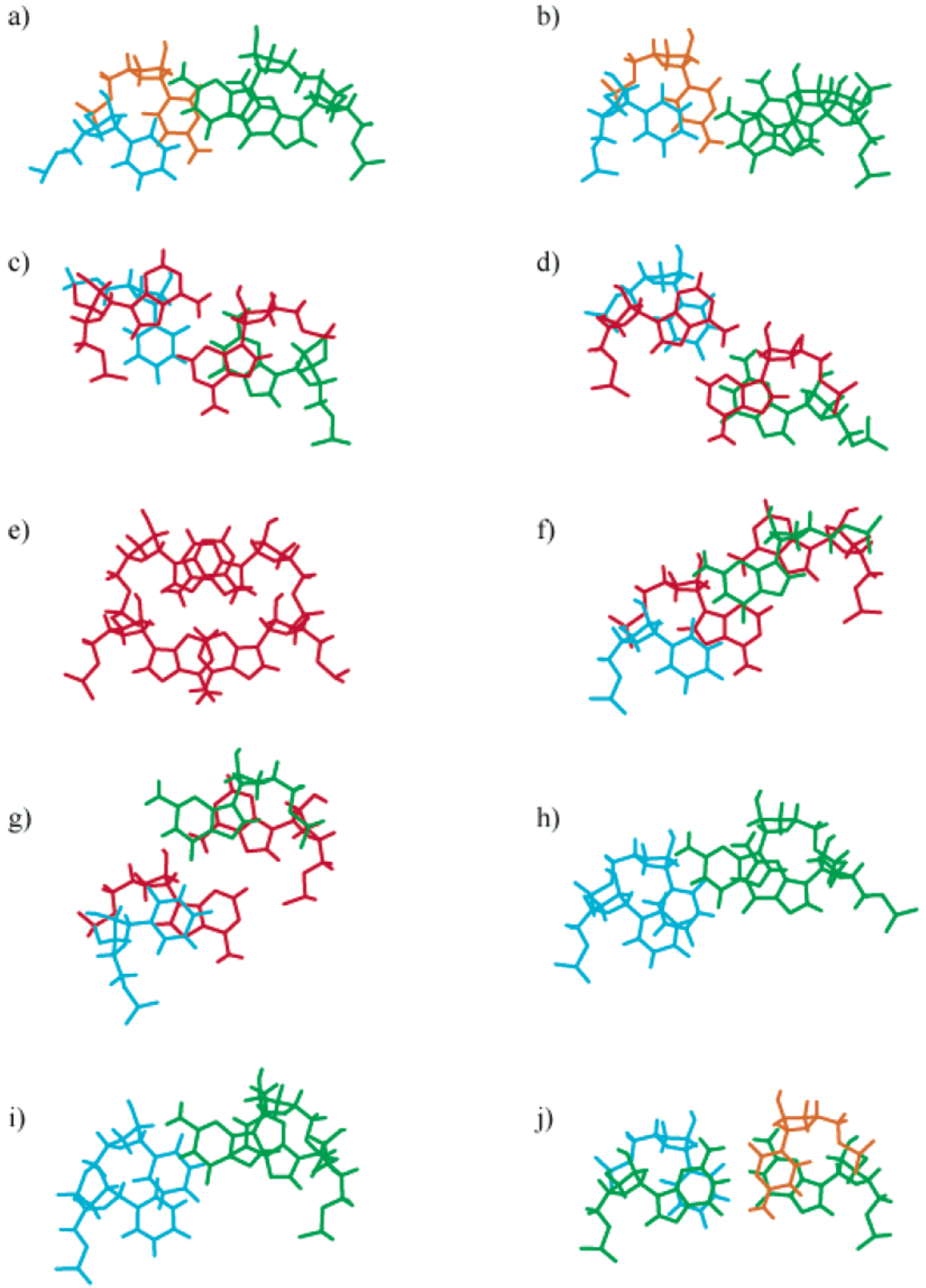

FIGURE 6: Stacking of (a) $^{19C-G3}_{18U\cdot G4}$ if G4·U18 is forming a standard wobble pair, (b) $^{19C-G3}_{18U\cdot G4}$ if G4·U18 is forming an amino−O4 pair, (c) $^{18U\cdot G4}_{17A\cdot A5}$ if G4·U18 is forming a standard wobble pair, (d) $^{18U\cdot G4}_{17A\cdot A5}$ if G4·U18 is forming an amino−O4 pair, (e) $^{17A\cdot A5}_{16A\cdot A6}$, (f) $^{16A\cdot A6}_{15U\cdot G7}$ if G7·U15 is forming a standard wobble pair, (g) $^{16A\cdot A6}_{15U\cdot G7}$ if G7·U15 is forming an O6−imino pair, (h) $^{15U\cdot G7}_{14U\cdot G8}$ if G7·U15 is forming a standard wobble pair, (i) $^{15U\cdot G7}_{14U\cdot G8}$ if G7·U15 is forming an O6−imino pair, and (j) $^{14U\cdot G8}_{13G-C9}$. Nucleotides are colored by base type: adenine (red), cytosine (orange), guanosine (green), and uridine (blue). Figures were created with 3DNA (*85*).

duplex stability as 1 M Na^+ in the absence of Mg^{2+} (Tables 1 and 2).

Addition of only 1 mM $Co(NH_3)_6^{3+}$ to an 80 mM Na^+ solution stabilized the mimic by 7.3 kcal/mol at 37 °C whereas addition of 50 mM Mg^{2+} or 0.92 M Na^+ stabilized it by only 3.7 and 3.8 kcal/mol, respectively. The corresponding increases in melting temperature are 25.2, 13.8, and 14.5 °C, respectively. In comparison, addition of 1 mM $Co(NH_3)_6^{3+}$ stabilized the 11 bp duplex by 3.5 kcal/mol and increased the melting temperature by 11.2 °C. The unusually large stabilization of the *Ca/Cd* mimic by $Co(NH_3)_6^{3+}$ suggests sequence- or structure-specific interactions with $Co(NH_3)_6^{3+}$, even though $Co(NH_3)_6^{3+}$ does not alter the structure of the *Ca/Cd* mimic.

In the SNOESY spectrum of the *Ca/Cd* mimic, cross-peaks are observed between hexaammine protons and the G8 and

U14 imino protons. The G8·U14 wobble pair appears to be a binding site for $Co(NH_3)_6^{3+}$. This is consistent with reports in the literature; $Co(NH_3)_6^{3+}$ binds the major groove of G·U pairs, due to the deep negative electrostatic potential there (*19*−*22*).

It is thought that the structure of $Co(NH_3)_6^{3+}$ mimics that of $Mg(H_2O)_6^{2+}$ and that the two complexes should have similar binding sites in RNA (*25*). The results in Table 2 indicate that $Co(NH_3)_6^{3+}$ stabilizes the *Ca/Cd* J4/5 loop much more than $Mg(H_2O)_6^{2+}$. This does not necessarily mean that $Mg(H_2O)_6^{2+}$ does not bind this loop. The additional charge on $Co(NH_3)_6^{3+}$ favors binding, the tighter complexation of NH_3 to Co^{3+} relative to H_2O to Mg^{2+} may reduce the entropic penalty for binding, and $Co(NH_3)_6^{3+}$ has more potential hydrogen bond donors. It is also possible, however, that the structural differences between $Co(NH_3)_6^{3+}$ and $Mg(H_2O)_6^{2+}$ lead to different binding. Thus, it will be interesting to compare RNA binding sites for $Mg(H_2O)_6^{2+}$ and $Co(NH_3)_6^{3+}$ in future crystal structures. If the binding sites are similar, then optical melting and/or temperature gradient gel electrophoresis studies (*49*, *50*) with $Co(NH_3)_6^{3+}$ may provide a fast assay for RNA motifs that bind $Mg(H_2O)_6^{2+}$.

Conformation of the Tandem A·A Pairs. NMR spectra of the *Ca/Cd* mimic are consistent with the formation of tandem sheared A·A pairs with a cross-strand stack. For example, A6H2−A17H1′ and A17H2−A6H1′ NOEs are observed in the NOESY spectrum. The A6H2−A17H1′ and A17H2−A6H1′ cross-peaks lead to NMR-derived distances of 3.4 and 3.8 Å, respectively. These cross-peaks suggest that the minor groove is narrower than expected for a head-to-head A·A pair in A-form geometry, where a 4.7 Å distance is found for equivalent protons in the same sequence. A narrowing of the minor groove is consistent with the formation of sheared A·A pairs. In a Watson−Crick duplex, the distance between AH2 protons and cross-strand H1′ protons can range from 3.2 to 4.1 Å (*51*). It is unlikely, however, that these adenosines in the *Ca/Cd* mimic are forming A-U pairs. The upfield shift of the G7H1′ resonance to ∼4.99 ppm is also consistent with sheared A·A pairs. In standard A-form helices, typical H1′ shifts are ∼5.5−6.1 ppm. Positioning of an H1′ proton above or below a base with a strong ring current, however, results in an upfield shift. Similar AH1′−AH2 cross-strand NOEs and upfield-shifted H1′ resonances were also observed for the tandem sheared A·A and G·A pairs, respectively, in (rGGCAAGCCU)$_2$ (*34*) and (rGGCGAGCC)$_2$ (*52*). In addition, although they were not used to generate restraints, other AH2 cross-peaks are consistent with the A·A conformations in the modeled structures.

The thermodynamic and structural changes observed in the 5/16P and 6/17P duplexes also support the hypothesis that the A·A pairs in the *Ca/Cd* mimic are in a sheared conformation. In sheared A·A pairs, the non-hydrogen-bonding amino groups are situated in the major groove (Figure 5a). Thus, removal of these amino groups by substitution with purine should not change the structure of the loop. When those amino groups were removed from the *Ca/Cd* mimic, to give the 5/16P duplex, the resulting NMR spectra suggest there is no conformational change. In tandem sheared A·A pairs, however, the adenine amino groups that are not situated in the major groove are involved in hydrogen bonding to the base and sugar of the cross-strand adenosine (Figure 5a). Removal of these amino groups is expected to destabilize the structure, due to the loss of hydrogen bonding. When these amino groups were removed from the *Ca/Cd* mimic to give the 6/17P duplex, the resulting NMR spectra indicated major structural changes. One observed change is the increased dynamics of the A5 and A16 nucleotides in comparison to the equivalent nucleotides in the *Ca/Cd* mimic. The H1′ resonances of the A5 and A16 nucleotides appear as doublets, indicating a C2′-*endo* sugar pucker for these sugars. Also, NOESY connectivities involving the purines are weak or missing. In addition, the A6H2−A17H1′ and A17H2−A6H1′ cross-peaks observed in the NOESY spectrum of the *Ca/Cd* mimic are not observed in the 6/17P spectrum. Lastly, the G7H1′ resonance with a chemical shift of ∼4.99 ppm in the spectrum of the *Ca/Cd* mimic has a chemical shift of 5.60 ppm in the 6/17P spectrum, suggesting that this proton is no longer beneath the P6 base. Optical melting studies revealed a destabilization of 0.8 kcal/mol at 37 °C for substitution of A6 and A17 with purine (Table 1). These optical melting results indicate a destabilization that is consistent with changes in the structure and dynamics of the loop.

The results observed for purine substitutions in the *Ca/Cd* mimic parallel those observed for purine substitutions in (rGGCAAGCCU)$_2$ (*53*), a duplex containing tandem sheared A·A pairs. When the non-hydrogen-bonded amino groups of the sheared A·A pairs were replaced with hydrogen, no structural changes were observed, similar to the lack of changes observed in the 5/16P spectra in comparison to the *Ca/Cd* mimic spectra. When the hydrogen-bonded amino groups were replaced with hydrogen, thermodynamic and structural consequences were similar to the changes observed in the 6/17P thermodynamics and spectra in comparison to the *Ca/Cd* mimic. There was a thermodynamic destabilization of 0.8 kcal/mol at 37 °C, the cross-strand AH2−AH1′ NOEs were not observed, and the H1′ adjacent to the loop was not shifted upfield. These NMR data resulted in a loop model containing face-to-face A·P rather than sheared A·P pairs.

Conformations of the Potential G·U Pairs. Wobble G·U pairs are characterized by a strong imino−imino cross-peak and imino chemical shifts of 10−12 ppm (*54*). The G8·U14 pair exhibits both of these characteristics and is evidently in a wobble conformation.

The conformation of the G4·U18 pair is ambiguous. No imino−imino cross-peak is observed for the G4·U18 pair. In addition, the NMR data suggest that U18 is dynamic. Pyrimidine H5−H6 NOESY and TOCSY cross-peaks are typically very strong. The U18H5−H6 cross-peaks in the NOESY and TOCSY spectra, however, are broad and weak. In addition, the U18H1′ resonance is never observed, resulting in a break in the NOESY walk. All other U18 cross-peaks are very weak or missing. This all suggests that U18 is dynamic.

While in the 40 models the G4 and U18 residues always form a *cis* Watson−Crick/Watson−Crick standard wobble pair (Figure 5b), the NMR data suggest that this pair is more dynamic. When the G4 and U18 imino protons are forced to be at least 4.0 Å apart, slight shifts in the bases result in different hydrogen-bonding patterns, including bifurcated wobble, amino−O4, and bifurcated amino:imino−O4 pairs (Figure 5c−e). When G4 is replaced by inosine, the I4 nucleotide appears more dynamic than the G4 nucleotide,

suggesting that the amino group of G4 may be hydrogen bonded to U18. In a standard wobble G·U pair, the amino group of G is not involved in hydrogen bonding. In the bifurcated wobble pair, the amino−O4 pair, and the bifurcated amino:imino−O4 pair, however, the G amino group is involved in hydrogen bonding. We thus speculate that the G4·U18 pair is in a dynamic equilibrium between structures such as those in Figure 5b−e. This may be due to the relative instability of a 5′GA/3′UA stack, as suggested by the thermodynamic instability of the motif 5′GAAU/3′UAAG (*42*).

A combination of hydrogen bonding, base stacking, and backbone distortion results in stabilizing and destabilizing forces associated with each G·U pair. The amino−O4 pair (Figure 5d) contains one hydrogen bond, the wobble and bifurcated amino:imino−O4 pairs (Figure 5b,e) contain two hydrogen bonds, and the bifurcated wobble pair contains three hydrogen bonds (Figure 5c). There is disagreement over the energetics of bifurcated hydrogen bonds (*55*, *56*), so it is not clear which hydrogen bonds are most favorable. Moreover, other energetic factors are probably important. For example, the wobble pairs may be destabilized due to backbone distortion. In a wobble conformation, the α and γ dihedral angles of G4 are outside the range of standard A-form values, whereas in an amino−O4 pair, for example, these angles fall within the range of standard A-form values. Although the amino−O4 pair only contains one hydrogen bond, the conformation of the amino−O4 pair may be stabilized by stacking. When in an amino−O4 conformation, the U18 base stacks below the A17 base to a greater extent than when U18 is in a standard wobble pair (Figure 6c,d). The A17 base is involved in a cross-strand stack with the A6 base, resulting in a U18-A17-A6 three base stack. Thus, neither the NMR data nor current understanding of nucleic acid energetics favors one G4·U18 conformation, and it is likely that the conformation is dynamic.

The conformation of the G7·U15 pair is also ambiguous. In the 40 models generated without $Co(NH_3)_6^{3+}$, the G7 and U15 nucleotides always form a *cis* Watson−Crick/Watson−Crick standard wobble conformation. On the basis of the NMR data, however, it is unlikely that a standard wobble pair is the only conformation for G7·U15. There is no observable imino−imino cross-peak, a common feature of G·U wobble pairs. In addition, the G7H1′ resonance is broad, its cross-peaks are very weak and only observable at high temperatures, and its coupling to G7H2′ indicates a mixed sugar pucker. This suggests that the G7 residue is dynamic. The NMR data suggest that the G7·U15 pair may be switching between a standard wobble pair and other conformations. In the 10 models generated with $Co(NH_3)_6^{3+}$, the converging structures suggest that G7 and U15 are forming an O6−imino pair (Figure 5f). This conformation creates a deep binding pocket lined with functional groups having partial negative charges (Figure 7). This conformation appears best suited to accommodate $Co(NH_3)_6^{3+}$, a large sphere of positive charge.

Cobalt Hexaammine Binding Site. $Co(NH_3)_6^{3+}$ is bound within the major groove of the tandem G·U pairs. The ion is situated more closely to the guanines and more distant from the uracil residues. The hexaammine−RNA NOEs, hexaammine−G8H1, −U14H3, and −U15H3, position the metal ion where it is able to form hydrogen bonds with the

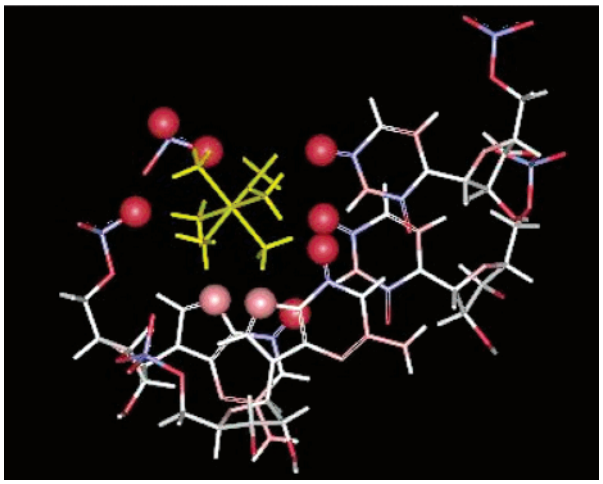

FIGURE 7: One possible configuration of the $Co(NH_3)_6^{3+}$ binding site. Shown in yellow is $Co(NH_3)_6^{3+}$. The G7, G8, U14, and U15 nucleotides and the A6 phosphate group are colored by a charge spectrum (red = negative, blue = positive). The A6 base and sugar were removed for clarity. Atoms shown as balls are hydrogen bond acceptors within hydrogen-bonding distance of the hexaammine protons.

N7 of both G7 and G8, O6 of G8, and O4 of both U14 and U15. There is also the possibility of hydrogen bonds between the hexaammine and the phosphate oxygens, A6O1P, A6O2P, and G7O2P (Figure 7).

Although this positioning of $Co(NH_3)_6^{3+}$ is most prevalent in the simulated structures, there is likely an ensemble of possibilities. The evidence that the G7·U15 pair is dynamic is discussed above. In addition, the single frequency of the 18 amino protons of the hexaammine suggests that the ion is exchanging rapidly with free $Co(NH_3)_6^{3+}$. Also, there are few NMR restraints between the hexaammine and the RNA. It appears that there is an ensemble of orientations of $Co(NH_3)_6^{3+}$ in the binding pocket, resulting in a variety of different hydrogen-bonding interactions. The $Co(NH_3)_6^{3+}$ binding site in the *Ca/Cd* mimic is similar to the $Mg(H_2O)_6^{2+}$ binding site in the P4−P6 region of the *T. thermophila* group I intron (*23*, *24*). Although the *T. thermophila* sequence does not have a tandem G·U pair in the J4/5 loop, the $Mg(H_2O)_6^{2+}$ binds near adjacent Gs, 5′GAAAGG3′/3′CAAUC5′, in a similar location with a sequence similar to that of the tandem G·U pair in the *Ca/Cd* mimic, 5′GAAGG3′/3′UAAUU5′. The $Co(NH_3)_6^{3+}$ binding site and interactions are also similar to $Co(NH_3)_6^{3+}$ binding sites described by the Tinoco laboratory (*21*, *57*−*59*).

Comparison to Other J4/5 Loops. The studies reported here provide a reasonable model for the structure of the *Ca/Cd* mimic in solution. While there may be other models consistent with all of the data, none are evident from homology modeling with known motifs. The NMR structure model can be compared to the crystal structures of the *T. thermophila* J4/5 loop, $^{5'GAAAG}_{3'CAA\ U}$, and the *Azoarcus* J4/5 loop, $^{5'GAAAG}_{3'CAACG}$, and the NMR structure model of the *Pneumocystis carinii* J4/5 loop, $^{5'GGAAG}_{3'UAG\ U}$. The tandem sheared A·A pairs of the *Ca/Cd* mimic are identical to the A113·A207 pair and similar to the A114·A206 pair (which is lacking the amino proton−2′OH hydrogen bond) in the *T. thermophila* structure. Tandem A·A pairs are a common feature of the J4/5 loop of group I introns (*9*, *14*, *60*), suggesting a functional role. For example, the A113 and A206 residues

of the *T. thermophila* group I intron, which are comparable to the A5 and A16 nucleotides in the *Ca/Cd* mimic, are adenosines in 71% and 83%, respectively, of group I introns. Similarly, the A114 and A207 residues of the *T. thermophila* group I intron, which are comparable to the A6 and A17 nucleotides in the *Ca/Cd* mimic, are adenosines in 99% and 100%, respectively, of group I introns (*9*).

One possible reason for conservation of these nucleotides is the availability of functional groups of sheared purine−purine pairs for hydrogen bonding in the major and minor grooves (*52*). In the *Ca/Cd* mimic, the N7 and NH6 atoms of the A5 and A16 nucleotides extend into the major groove. With the N7 and NH6 atoms available for hydrogen bonding, the A5 and A16 nucleotides 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 A6 and A17 extend into the minor groove. These groups are also available to form tertiary or binding contacts with proteins or with other nucleotides. For example, in *T. thermophila* (*23*), 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 (*60*, *61*). 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 (*60*, *61*). 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. Although the A·A pairs of the *Ca/Cd* mimic are G·A pairs in the *P. carinii* loop, it has been proposed that these G·A pairs form sheared G·A pairs (*62*) and have a geometry similar to that of the A·A pairs reported here.

Other conformations are conceivable for J4/5 loops (*63*). For example, A5 or A16 could be *syn*. The NMR data of the *Ca/Cd* mimic are not consistent with a *syn* conformation for A5 or A16, however. In particular, one major consequence of the *syn* orientation would be a distance from A5H8 to G4H2′ or A16H8 to U15H2′ of about 6 Å. The NMR spectra for the *Ca/Cd* mimic, however, have strong NOE cross-peaks between A5H8−G4H2′ and A16H8−U15H2′ protons that correspond to distances of 2.5 and 2.9 Å, respectively. The AH8−AH1′ distances for A5 and A16 are 3.5 and 3.7 Å, respectively, while they are expected to be less than 3.0 Å for a *syn* conformation. The A16H8−A16H3′ distance is 3.0 Å in contrast to an expected distance of more than 4 Å. Evidently, both A5 and A16 are in *anti* conformations.

It is interesting to compare the conformation of base pairs adjacent to A·A pairs in all four structures. The *T. thermophila*, *P. carinii*, and *Azoarcus* J4/5 loops contain a G-C, G·U, and G-C pair, respectively, in the equivalent position as the G4·U18 pair in the *Ca/Cd* mimic. It is straightforward to make a comparison between the G4·U18 pair and the equivalent pair in *P. carinii*. In *P. carinii*, only a weak imino−imino cross-peak is observed in the H_2O spectrum, leading to the suggestion that this G·U pair is dynamic (*62*). As discussed above, the G4·U18 pair in the *Ca/Cd* mimic also exhibits dynamic characteristics even though it closes a size symmetric rather than size asymmetric loop.

The *T. thermophila* and *P. carinii* sequences contain a G·U pair equivalent to the G8·U14 pair in the *Ca/Cd* mimic. The G8·U14 pair apparently has a standard wobble conformation in the *Ca/Cd* mimic. For the *P. carinii* J4/5 mimic, a strong imino−imino cross-peak is observed in the NOESY spectrum so similar artificial hydrogen-bonding restraints were used that forced the G·U pair into a standard wobble conformation (*62*). The equivalent pair in the crystal structure of *T. thermophila* forms a bifurcated amino:imino−O2 conformation (*23*) (Figure 5g), a slight shift of the bases in comparison to the standard wobble pairs observed in the *Ca/Cd* mimic and the *P. carinii* mimic. While this conformation also brings the imino protons close in space, the U imino proton would be expected to exchange with water and have a resonance that is broader and further upfield than observed with the *P. carinii* and *Ca/Cd* J4/5 loops.

The characteristics of the G7·U15 pair in the 2 × 2 nt *Ca/Cd* J4/5 loop are similar to those of the additional loop nucleotide in the 2 × 3 nt *T. thermophila* and *P. carinii* J4/5 loops. In *P. carinii*, the extra A nucleotide is dynamic and exhibits weak cross-peaks to other protons, similar to G7 in the *Ca/Cd* mimic. In the crystal structure of the *T. thermophila* J4/5 loop, A115 is situated below the base of A114. Similarly, G7 in the *Ca/Cd* mimic is situated below the adjacent base, as evidenced by the upfield-shifted G7H1′. The A·C pair of the *Azoarcus* 3 × 3 nt J4/5 loop is similar in conformation to the proposed conformations of G7·U15 in the *Ca/Cd* mimic. The A·C pair resembles the conformation of an A^+·C pair (*63*), which is isosteric with a standard wobble pair as well as G·U conformations with a slight shift from the standard wobble conformation.

Taken together, these data show that G·U pairs adjacent to noncanonical base pairs may be flexible and may not always prefer the standard wobble conformation. There are numerous examples in the literature of terminal G·U pairs that are in a conformation other than the standard wobble G·U pair (*46*, *64*−*68*). The Fox laboratory has compiled a database of non-Watson−Crick pairs including many types of G·U pairs (*69*, *70*).

Biological Implications. Many RNAs are being considered as targets for therapeutics (*71*−*76*). Designing such therapeutics requires an understanding of the interactions that are important for determining RNA structure, stability, flexibility, and tertiary contacts. This study shows that the tandem sheared A·A pairs in the *C. albicans* and *C. dubliniensis* J4/5 loops are likely to be similar to the tandem A·A pairs of *T. thermophila* and of one A·A pair in *Azoarcus* and may be similar to the tandem G·A pairs in *P. carinii*. The G·U pair not adjacent to the tandem A·A pairs in *C. albicans* and *C. dubliniensis* is likely similar in structure to the equivalent *P. carinii* G·U pair and slightly shifted in conformation relative to the *T. thermophila* G·U pair. Another G·U pair in *C. albicans* and *C. dubliniensis* is similar in dynamics to the equivalent *P. carinii* pair. The final G·U pair in *C. albicans* and *C. dubliniensis* is similar in dynamics to the extra loop nucleotide in *P. carinii* and similar in structure to the extra loop nucleotide in *T. thermophila* and *P. carinii*. Evidently, local hydrogen-bonding and stacking interactions shape this structure. Thus, it may be possible to predict and target this and similar sites on the basis of local structure and flexibility. Tandem sheared purine−purine pairs

may be an attractive motif to target with therapeutics. Loops with terminal G·U pairs may be good therapeutic targets due to the ability of the flexible nucleotides to structurally accommodate a variety of possible ligands. 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 (*60*, *61*, *77*). It is feasible that a therapeutic could selectively target functional groups in such loops and interrupt tertiary interactions, resulting in an inactive RNA.

SUPPORTING INFORMATION AVAILABLE

Two tables listing proton chemical shifts of the *Ca/Cd* mimic in a 3 mM $Co(NH_3)_6^{3+}$ solution and NMR-derived distance, sugar phosphate backbone torsion angle, and endocylic torsion angle restraints used in the simulations, one table summarizing the NMR-derived distance restraints, and ten figures showing the imino spectrum of the s-*Ca/Cd* mimic, the (H8/H6/H2)−(H1′/H5) region of the NOESY spectrum of the s-*Ca/Cd* mimic, the imino−amino region of the SNOESY spectrum of the *Ca/Cd* mimic, the H1′−sugar region of the TOCSY spectrum of the *Ca/Cd* mimic, the (H8/H6/H2)−(H1′/H5) region of the NOESY spectrum of 4I, the imino−amino region of the SNOESY spectrum of 4I, the (H8/H6/H2)−(H1′/H5) region of the NOESY spectra of 7I, 5/16P, and 6/17P, and an HSQC spectrum of the *Ca/Cd* mimic. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Fourmy, D., Recht, M. I., and Puglisi, J. D. (1998) Binding of neomycin-class aminoglycoside antibiotics to the A-site of 16 S rRNA, *J. Mol. Biol. 277*, 347−362.
2. Fourmy, D., Yoshizawa, S., and Puglisi, J. D. (1998) Paromomycin binding induces a local conformational change in the A-site of 16 S rRNA, *J. Mol. Biol. 277*, 333−345.
3. Woodcock, J., Moazed, D., Cannon, M., Davies, J., and Noller, H. F. (1991) Interaction of antibiotics with A-site-specific and P-site-specific bases in 16S ribosomal-RNA, *EMBO J. 10*, 3099−3103.
4. Moazed, D., and Noller, H. F. (1987) Interaction of antibiotics with functional sites in 16S ribosomal-RNA, *Nature 327*, 389−394.
5. Davies, J., and Davies, B. D. (1968) Misreading of ribonucleic acid code words induced by aminoglycoside antibiotics. The effect of drug concentration, *J. Biol. Chem. 243*, 3312−3316.
6. Schroeder, S. J., Burkard, M. E., and Turner, D. H. (2001) The energetics of small internal loops in RNA, *Biopolymers 52*, 157−167.
7. Lambowitz, A. M., and Belfort, M. (1993) Introns as mobile genetic elements, *Annu. Rev. Biochem. 62*, 587−622.
8. Cech, T. R. (1988) Conserved sequences and structures of group-I introns−Building an active-site for RNA catalysis−A review, *Gene 73*, 259−271.
9. Michel, F., and Westhof, E. (1990) Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis, *J. Mol. Biol. 216*, 585−610.
10. Damberger, S. H., and Gutell, R. R. (1994) A comparative database of group I intron structures, *Nucleic Acids Res. 22*, 3508−3510.
11. Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y. S., Feng, B., Lin, N., Madabusi, L. V., Muller, K. M., Pande, N., Shang, Z. D., Yu, N., and Gutell, R. R. (2002) The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs, *BMC Bioinf. 3*.
12. Strobel, S. A., and Ortoleva-Donnelly, L. (1999) A hydrogen-bonding triad stabilizes the chemical transition state of a group I ribozyme, *Chem. Biol. 6*, 153−165.
13. Cech, T. R., Tanner, N. K., Tinoco, I., Weir, B. R., Zuker, M., and Perlman, P. S. (1983) Secondary structure of the *Tetrahymena* ribosomal-RNA intervening sequence−Structural homology with fungal mitochondrial intervening sequences, *Proc. Natl. Acad. Sci. U.S.A. 80*, 3903−3907.
14. Ortoleva-Donnelly, L., Szewczak, A. A., Gutell, R. R., and Strobel, S. A. (1998) The chemical basis of adenosine conservation throughout the *Tetrahymena* ribozyme, *RNA 4*, 498−519.
15. Schaberg, D. R., Culver, D. H., and Gaynes, R. P. (1991) Major trends in the microbial etiology of nosocomial infection, *Am. J. Med. 91*, S72−S75.
16. Coleman, D. C., Sullivan, D. J., Bennett, D. E., Moran, G. P., Barry, H. J., and Shanley, D. B. (1997) Candidiasis: The emergence of a novel species, *Candida dubliniensis*, *AIDS 11*, 557−567.
17. Mercure, S., Montplaisir, S., and Lemay, G. (1993) Correlation between the presence of a self-splicing intron in the 25S RNA of *Candida-albicans* and strains susceptibility to 5-fluorocytosine, *Nucleic Acids Res. 21*, 6020−6027.
18. Nikolcheva, T., and Woodson, S. A. (1997) Association of a group I intron with its splice junction in 50S ribosomes: Implications for intron toxicity, *RNA 3*, 1016−1027.
19. Allain, F. H. T., and Varani, G. (1995) Divalent metal-ion binding to a conserved wobble pair defining the upstream site of cleavage of group-I self-splicing introns, *Nucleic Acids Res. 23*, 341−350.
20. Cate, J. H., and Doudna, J. A. (1996) Metal-binding sites in the major groove of a large ribozyme domain, *Structure 4*, 1221−1229.
21. Kieft, J. S., and Tinoco, I. (1997) Solution structure of a metal-binding site in the major groove of RNA complexed with cobalt(III) hexammine, *Structure 5*, 713−721.
22. Varani, G., and McClain, W. H. (2000) The G·U wobble base pair−A fundamental building block of RNA structure crucial to RNA function in diverse biological systems, *EMBO J. 1*, 18−23.
23. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) Crystal structure of a group I ribozyme domain: Principles of RNA packing, *Science 273*, 1678−1685.
24. Juneau, K., Podell, E., Harrington, D. J., and Cech, T. R. (2001) Structural basis of the enhanced stability of a mutant ribozyme domain and a detailed view of RNA-solvent interactions, *Structure 9*, 221−231.
25. Cowan, J. A. (1993) Metallobiochemistry of RNA−$Co(NH_3)_6^{3+}$ as a probe for Mg^{2+}(aq) binding-sites, *J. Inorg. Biochem. 49*, 171−175.
26. Usman, N., Ogilvie, K. K., Jiang, M. Y., and Cedergren, R. J. (1987) Automated chemical synthesis of long oligoribonucleotides using 2′-O-silylated ribonucleoside 3′-O-phosphoramidites on a controlled-pore glass support−synthesis of a 43-nucleotide sequence similar to the 3′-half molecule of an *Escherichia-coli* formylmethionine transfer-RNA, *J. Am. Chem. Soc. 109*, 7845−7854.
27. Wincott, F., DiRenzo, A., Shaffer, C., Grimm, S., Tracz, D., Workman, C., Sweedler, D., Gonzalez, C., Scaringe, S., and Usman, N. (1995) Synthesis, deprotection, analysis and purification of RNA and ribozymes, *Nucleic Acids Res. 23*, 2677−2684.
28. Stawinski, J., Stromberg, R., Thelin, M., and Westman, E. (1988) Evaluation of the use of the *tert*-butyldimethylsilyl group for 2′-protection in RNA-synthesis via the H-phosphonate approach, *Nucleosides Nucleotides 7*, 779−782.
29. Chou, S. H., Flynn, P., and Reid, B. (1989) Solid-phase synthesis and high-resolution NMR-studies of 2 synthetic double-helical RNA Dodecamers−r(CGCGAAUUCGCG) and r(CGCGUAU-ACGCG), *Biochemistry 28*, 2422−2435.
30. Petersheim, M., and Turner, D. H. (1983) Base-stacking and base-pairing contributions to helix stability: thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp, *Biochemistry 22*, 256−263.
31. McDowell, J. A., and Turner, D. H. (1996) Investigation of the structural basis for thermodynamic stabilities of tandem GU mismatches: Solution structure of $(rGAGGUCUC)_2$ by two-dimensional NMR and simulated annealing, *Biochemistry 35*, 14077−14089.
32. Borer, P. N., Dengler, B., Tinoco, I., and Uhlenbeck, O. C. (1974) Stability of ribonucleic-acid double-stranded helices, *J. Mol. Biol. 86*, 843−853.
33. Smallcombe, S. H. (1993) Solvent suppression with symmetrically-shifted pulses, *J. Am. Chem. Soc. 115*, 4776−4785.
34. Znosko, B. M., Burkard, M. E., Schroeder, S. J., Krugh, T. R., and Turner, D. H. (2002) Sheared $A_{(anti)}$·$A_{(anti)}$ base pairs in a

destabilizing 2x2 internal loop: the NMR structure of 5′-(rGGCAAGCCU)$_2$, *Biochemistry 41*, 14969−14977.
35. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules, *J. Am. Chem. Soc. 117*, 5179−5197.
36. Shi, S. H., Yan, L., Yang, Y., Fisher-Shaulsky, J., and Thacher, T. (2003) An extensible and systematic force field, ESFF, for molecular modeling of organic, inorganic, and organometallic systems, *J. Comput. Chem. 24*, 1059−1076.
37. Burkard, M. E., and Turner, D. H. (2000) NMR structures of r(GCAGGCGUGC)$_2$ and determinants of stability for single guanosine-guanosine base pairs, *Biochemistry 39*, 11748−11762.
38. Turner, D. H. (2000) in *Nucleic Acids: Structures, Properties, and Functions* (Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Eds.) University Science Books, Sausalito, CA.
39. Xia, T., SantaLucia, J., Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson−Crick base pairs, *Biochemistry 37*, 14719−14735.
40. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure, *J. Mol. Biol. 288*, 911−940.
41. Wu, M., McDowell, J. A., and Turner, D. H. (1995) A periodic table of symmetric tandem mismatches in RNA, *Biochemistry 34*, 3204−3211.
42. Schroeder, S. J., and Turner, D. H. (2001) Thermodynamic stabilities of internal loops with GU closing pairs in RNA, *Biochemistry 40*, 11509−11517.
43. Wijmenga, S. S., Mooren, M. M. W., and Hilbers, C. W. (1993) in *NMR of Macromolecules: A Practical Approach* (Roberts, G. C. K., Ed.) Oxford University Press, Oxford.
44. Varani, G., and Tinoco, I. (1991) RNA structure and NMR-spectroscopy, *Q. Rev. Biophys. 24*, 479−532.
45. Varani, G., Aboulela, F., and Allain, F. H. T. (1996) NMR investigation of RNA structure, *Prog. Nucl. Magn. Reson. Spectrosc. 29*, 51−127.
46. Varani, G., Cheong, C., and Tinoco, I., Jr. (1991) Structure of an unusually stable RNA hairpin, *Biochemistry 30*, 3280−3289.
47. Gorenstein, D. (1984) *^{31}P NMR, Principles and Applications*, Academic Press, New York.
48. Serra, M. J., Baird, J. D., Dale, T., Fey, B. L., Retatagos, K., and Westhof, E. (2002) Effects of magnesium ions on the stabilization of RNA oligomers of defined structures, *RNA 8*, 307−323.
49. Proctor, D. J., Schaak, J. E., Bevilacqua, J. M., Falzone, C. J., and Bevilacqua, P. C. (2002) Isolation and characterization of a family of stable RNA tetraloops with the motif YNMG that participate in tertiary interactions, *Biochemistry 41*, 12062−12075.
50. Shu, Z. Y., and Bevilacqua, P. C. (1999) Isolation and characterization of thermodynamically stable and unstable RNA hairpins from a triloop combinatorial library, *Biochemistry 38*, 15369−15379.
51. Klosterman, P. S., Shah, S. A., and Steitz, T. A. (1999) Crystal structures of two plasmid copy control related RNA duplexes: An 18 base pair duplex at 1.20 Å resolution and a 19 base pair duplex at 1.55 Å resolution, *Biochemistry 38*, 14784−14792.
52. SantaLucia, J., Jr., and Turner, D. H. (1993) Structure of (rGGCGAGCC)$_2$ in solution from NMR and restrained molecular dynamics, *Biochemistry 32*, 12612−12623.
53. Znosko, B. M., Burkard, M. E., Krugh, T. R., and Turner, D. H. (2002) Molecular recognition in purine-rich internal loops: Thermodynamic, structural, and dynamic consequences of purine for adenine substitutions in 5′(rGGCAAGCCU)$_2$, *Biochemistry 41*, 14978−14987.
54. Johnston, P. D., and Redfield, A. G. (1978) Pulsed FT-NMR double-resonance studies of yeast transfer RNAphe−Specific nuclear overhauser effects and reinterpretation of low-temperature relaxation data, *Nucleic Acids Res. 5*, 3913−3927.
55. Yang, J. H., and Gellman, S. H. (1998) Energetic superiority of two-center hydrogen bonding relative to three-center hydrogen bonding in a model system, *J. Am. Chem. Soc. 120*, 9090−9091.
56. Lan, T., and McLaughlin, L. W. (2001) The energetic contribution of a bifurcated hydrogen bond to the binding of DAPI to dA-dT rich sequences of DNA, *J. Am. Chem. Soc. 123*, 2064−2065.
57. Gonzalez, R. L., and Tinoco, I. (1999) Solution structure and thermodynamics of a divalent metal ion binding site in an RNA pseudoknot, *J. Mol. Biol. 289*, 1267−1282.
58. Colmenarejo, G., and Tinoco, I. (1999) Structure and thermodynamics of metal binding in the P5 helix of a group I intron ribozyme, *J. Mol. Biol. 290*, 119−135.
59. Schmitz, M., and Tinoco, I. (2000) Solution structure and metal-ion binding of the P4 element from bacterial RNase P RNA, *RNA 6*, 1212−1225.
60. Strauss-Soukup, J. K., and Strobel, S. A. (2000) A chemical phylogeny of group I introns based upon interference mapping of a bacterial ribozyme, *J. Mol. Biol. 302*, 339−358.
61. Strobel, S. A., Ortoleva-Donnelly, L., Ryder, S. P., Cate, J. H., and Moncoeur, E. (1998) Complementary sets of noncanonical base pairs mediate RNA helix packing in the group I intron active site, *Nat. Struct. Biol. 5*, 60−66.
62. Schroeder, S. J., Fountain, M. A., Kennedy, S. D., Lukavsky, P. J., Puglisi, J. D., Krugh, T. R., and Turner, D. H. (2003) Thermodynamic stability and structural features of the J4/5 loop in a *Pneumocystis carinii* group I intron, *Biochemistry 42*, 14184−14196.
63. Adams, P. L., Stahley, M. R., Kosek, A. B., Wang, J., and Strobel, S. A. (2004) Crystal structure of a self-splicing group I intron with both exons, *Nature 430*, 45−50.
64. Bullock, T. L., Sherlin, L. D., and Perona, J. J. (2000) Tertiary core rearrangements in a tight binding transfer RNA aptamer, *Nat. Struct. Biol. 7*, 497−504.
65. Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 angstrom resolution, *Science 289*, 905−920.
66. Basavappa, R., and Sigler, P. B. (1991) The 3 Å crystal structure of yeast initiator tRNA: functional implications in initiator/elongator discrimination, *EMBO J. 10*, 3105−3111.
67. Rath, V. L., Silvian, L. F., Beijer, B., Sproat, B. S., and Steitz, T. A. (1998) How glutaminyl-tRNA synthetase selects glutamine, *Struct. Folding Des. 6*, 439−449.
68. Molinaro, M., and Tinoco, I. (1995) Use of ultra-stable UNCG tetraloop hairpins to fold RNA structures−Thermodynamic and spectroscopic applications, *Nucleic Acids Res. 23*, 3056−3063.
69. Nagaswamy, U., Voss, N., Zhang, Z. D., and Fox, G. E. (2000) Database of non-canonical base pairs found in known RNA structures, *Nucleic Acids Res. 28*, 375−376.
70. Nagaswamy, U., Larios-Sanz, M., Hury, J., Collins, S., Zhang, Z. D., Zhao, Q., and Fox, G. E. (2002) NCIR: a database of non-canonical interactions in known RNA structures, *Nucleic Acids Res. 30*, 395−397.
71. Pearson, N. D., and Prescott, C. D. (1997) RNA as a drug target, *Chem. Biol. 4*, 409−414.
72. Hermann, T., and Westhof, E. (1998) RNA as a drug target: chemical, modelling, and evolutionary tools, *Curr. Opin. Biotechnol. 9*, 66−73.
73. Testa, S. M., Disney, M. D., Turner, D. H., and Kierzek, R. (1999) Thermodynamics of RNA-RNA duplexes with 2- or 4-thiouridines: implications for antisense design and targeting a group I intron, *Biochemistry 38*, 16655−16662.
74. Disney, M. D., Testa, S. M., and Turner, D. H. (2000) Targeting a *Pneumocystis carinii* group I intron with methylphosphonate oligonucleotides: backbone charge is not required for binding or reactivity, *Biochemistry 39*, 6991−7000.
75. Lukavsky, P. J., Otto, G. A., Lancaster, A. M., Sarnow, P., and Puglisi, J. D. (2000) Structures of two RNA domains essential for hepatitis C virus internal ribosome entry site function, *Nat. Struct. Biol. 7*, 1105−1110.
76. Gallego, J., and Varani, G. (2001) Targeting RNA with small-molecule drugs: Therapeutic promise and chemical challenges, *Acc. Chem. Res. 34*, 836−843.
77. Lehnert, V., Jaeger, L., Michel, F., and Westhof, E. (1996) New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: a complete 3D model of the *Tetrahymena thermophila* ribozyme, *Chem. Biol. 3*, 993−1009.
78. Carter, A. P., Clemons, W. M., Brodersen, D. E., Morgan-Warren, R. J., Wimberly, B. T., and Ramakrishnan, V. (2000) Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics, *Nature 407*, 340−348.
79. Collier, A. J., Gallego, J., Klinck, R., Cole, P. T., Harris, S. J., Harrison, G. P., Aboul-ela, F., Varani, G., and Walker, S. (2002) A conserved RNA structure within the HCVIRES eIF3-binding site, *Nat. Struct. Biol. 9*, 375−380.

80. Fukai, S., Nureki, O., Sekine, S., Shimada, A., Tao, J. S., Vassylyev, D. G., and Yokoyama, S. (2000) Structural basis for double-sieve discrimination of L-valine from L-isoleucine and L-threonine by the complex of tRNA(Val) and valyl-tRNA synthetase, *Cell 103*, 793−803.

81. Kitamura, A., Muto, Y., Watanabe, S., Kim, I., Ito, T., Nishiya, Y., Sakamoto, K., Ohtsuki, T., Kawai, G., Watanabe, K., Hosono, K., Takaku, H., Katoh, E., Yamazaki, T., Inoue, T., and Yokoyama, S. (2002) Solution structure of an RNA fragment with the P7/P9.0 region and the 3′-terminal guanosine of the *Tetrahymena* group I intron, *RNA 8*, 440−451.

82. Nikulin, A., Serganov, A., Ennifar, E., Tishchenko, S., Nevskaya, N., Shepard, W., Portier, C., Garber, M., Ehresmann, B., Ehresmann, C., Nikonov, S., and Dumas, P. (2000) Crystal structure of the S15-rRNA complex, *Nat. Struct. Biol. 7*, 273−277.

83. Jiang, L. C., and Patel, D. J. (1998) Solution structure of the tobramycin-RNA aptamer complex, *Nat. Struct. Biol. 5*, 769−774.

84. Leontis, N. B., and Westhof, E. (2001) Geometric nomenclature and classification of RNA base pairs, *RNA 7*, 499−512.

85. Lu, X. J., Shakked, Z., and Olson, W. K. (2000) A-form conformational motifs in ligand-bound DNA structures, *J. Mol. Biol. 300*, 819−840.

BI049256Y