BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

The Nuclear Magnetic Resonance of CCCC RNA Reveals a Right-Handed Helix, and Revised Parameters for AMBER Force Field Torsions Improve Structural Predictions from Molecular Dynamics

Jason D. Tubbs,[†,‡] David E. Condon,[†,‡] Scott D. Kennedy,[‡,§] Melanie Hauser,[⊥,∥] Philip C. Bevilacqua,[⊥] and Douglas H. Turner*[,†,‡]

[†]Department of Chemistry, University of Rochester, Rochester, New York 14627, United States

[‡]Center for RNA Biology, University of Rochester, Rochester, New York 14642, United States

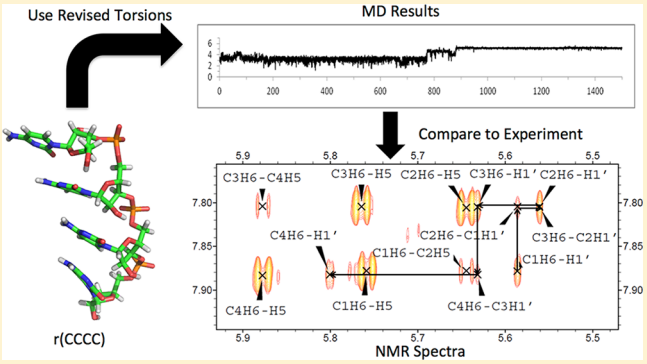
[§]Department of Biochemistry and Biophysics, University of Rochester, Rochester, New York 14642, United States

[∥]Department of Chemistry, Buena Vista University, Storm Lake, Iowa 50588, United States

[⊥]Department of Chemistry and Center for RNA Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania 16802, United States

Supporting Information

ABSTRACT: The sequence dependence of RNA energetics is important for predicting RNA structure. Hairpins with C_n loops are consistently less stable than hairpins with other loops, which suggests the structure of C_n regions could be unusual in the "unfolded" state. For example, previous nuclear magnetic resonance (NMR) evidence suggested that polycytidylic acid forms a left-handed helix. In this study, UV melting experiments show that the hairpin formed by r(5′GGACCCCCGUCC) is less stable than r(5′GGACUUUUGUCC). NMR spectra for single-stranded C_4 oligonucleotide, mimicking the unfolded hairpin loop, are consistent with a right-handed A-form-like helix. Comparisons between NMR spectra and molecular dynamics (MD) simulations suggest that recent reparametrizations, parm99χ_YIL and parm99TOR, of the AMBER parm99 force field improve the agreement between structural features for C_4 determined by NMR and predicted by MD. Evidently, the force field revisions to parm99 improve the modeling of RNA energetics and therefore structure.



Ribonucleic acids (RNA) are important biological molecules that have a variety of functions, including catalysis,[1,2] regulation of gene expression,[3,4] and use as a template for reverse transcription.[5] The function of RNA is often dictated by its structure, and loops are an important component of structure.

The secondary structure of an RNA can be predicted with thermodynamic parameters from the nearest neighbor model, and predictions are reasonably accurate when constrained by experimental data.[6−8] UV melting experiments with RNAs have revealed a variety of stabilizing and destabilizing structures. One destabilizing structure is a hairpin loop composed of only cytidine. Groebe and Uhlenbeck showed that hairpin loops composed of only cytidine are more destabilizing that the respective adenosine and uridine hairpin loops.[9] This led to a penalty in nearest neighbor parameters for homocytidine loops.[6,10] One possible reason for the destabilizing effect is that single-stranded oligocytidylic acid is unusually structured, making folding into the hairpin more thermodynamically unfavorable.

Polycytidylic acid (polyC) has been studied by numerous techniques.[11−32] An X-ray fiber diffraction study revealed that at neutral pH, polyC is a right-handed helix with a C3′-endo sugar pucker and base stacking.[11] A later nuclear magnetic resonance (NMR) study concluded that polyC in solution is a left-handed helix without base stacking and stabilized by direct hydrogen bonding between the carbonyl and amino groups of adjacent bases.[13] Temperature jump experiments showed that stacking and unstacking in polyC and CpC occur on the time scale of 100 ns.[30−32]

In this paper, UV melting shows that the hairpin formed by r(5′GGACCCCCGUCC), termed HPC_4, is less stable than r(5′GGACUUUUGUCC), termed HPU_4. NMR spectra reveal that r(CCCC) has an A-form-like folded-state structure with C3′-endo sugar puckers, an anti orientation of bases relative to their sugars, and base−base stacking. The NMR data provide

Received: July 31, 2012
Revised: December 27, 2012
Published: January 3, 2013

benchmarks for testing recent revisions to torsion parameters[33,34] in the parm99[35,36] force field, which is commonly used for molecular dynamics (MD) simulations of RNA. For this purpose, the NMR results are compared to predictions from 1500 ns MD simulations with parm99,[35,36] parm99χ_YIL,[33] and parm99TOR[34] force fields. The rapid rate of conformational changes allows several transitions during the simulation time, so that simulations can be started far from the expected average structure and have the possibility of finding structures consistent with NMR spectra. The results indicate that revisions in parm99χ_YIL[33] and parm99TOR[34] force fields improve the modeling of RNA, but further revisions are necessary.

MATERIALS AND METHODS

UV Melting. The oligoribonucleotides r(5′GGACCCCCGUCC) and r(5′GGACUUUUGUCC) were purchased from Dharmacon Inc. and deblocked according to the manufacturer's instructions. The RNAs were then purified by PAGE on a 20% gel.

Thermodynamic parameters were measured in 10 mM MOPS, 0.1 mM EDTA (pH 7.0), and a salt concentration varying between 5 and 2000 mM for NaCl, 10 and 1000 mM for KCl and LiCl, or 2 and 20 mM for $MgCl_2$. Absorbance versus temperature melting curves were measured at 260, 270, and 280 nm with a heating rate of 1 °C/min, forward and reverse, from 5 to 99 °C on a Gilford Response spectrophotometer. Typical melting curves were measured over a 10-fold range of oligonucleotide concentrations. Melting temperatures were independent of concentration, consistent with hairpin formation.

Thermodynamic Analysis of Hairpin Formation. Thermodynamic parameters were obtained using MeltWin,[37] which assumes a two-state model. Melting curves were fit to the following equations:

$$A = fA_F + (1 - f)A_U \quad (1)$$

$$A_F = \varepsilon_F C_F L = (m_F T + b_F) C_F L \quad (2)$$

$$A_U = \varepsilon_U C_U L = (m_U T + b_U) C_U L \quad (3)$$

$$K = \exp\left(-\frac{\Delta H^\circ}{RT} + \frac{\Delta S^\circ}{R}\right) \quad (4)$$

where A is the absorbance, f is the fraction of strands in a hairpin, A_F is the absorbance of folded species, A_U is the absorbance of unfolded species, ε_F and ε_U are the extinction coefficients for the folded and unfolded species, respectively, C_F and C_U are the concentrations of the folded and unfolded species, respectively, L is the path length of the cell, m_F and m_U are the slopes for the temperature dependence of extinction coefficients of the folded and unfolded species, respectively, b_F and b_U are the corresponding intercepts, ΔH° is the standard enthalpy change, and ΔS° is the standard entropy change. The six parameters fit are b_F, b_U, m_F, m_U, ΔH°, and ΔS°.

NMR. r(CCCC) was purchased from Dharmacon Inc. The NMR sample consisted of 3.2 mM r(CCCC) in 80 mM NaCl, 10 mM sodium phosphate (pH 7.0), and 0.5 mM EDTA. Two deuterium exchanges with 99.9% D_2O (Cambridge Isotopes Laboratories) were performed on the sample followed by a final exchange with 99.990% D_2O (Sigma Aldrich).

Nonexchangeable proton assignments were made from two-dimensional (2D) NOESY, $^1H-^{31}P$ HETCOR, $^{13}C-^1H$ HMQC, and TOCSY spectra collected at 5 °C with a Varian Inova 600 MHz NMR spectrometer (Table S1 of the Supporting Information). NOESY spectra were recorded with mixing times of 200, 400, and 800 ms and were ^{31}P-decoupled. Acquisition parameters are listed in Table S2 of the Supporting Information. NMR spectra were processed with NMRPipe[38] and analyzed with SPARKY.[39] Chemical shifts for the H5′ and H5″ protons were assigned by assuming the higher and lower chemical shifts are H5′ and H5″ resonances, respectively. The H5′ and H5″ labels are those used in AMBER.[35,36]

Generation of Starting Structures for Molecular Dynamics Simulations of r(CCCC). Starting structures of r(CCCC) for MD were generated with the nucgen program in AMBER9.[40] Four different starting structures were used: A-form (C3′-endo/anti), C3′-endo/syn, C2′-endo/anti, and C2′-endo/syn, where C3′-endo and C2′-endo represent possible sugar puckers and anti and syn represent possible glycosydic bond torsions. The C3′-endo/syn structure was generated with Avogadro 1.1.0[41] by manually rotating the four torsions to an approximate syn value (Table S3 of the Supporting Information). C2′-endo/anti and C2′-endo/syn structures were generated by simulated annealing. Sufficient simulated annealing calculations were performed so that the structure would fit the desired torsion and pseudorotation phase angle constraints for each starting structure, with each succeeding simulation starting from the previous one. The generalized Born implicit solvent model[42−44] with a salt concentration of 1 M was used in the simulated annealing protocol. A 20 Å cutoff for long-range nonbonded interactions was used. The velocity limit was set to 10. Berendsen temperature control[45] with a coupling time of 1 ps was utilized in all simulated annealing calculations. Distance and torsion restraints of 32 kcal mol^{-1} $Å^{-2}$ and 32 kcal mol^{-1} rad^{-2}, respectively, were applied with square bottom wells with parabolic sides, which forced the structure into the intended starting structure. Starting backbone torsions measured with Pymol[46] are presented in Table S3 of the Supporting Information. Torsion angle definitions were taken from Murray et al.[47] Figure 1 shows the atoms defining torsions. AMBER input files and starting structural coordinates (PDB format) are provided in the Supporting Information (Tables S4−S11).

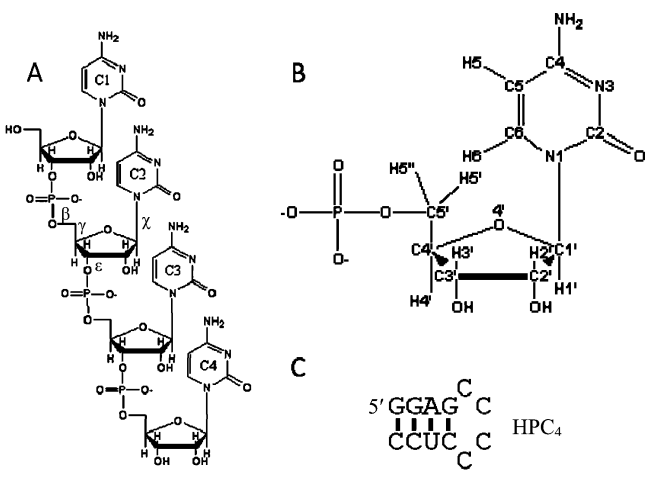


Figure 1. (A) Single-stranded r(CCCC) with the β, γ, ε, and χ torsion angles labeled. (B) Atom notation used in cytidine and D-ribose. (C) Hairpin formed by r(5′GGACCCCCGUCC), termed HPC_4.

dx.doi.org/10.1021/bi3010347 | Biochemistry 2013, 52, 996−1010

Table 1. Thermodynamic Parameters Calculated from UV Melting in NaCl, KCl, and LiCl at pH 7.0 Buffered with 10 mM MOPS and 0.1 mM EDTA

hairpin	salt	T_M (°C)[a]	$\Delta H°$ (kcal/mol)	$\Delta S°$ (eu)	$\Delta G°_{37}$ (kcal/mol)[a,b]	$\Delta G°_{37,HL}$ (kcal/mol)[c]
r(5′GGACCCCCGUCC)	1 M NaCl	61.9	−31.8	−94.9	−2.37	6.48
	1 M KCl	62.7	−30.6	−91.1	−2.35	6.50
	1 M LiCl	64.8	−31.5	−93.2	−2.59	6.26
r(5′GGACUUUUGUCC)	1 M NaCl	67.9	−42.7	−125.2	−3.87	5.18
	1 M KCl	66.1	−40.3	−118.8	−3.45	5.60
	1 M LiCl	68.6	−41.8	−122.3	−3.87	5.18
r(5′GGACCCCCGUCC)	5 mM NaCl	57.3	−30.3	−91.7	−1.86	–
	10 mM KCl	57.1	−32.4	−98.1	−1.97	–
	10 mM LiCl	58.5	−32.4	−97.7	−2.10	–
r(5′GGACUUUUGUCC)	5 mM NaCl	58.8	−38.7	−116.6	−2.54	–
	10 mM KCl	61.6	−39.3	−117.4	−2.89	–
	10 mM LiCl	60.2	−38.5	−115.5	−2.68	–

[a] T_m and $\Delta G°_{37}$ were calculated after averaging the $\Delta H°$ and $\Delta S°$ results. [b] Folding free energy for the stem−loop structure. [c] The hairpin loop component of the folding free energy. INN−HB thermodynamic parameters were measured in 1 M NaCl[59,60] and are not applicable to 5 mM NaCl, 10 mM KCl, or 10 mM LiCl.

Molecular Dynamics Simulations of r(CCCC). AMBER9's tleap program was used to generate starting structures neutralized with three Na^+ ions[40,48] and solvated in an 8.65 Å truncated octahedral box with 1215 TIP3P water molecules.[49] The system was minimized by steepest descent for 500 steps and then minimized by conjugate gradient descent for another 500 steps. Periodic boundaries were used, and Cartesian restraints were activated. Long-range interactions were not calculated past 10 Å. To keep the RNA immobile, a 500 kcal/mol restraint was placed on all residues. After minimization of the solvent, the entire system was minimized by steepest descent for 1000 steps and then minimized by conjugate gradient descent for 1500 steps, which were otherwise identical to those of the first minimization routine.

The RNA was then kept fixed for a 2 ns MD run using a 2 fs time step. Minimization was turned off; nonbonded interactions beyond 10 Å were neglected, and periodic boundaries were activated. The initial temperature was set to 0 K, so the velocities were calculated from the forces. The temperature was gradually increased from 0 to 275 K through 100000 steps. Langevin dynamics with a collision frequency of 1.0 ps^{-1} was used. The system was again equilibrated for 100 ps at 275 K. After minimization, the non-A-form starting structures had already moved toward A-form.

Production runs were conducted with a 3 fs time step at 278 K using particle mesh Ewald (PME)[50,51] with a 10 Å nonbonded cutoff, a constant pressure, isotropic position scaling, a 2 ps pressure relaxation time, and no position restraints. To test if energy is conserved with a 3 fs time step, 2 and 3 fs time steps were simulated for 100 ns. There were no anomalous energy spikes observed in either simulation (Figure S1 of the Supporting Information). SHAKE[52] was used to provide constraints of bonds involving hydrogen. All simulations were run for 1500 ns, which allows for several conformational changes. Restart files were saved every 0.3 ns, and the trajectory was written every 0.15 ns. The same random seed number generator was used for each restart. To counteract potential problems, identified by Cerutti et al.,[53] the simulations were restarted at long (~100 ns on average), uneven time intervals. For comparison to NMR results, the MD results were averaged from these trajectories.

Calculation of Experimental Values. The ^3J scalar coupling values were measured from peak splittings in the NMR spectra. An error of ±0.5 Hz was assumed for scalar couplings of >1.0 Hz. Distances were calculated from measured NOE volumes using eq 5.

$$r_{NOE} = \left(\frac{c}{NOE_{ij}}\right)^{1/6} \quad (5)$$

where r_{NOE} is the NMR distance between protons i and j, NOE_{ij} is the volume for an NOE peak between protons i and j, and c is an average constant determined from NOE_{ij} H1′−H2′ cross-peaks of C1 and C3 and H1′−H4′ cross-peaks of C2 and C3 assuming corresponding distances of 2.75 and 3.29 Å, respectively. The H1′−H2′ and H1′−H4′ distances are assumed to be consistent because there is no observable splitting of the C1−C3 H1′ resonances (Figure S2 of the Supporting Information), indicating that C1−C3 are in the C3′-endo sugar conformation. The more commonly used H5−H6 distance was not used as a reference because of zero quantum coherence, which results in reduction of NOE peak volumes, sometimes even leading to negative NOE volumes.

The 200 ms mixing time NOESY spectrum was used for all NMR cross-peak volume measurements. Because r(CCCC) is small, 200 ms includes the linear range for NOE buildup. Comparison of NOE cross-peak volumes for the 200 and 400 ms mixing time NOESY spectra (Table S12 of the Supporting Information) gave a ratio of 1.8, consistent with 200 ms being in the linear buildup range. Error limits on NMR distances were calculated from

$$\text{NMR upper limit:} \quad r_{upper} = \left(\frac{c + 2c_{SD}}{NOE_{ij} - 2V_{err}}\right)^{1/6} \quad (6)$$

$$\text{NMR lower limit:} \quad r_{lower} = \left(\frac{c - 2c_{SD}}{NOE_{ij} + 2V_{err}}\right)^{1/6} \quad (7)$$

where c_{SD} is the standard deviation in the average c and V_{err} is the standard deviation calculated from 12 volumes of spectral noise (areas with no NOEs). Table S12 of the Supporting Information contains the peak volumes, average c values, c_{SD} values, noise volumes, and the noise standard deviation (V_{err}) used in this paper.

^3J scalar coupling values predicted from MD were calculated from torsion angle trajectories. Each torsion angle was converted to a ^3J scalar coupling using equations from Marino et al. (Table S13 of the Supporting Information)[54] and then averaged to give the MD-predicted ^3J scalar coupling. The average "distances" predicted by MD, r_{MD}, were calculated from interproton distance trajectories:

$$r_{MD} = \left(\frac{1}{n} \sum_{i=1}^{n} \frac{1}{r_i^6} \right)^{-1/6} \quad (8)$$

where r_{MD} is the distance between two protons averaged over n snapshots, i, of the trajectory.

C3′-endo and C2′-endo conformations predicted by MD were defined by pseudorotation phase angle ranges of 0−36° and 144−180°, respectively.[55] Anti, high anti, and syn conformations were defined by torsion angle, χ, ranges of 180−239°, 240−300°, and 0−120°, respectively.[56]

A-Form distance values are the average of distances measured from four CC doublets in the A-form duplex structure, COPY19 RNA, entry 1QC0[57] as deposited in the Protein Data Bank (PDB).[58] The A-form torsions were obtained in a similar fashion.

RESULTS

Thermodynamics of C_4 and U_4 Hairpin Formation. Table 1 lists thermodynamic parameters for hairpin formation by HPC_4 and HPU_4 in 1 M and 5 mM NaCl and 1 M and 10 mM KCl and LiCl. Table 1 also lists the free energy increments, $\Delta G^\circ_{37,HL}$, of the hairpin loop closure at a cation concentration of 1 M under the assumption that the nearest neighbor parameters for Watson−Crick helices[59] are the same for Na^+, K^+, and Li^+.

For example

$$\Delta G^\circ_{37,HL}(C_4) = \Delta G^\circ_{37}(5'\text{GGAC}\underline{\text{CCCC}}\text{GUCC}) - \Delta G^\circ_{37}(5'\text{GG}/3'\text{CC}) - \Delta G^\circ_{37}(5'\text{GA}/3'\text{CU}) - \Delta G^\circ_{37}(5'\text{AC}/3'\text{UG}) - \Delta G^\circ_{37}(5'\text{CC}/3'\text{GC}) = -2.37 + 3.26 + 2.35 + 2.24 + 1.0 = 6.48 \text{ kcal/mol}$$

where $\Delta G^\circ_{37}(5'\text{GGAC}\underline{\text{CCCC}}\text{GUCC})$ is the measured folding free energy change for the hairpin in 1 M NaCl and the other ΔG°_{37} values are nearest neighbor values.[59,60]

The $\Delta\Delta G^\circ_{37}$ values between C_4 and U_4 hairpins with stems of four Watson−Crick base pairs are 1.50, 1.10, and 1.28 kcal/mol in 1 M NaCl, KCl, and LiCl, respectively, and 0.68, 0.92, and 0.58 kcal/mol in 5 mM NaCl, 10 mM KCl, and 10 mM LiCl, respectively. The Supporting Information contains melting data for a wide range of salt conditions (Tables S14 and S15 of the Supporting Information). Values of ΔG°_{37} in 2−20 mM Mg^{2+} are similar to those in 1 M NaCl.

NMR. $^3J_{1'-2'}$ scalar coupling constants from one-dimensional (1D) ^{1}H NMR spectra of r(CCCC) revealed that all residues prefer a C3′-endo sugar pucker from 0 to 70 °C (Tables 2 and 3 and Tables S16 and S17 and Figure S2 of the Supporting Information). Figure 2 shows the NOESY walk region from a 2D ^{1}H−^{1}H spectrum. The intranucleoside H6−H1′ cross-peaks for C1, C2, and C4 are weak compared to the intranucleoside H6−H5 cross-peaks, implying that the bases are anti relative to the ribose (Tables 2 and 4). In addition, the volume of the peak containing overlapped cross-peaks from H6 to H1′ of C3 and from H6 to H5 of C2 is also smaller than expected if C3 was syn relative to the ribose. NOEs between C2 H6 and C1 H2′, C3 H6 and C2 H2′, C4 H6 and C3 H2′, and C4 H6 and C3 H3′ (Figure 3) are consistent with A-form base stacking. Because of the overlap between C3 H6 and C2 H6 and between C2 H2′ and C2 H3′, however, the stacking between C2 and C3 cannot be evaluated quantitatively. Chemical shifts of the phosphorus resonances are all within 0.3 ppm, indicating that there is no significant deviation from accepted A-form RNA backbone torsions (Figure S3 of the Supporting Information).

Table 2. Percentages at 278 K of C3′-Endo Ribose and Anti Glycosidic Angles from NMR and Predicted from 1500 ns of MD Starting with C3′-Endo/Anti Starting Structures

	% C3′-Endo[a]			
residue	NMR[b]	parm99	parm99χ_YIL	parm99TOR
C1	90−100	19.5	70.3	69.4
C2	90−100	32.0	83.6	79.3
C3	90−100	40.8	82.5	89.7
C4	70−80	15.3	72.2	61.9
	% Anti			
residue	NMR	parm99[c]	parm99χ_YIL	parm99TOR
C1	90−100	51.0	99.9	100.0
C2	90−100	93.8	100.0	100.0
C3	90−100	99.0	100.0	100.0
C4	90−100	94.1	99.9	99.4

[a]In MD simulations, ribose was considered C2′-endo or C3′-endo if the pseudorotation phase angle was between 144° and 180° or between 0° and 36°, respectively. [b]C3′-endo ranges were approximated from the H1′−H2′ scalar couplings, assuming that C3′-endo and C2′-endo have scalar couplings of 0 and 10 Hz, respectively. [c]Includes high anti in the percentage. Anti, high anti, and syn are defined as having O4′−C1′−N1−C2 dihedral angles of 180−239°, 240−300°, and 0−120°, respectively.

In a 200 ms NOESY spectrum, C3 H6−C4 H2′ and C3 H3′−C4 H2′ cross-peaks correspond to distances of 4.00 and 3.32 Å, respectively, which suggests a population with the C4 sugar inverted because A-form distances[57] would range from 7.91 to 9.15 Å and from 6.38 to 7.62 Å, respectively (Table 4). Sugar inversion has been observed previously[56,61−65] but is rare. An approximate percentage of terminal sugar inversion can be calculated by comparing the cross-peak volumes of C3 H2′−C4 H6 and C1 H2′−C2 H6 cross-peaks because sugar inversion increases the distance between Cn H2′ and Cn+1 H6 to >5 Å. If the cross-peak volume for the C3 H2′−C4 H6 cross-peak is assumed to be zero when the sugar is inverted, then the population of inverted terminal sugars is approximately 13% at 5 °C.

Broido and Kearns proposed a left-handed helix for polyC caused by a large H5−H1′ NOE corresponding to a distance of 2.3 Å.[13] Figure S4 of the Supporting Information shows that C_4 has no significant peaks corresponding to H5−H1′ NOEs in a 200 ms NOESY ^{31}P-decoupled spectrum.

Dynamics of the Initial A-Form r(CCCC) Structure in MD Simulations with the parm99, parm99χ_YIL, and parm99TOR Force Fields. Figures S5 and S6 of the Supporting Information and Figure 4 show 1500 ns trajectories for the heavy atoms and backbone rmsds, χ dihedrals, and δ dihedrals, respectively, for all cytidines in the parm99,[35,36] parm99χ_YIL,[33] and parm99TOR[34] force field simulations with an A-form starting structure. The 1500 ns simulation with

dx.doi.org/10.1021/bi3010347 | Biochemistry 2013, 52, 996−1010

Table 3. Measured NMR and MD-Predicted ^3J Couplings (hertz) at 278 K for Each Force Field with the A-Form (C3′-endo/anti) Starting Structure

torsion	NMR scalar coupling[a]	parm99[b]	parm99χ[b]	parm99TOR[b,c]	A-form[d]	angle (deg)[e]
C1 H1′−H2′	≤1.0	5.6	0.7	0.5 (0.6)	0.0−1.0	91−104
C2 H1′−H2′	≤1.0	4.6	0.5	0.5 (0.7)	0.0−1.0	91−104
C3 H1′−H2′	≤1.0	3.6	0.6	0.5 (0.6)	0.0−1.0	91−104
C4 H1′−H2′	2.6	6.7	1.5	1.8 (1.9)	0.0−1.0	116
C1 H2′−H3′	3.6	5.1	4.4	4.3 (4.3)	3.8−4.8	50
C2 H2′−H3′	4.3	5.0	4.4	4.4 (4.3)	3.8−4.8	42
C3 H2′−H3′	4.7	5.1	4.5	4.4 (4.4)	3.8−4.8	38
C4 H2′−H3′	4.9	5.1	4.7	4.9 (4.8)	3.8−4.8	36
C1 H3′−H4′	8.7	5.2	7.7	7.8 (7.8)	7.6−8.6	*f*
C3 H3′−H4′	8.8	5.8	7.7	7.9 (7.9)	7.6−8.6	*f*
C4 H3′−H4′	7.2	3.6	7.1	6.7 (7.0)	7.6−8.6	205
C2 H4′−H5′	≤1.0	2.5	2.6	2.8 (3.0)	1.6−2.6	*f*
C3 H4′−H5′	2.1	2.3	2.5	2.8 (2.8)	1.6−2.6	63
C4 H4′−H5′	2.0	2.3	2.5	3.7 (2.9)	1.6−2.6	62
C2 H4′−H5″	≤1.0	4.9	3.9	3.2 (3.1)	2.9−4.4	*f*
C3 H4′−H5″	≤1.0	3.8	3.5	3.1 (3.2)	2.9−4.4	*f*
C4 H4′−H5″	1.4	4.3	3.8	3.4 (3.5)	2.9−4.4	292
C1 H3′−C2 P	8.8	5.1	8.7	7.1 (8.3)	6.9−9.4	1
C2 H3′−C3 P	9.3	5.6	9.0	3.8 (5.2)	6.9−9.4	3
C3 H3′−C4 P	9.3	5.8	8.6	4.7 (5.0)	6.9−9.4	3
C2 P−C2 H5′	3.8	3.1	3.8	4.2 (5.0)	3.4−6.6	309
C3 P−C3 H5′	3.9	3.8	3.8	3.7 (4.1)	3.4−6.6	309
C4 P−C4 H5′	3.8	2.7	3.6	6.2 (4.3)	3.4−6.6	309
C2 P−C2 H5″	1.2	3.0	2.3	2.6 (2.4)	0.9−1.9	72
C3 P−C3 H5″	0.5	2.4	2.2	2.5 (2.2)	0.9−1.9	*f*
C4 P−C4 H5″	1.1	3.3	2.4	4.2 (3.7)	0.9−1.9	74

[a]The error limit is ±0.5 Hz for scalar couplings of >1.0 Hz. [b]Marino et al.[54] equations (Table S13 of the Supporting Information) were used to calculate ^3J couplings for each time point of the MD trajectory, and then the couplings were averaged. [c]Parentheses indicate MD predicted ^3J scalar coupling values calculated from 0 to 770 ns. After 770 ns, C1 intercalates between C3 and C4. [d]Range of values measured from four CC doublets in PDB entry 1QC0.[57] [e]Torsion angles determined via NMR (Table S13 of the Supporting Information). For scalar couplings of <1.0 Hz, an angle range covering 0−1 Hz is shown. [f]Could not be calculated with Marino et al.[54] equations (Table S13 of the Supporting Information).

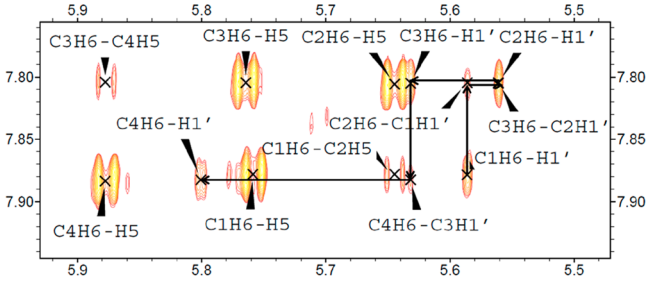


Figure 2. NOESY walk region from an 800 ms mixing time ^{31}P-decoupled spectrum. Arrows indicate the direction of the walk from the 5′ to 3′ end. The C3 H6−C2 H1′ cross-peak overlaps with the C2 H6−C2 H1′ resonance.

the parm99 force field shows rapid fluctuations (Figure S5 of the Supporting Information). The χ torsion for C1 in parm99 makes frequent transitions between syn (~45°) and high anti (~265°). C2−C4 have a few transitions to syn but stay high anti for most of the simulation (Table 2). For all the cytidines, the δ dihedral shows frequent transitions between C2′-endo (140−152°) and C3′-endo (78−90°) (Table 2 and Figure S5 of the Supporting Information).

With the parm99χ_YIL force field simulation (Table 2 and Figure S6 of the Supporting Information), C1−C4 are anti (180−239°), with a few transitions to high anti. C1−C3 are largely C3′-endo, with a few C2′-endo transitions, but C4 has rapid fluctuations between C2′-endo and C3′-endo (Table 2 and Figure S6 of the Supporting Information).

With the parm99TOR force field simulation, C1 intercalates between C3 and C4 after 770 ns (Table 5 and Figures 4 and 5) and fluctuations in the rmsd plots significantly decrease. The intercalation is stabilized by C1 base stacking with C3 and C4. There is no NMR evidence of this conformation (Figure 5). parm99TOR, with an A-form starting structure, was thus evaluated for the first 770 ns in addition to the whole trajectory (Tables 3 and 4 and Figure 4). C1 is anti and C3′-endo with a few fluctuations to high anti and C2′-endo; C2 and C3 are anti and C3′-endo without fluctuations, and C4 is anti, with a few syn and high anti transitions, and has rapid fluctuations between C2′-endo and C3′-endo. Interestingly, the intercalation of C1 between C3 and C4 after 770 ns is not accompanied by any clear changes in χ or δ (Figure 4).

Dynamics of Non-A-Form r(CCCC) in MD Simulations with the parm99, parm99χ_YIL, and parm99TOR Force Fields. Figure 6 shows the heavy atom rmsd trajectories relative to an A-form (C3′-endo/anti) structure for production runs with C3′-endo/syn, C2′-endo/anti, and C2′-endo/syn starting structures and parm99, parm99χ_YIL, and parm99-TOR force fields. The evaluation of these simulations was started at the first point at which the simulations show an A-from-like structure as defined by reasonable agreement with NMR distances (Table 4) and expectations for torsion angles based on ranges[61] observed in a crystal structure (PDB entry

dx.doi.org/10.1021/bi3010347 | *Biochemistry* 2013, 52, 996−1010

Table 4. Distances (angstroms) Determined from NMR NOEs in a 200 ms NOESY Spectrum Measured at 278 K and Predicted via MD with the C3′-Endo/Anti Starting Structures

cross-peak	NMR distance	NMR lower limit	NMR upper limit	parm99	parm99χ	parm99TOR[a]	A-form[b]
C1 H2′−C2 H1′	4.21	3.77	4.61	5.35	4.20	4.20 (4.01)	3.51−4.75
C1 H2′−C2 H5′	2.82	2.61	2.98	3.97	2.50	2.50 (2.38)	2.38−3.62
C1 H5−C2 H5	3.79	3.45	4.07	4.52	3.44	3.97 (3.55)	3.18−4.42
C1 H6−C1 H1′	3.46	3.17	3.68	2.54	3.50	3.52 (3.50)	2.82−4.06
C1 H6−C1 H2′	3.38	3.10	3.60	2.72	3.96	3.93 (3.95)	3.18−4.42
C1 H6−C2 H5	3.66	3.34	3.91	4.27	3.44	3.77 (3.38)	3.56−4.80
C2 H5−C1 H2′	4.43	3.94	4.94	3.45	4.33	4.33 (4.72)	3.13−4.35
C2 H5−C3 H5	3.80	3.46	4.08	4.15	3.28	3.59 (3.58)	3.18−4.42
C2 H6−C1 H2′	2.73	2.52	2.88	3.18	2.63	3.28 (2.94)	1.66−2.90
C2 H6−C2 H1′	3.43	3.14	3.65	3.24	3.57	3.53 (3.56)	2.82−4.06
C2 H6−C2 H5″	3.49	3.20	3.72	3.81	3.43	4.04 (3.91)	3.46−4.70
C3 H3′−C4 H2′	3.32	3.05	3.52	2.99	4.44	3.96 (3.62)	6.38−7.62
C3 H5−C4 H5	3.67	3.35	3.93	4.60	3.39	4.25 (3.88)	3.18−4.42
C3 H6−C3 H2′	3.02	2.78	3.19	2.40	3.78	3.77 (3.75)	3.18−4.42
C3 H6−C3 H3′	2.62	2.42	2.77	2.40	2.84	2.66 (2.66)	2.06−3.30
C3 H6−C3 H5″	3.85	3.50	4.15	4.23	3.82	4.02 (4.00)	3.46−4.70
C3 H6−C4 H2′	4.00	3.62	4.34	3.46	5.70	4.72 (4.30)	7.91−9.15
C3 H6−C4 H5	3.84	3.49	4.13	4.31	3.33	3.64 (3.34)	3.56−4.80
C4 H1′−C3 H2′	4.39	3.91	4.88	5.80	4.11	4.11 (4.10)	3.51−4.75
C4 H5−C3 H2′	3.54	3.24	3.78	3.41	4.11	4.19 (3.85)	3.13−4.35
C4 H5−C3 H3′	3.37	3.09	3.58	3.94	3.06	3.16 (2.95)	3.08−4.32
C4 H6−C3 H2′	2.79	2.58	2.95	3.41	2.61	3.32 (3.06)	1.66−2.90
C4 H6−C3 H3′	2.76	2.55	2.92	3.38	2.43	2.57 (2.36)	2.66−3.90
C4 H6−C4 H2′	2.88	2.66	3.05	2.35	3.35	3.42 (3.30)	3.18−4.42
C4 H6−C4 H3′	2.61	2.41	2.75	2.85	2.76	2.80 (2.64)	2.06−3.30
C4 H6−C4 H4′	3.96	3.58	4.28	4.55	4.13	4.13 (4.14)	3.38−4.62
C4 H6−C4 H5″	3.98	3.60	4.31	3.87	3.56	3.56 (3.63)	3.46−4.70

[a]Parentheses indicate distances taken from 0 to 770 ns. After 770 ns, C1 intercalates between C3 and C4. [b]Range of values measured from four CC doublets in PDB entry 1QC0.[57]

1JJ2)[62] of 23S and 5S rRNA. Reasonable agreement with NMR distances and expectations for torsion angles was defined as greater than 60% agreement for the combination of 27 NOEs within error limits (Table 4), three β torsions (between 150° and 210°), and three γ torsions (between 40° and 80°). Figures S8−S16 of the Supporting Information show additional details. Structures at the time points for starting analysis are shown in Figure S17 of the Supporting Information.

All the parm99 trajectories with non-A-form starting structures do not agree with the NMR data (Table 5), so these trajectories were not evaluated. The C3′-endo/syn and C2′-endo/syn simulations for parm99 form looplike structures with C1 and C4 (Figure S18 of the Supporting Information) stacked with amino and carbonyls overlapping after 656 and 931 ns, respectively.

With parm99χ_YIL and parm99TOR, the MD for non-A-form starting structures were evaluated from the time they show an A-form-like structure until a structure inconsistent with NMR was generated or the simulation ended (Table 5). With parm99χ_YIL, base intercalation of C1 between C3 and C4 occurs after 597 ns for the C3′-endo/syn starting structure so only the time range from 65 to 597 ns was analyzed. Evidently, this base intercalation is not unique to the parm99TOR force field. Table S17 of the Supporting Information shows the percentage of C3′-endo and percentage of anti for the entire simulation for each parm99χ_YIL non-A-form starting structure simulation.

The three non-A-form starting structures for parm99TOR reached reasonable agreement with NMR distances and expectations for torsions only after 290, 1093, and 991 ns (Table 5) and were then evaluated until the end of each 1500 ns simulation. The percentage of C3′-endo and percentage of anti for each 1500 ns non-A-form starting structure simulation are listed in Table S17 of the Supporting Information.

Comparison of NMR to MD Simulations. Several criteria can be used to benchmark force fields. One criterion is the length of time a starting structure, known to be accurate, is maintained during MD simulations.[66−68] Another is comparison of structural details determined by crystallography or NMR with those predicted by MD.[33,65−67,69,70] The rapid transitions between structures of unpaired RNAs[30−32,71] allow another test, namely, whether a force field generates a realistic structure if a starting conformation is unrealistic. The results presented here allow all three types of comparisons. Because force fields for RNA are not perfect, however, it is necessary to choose the time intervals in MD simulations that will provide the most insights into force field performance (Table 5). For example, if the starting structure is very inconsistent with NMR spectra, then our comparisons include only those MD time intervals after a structure is generated that is A-form-like. If, after a structure relatively consistent with NMR spectra is generated, unrealistic structures are generated, then the simulations are evaluated with and without the unrealistic data (Tables 3 and 4 and Tables S18−S23 of the Supporting Information). For the MD simulations reported here, the unrealistic structures involve base stacking by intercalation or loop formation (see the footnotes of Table 5, Figures 5 and 6, and Figure S18 of the Supporting Information). To use as

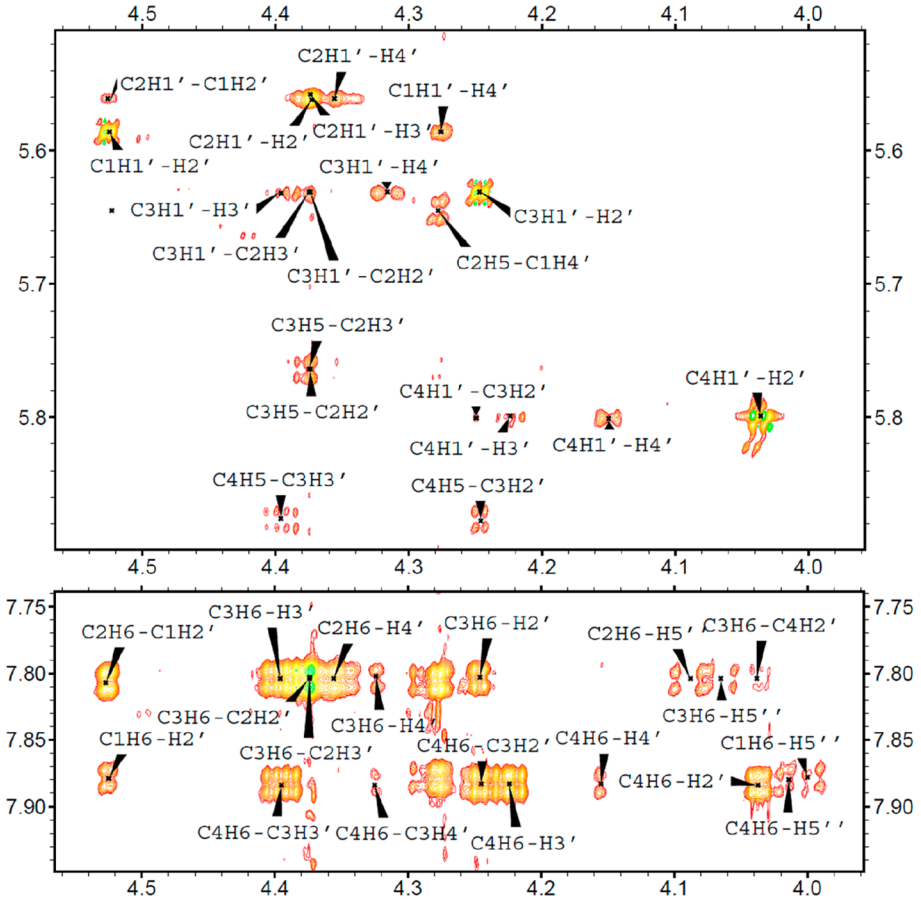


Figure 3. 200 ms NOESY spectrum of r(CCCC) (top) at 5 °C showing the cross-peaks of H5 and H1′ protons to sugar protons and 200 ms NOESY spectrum of r(CCCC) (bottom) at 5 °C showing the cross-peaks from the H6 protons to the sugar protons. Intense cross-peaks between C2 H6 and C1 H2′, C3 H6 and C2 H2′, and C4 H6 and C3 H2′ indicate base−base stacking. The weak cross-peak between C3 H6 and C4 H2′ indicates a population of r(CCCC) where the 3′ sugar is inverted.

much simulation data as possible for a given force field, results conforming to the conditions described above were combined to give an "extended MD" for comparison to NMR data. For example, with parm99TOR, segments were combined from 0 to 770, 290 to 1500, 1093 to 1500, and 991 to 1500 ns of C3′-endo/anti, C3′-endo/syn, C2′-endo/anti, and C2′-endo/syn simulations, respectively, to give a total of 2896 ns (Table 5). The total "extended time" for parm99χ_YIL was 4975 ns. For parm99, only the C3′-endo/anti simulation was used because the non-A-form starting structure simulations never reached good agreement with NMR data. Kinetic studies of polyC indicate that stacking and unstacking are largely noncooperative with half-lives of roughly 40 and 300 ns, respectively, at 5 °C.[30] Because there are three possible stacks to break in r(CCCC), the extended A-form (C3′-endo/anti) simulations for the parm99TOR and parm99χ_YIL force fields allow more than 20 and 40 such transitions, respectively. This is illustrated in plots of the distance between Cn N3 and Cn+1 N3 versus time (Figures S19 and S20 of the Supporting Information).

To allow comparisons to torsion angles generated by MD simulations, ^{31}P-decoupled and ^{31}P-coupled 2D NOESY spectra were used to measure 17 $^{3}J_{H-H}$ and nine $^{3}J_{H-P}$ coupling constants (Table 3). The measured ^{3}J coupling constants were compared to torsion angle predictions from MD by converting predicted torsion angles to predicted ^{3}J coupling constants using a best-fit Karplus relation (Table S13 of the Supporting Information).[54] To determine the accuracy of the MD predictions, the calculated values were compared to the NMR-measured ^{3}J coupling values ±0.5 Hz (Figure 7A). The average difference between MD-calculated ^{3}J scalar coupling and the NMR-measured ^{3}J scalar coupling is shown in Figure 7B. Tables S18−S20 of the Supporting Information show the values of the NMR-measured and MD-predicted ^{3}J scalar couplings for each force field.

To allow comparisons to distances generated by MD simulations (Figure 7C), 27 interproton distances were calculated from NOEs measured in a 200 ms mixing time NOESY spectrum (Table 4). The upper and lower limits used in the NOE comparison were calculated as described in Materials and Methods. In Figure 7D, the average difference between the measured NOE distance and the predicted MD distance is shown for each force field. Tables S21−S23 of the Supporting Information list the measured NMR NOEs and the predicted MD distances for each force field.

As mentioned previously, NMR indicates the terminal sugar is inverted ∼13% of the time. For simulations with A-form (C3′-endo/anti) r(CCCC), parm99, parm99χ_YIL, and parm99TOR force fields predict 27, 2, and 5%, respectively.

DISCUSSION

Modeling the sequence dependence of folding is important for predicting RNA structure.[6,10,72−74] C$_n$ hairpin loops are consistently less stable than other loops,[9] and structure in "unfolded" C$_n$ regions could account for the relative instability.

dx.doi.org/10.1021/bi3010347 | *Biochemistry* 2013, 52, 996−1010

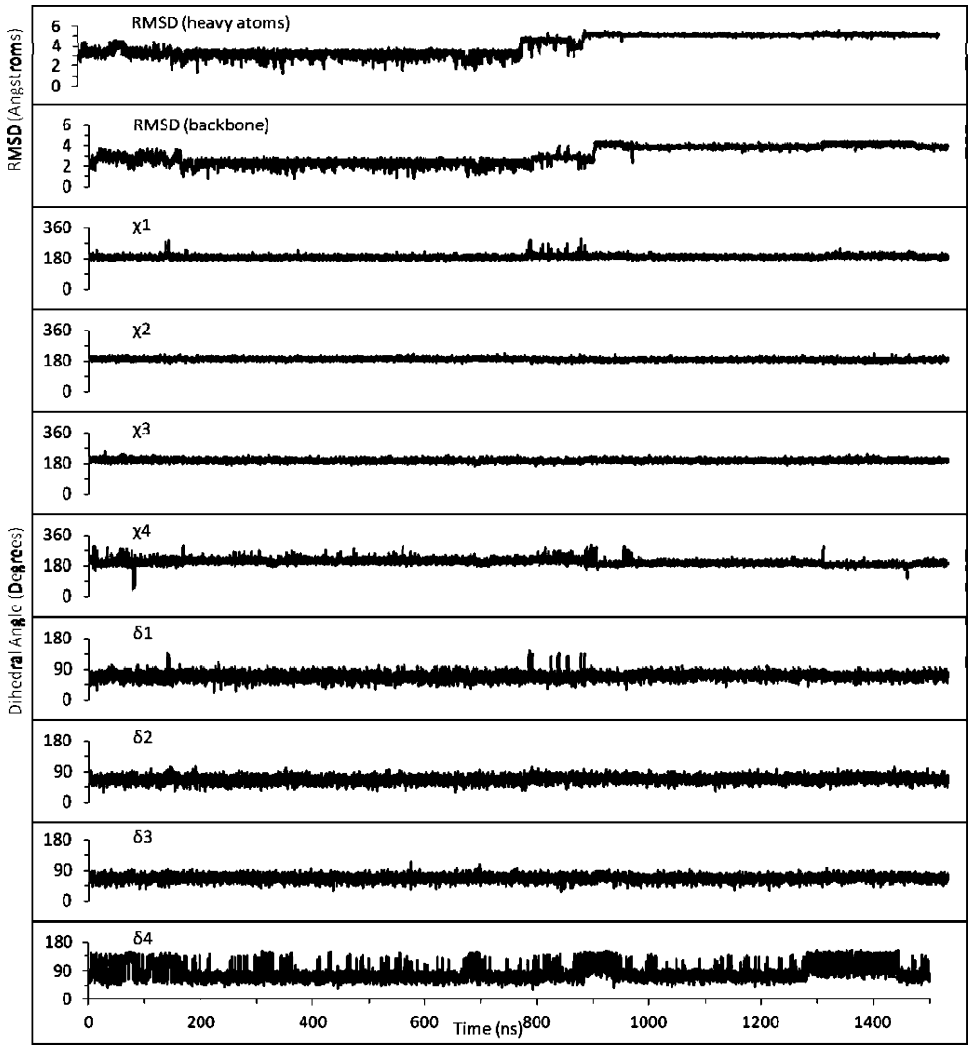


Figure 4. Time evolution (in nanoseconds) for the minimized A-form starting structure of r(CCCC) with the parm99TOR force field. The top two plots show the rmsds of the heavy atoms for the whole structure and of the backbone, respectively, relative to A-form r(CCCC). After 770 ns, C1 intercalates between C3 and C4. The remaining plots correspond to the χ and δ dihedral angles for each residue. δ dihedral angles of 78−90° and 140−152° correspond to C3′-endo and C2′-endo sugar puckers, respectively.[56] Anti, high anti, and syn conformations were defined by χ dihedral angles of 180−239°, 240−300°, and 0−120°, respectively.[56]

For example, on the basis of NMR, it has been suggested that polyC forms a left-handed helix.[13] Moreover, polyC stacks at 37 °C,[30] whereas polyU is unstacked.[75,76] In fact, polyC stacking at 37 °C is even more favorable than stacking in polyA as quantified by equilibrium constants of 1.7 and 0.88, respectively, in 0.05 M sodium cacodylate.[30,71] Here, UV melting studies provide additional evidence of the reduced stability of a C_4 hairpin; NMR provides insight into structural features of the isolated C_4 oligonucleotide, and comparison between NMR spectra and MD simulations tests our understanding of interactions determining RNA structure.

UV Melting of C_4 and U_4 Hairpins Suggests Revision of the Penalty for C_4 Loops in 1 M NaCl. In Table 1, the difference in $\Delta G^\circ_{37,HL}$ between C_4 and U_4 hairpins in 1 M NaCl is 1.30 kcal/mol. This contrasts with the $\Delta\Delta G^\circ_{37,HL}$ of 3.8 kcal/mol[10] calculated from the data of Groebe and Uhlenbeck[9] for hairpins formed by r(pppGGGAUACY$_4$GUAUCCA), where ppp is triphosphate and Y is C or U. A $\Delta\Delta G^\circ_{37,HL}$ of 0.8 kcal/mol has been reported by Shu and Bevilacqua[77] between C_3 and U_3 hairpins at 10 mM NaCl. With different stems, Groebe and Uhlenbeck measured a $\Delta\Delta G^\circ_{37,HL}$ of 1.4 kcal/mol for C_3 and U_3 hairpins in 10 mM Na^+.[9] Interestingly, the difference in ΔG°_{37} between C_4 and U_4 hairpin formation is 0.68 and 0.92 kcal/mol in 5 mM NaCl and 10 mM KCl, respectively, buffered at pH 7.0 with 10 mM sodium MOPS and 0.1 mM sodium EDTA (Table 1). This is within experimental error of the difference of 0.8 kcal/mol measured for the C_4 and U_4 hairpins by Groebe and Uhlenbeck in 10 mM Na_2PO_4 at pH 7.0.[9] The results may reflect a non-nearest neighbor interaction, possibly because of a length dependence for counterion condensation.[78−80] All the comparisons agree, however, that C_n loops are relatively unstable and therefore should be penalized when predicting RNA secondary structure. While the penalty for C_3 and C_4 hairpin loops should perhaps be reduced, it will still disfavor such hairpins, which have not been seen in libraries of known RNA secondary and three-dimensional structures.[81,82]

NMR Spectra Are Consistent with C_4 Being Largely a Right-Handed A-Form-like Helix at pH 7. In NMR spectra of C_4, H6−H1′ cross-peaks for C1, C2, and C4 indicate that the ensemble average has each base in the anti conformation relative to the ribose. A-Form base stacking is consistent with the presence of NOEs between C2 H6 and C1 H2′, C3 H6 and

dx.doi.org/10.1021/bi3010347 | Biochemistry 2013, 52, 996−1010

Table 5. Time It Took for Non-A-Form Starting Structures To Reach a Structure in Reasonable Agreement with NMR Spectra

starting structure	force field	time segment of simulation used for analysis of MD (ns)[a]	no. of major transitions observed
C3′-endo/anti	parm99	0−1500[b]	many
	parm99χ_YIL	0−1500[b]	0
	parm99TOR	0−770[b,d]	2
C3′-endo/syn	parm99	−[c]	many
	parm99χ_YIL	65−597[e]	3
	parm99TOR	290−1500	2
C2′-endo/anti	parm99	−[c]	many
	parm99χ_YIL	35−1500	3
	parm99TOR	1093−1500	2
C2′-endo/syn	parm99	−[c]	many
	parm99χ_YIL	22−1500	2
	parm99TOR	991−1500	2

[a]The time period evaluated for MD simulations was chosen from the point at which simulations show reasonable agreement with the NMR distances and expected torsion angles until a structure inconsistent with NMR data is observed or the simulation reaches its end. Reasonable agreement with the NMR distances and expected torsion angles is defined as 60% agreement with the 27 NOEs within error limits (Table 4), three β dihedrals (between 150° and 210°), and four γ dihedrals (between 40° and 80°). [b]Simulations started as A-form and were evaluated for the entire simulation unless otherwise noted. [c]Simulations did not show reasonable agreement with the NMR distances and expected torsion angles, so these trajectories were not evaluated. [d]Only the first 770 ns was used because after 770 ns, C1 intercalated between C3 and C4. [e]After 597 ns, C1 intercalated between C3 and C4.

C2 H2′, C4 H6 and C3 H2′, and C4 H6 and C3 H3′. Chemical shifts of the phosphorus resonances are within 0.3 ppm of each other, indicating no significant deviation from accepted A-form RNA backbone torsions. Figure S4 of the Supporting Information shows a lack of a large H5−H1′ cross-peak expected for a left-handed helix.[13] Thus, the C_4 NMR data are consistent with a right-handed A-form-like helix at neutral pH. This is in agreement with X-ray fiber diffraction data for polyC.[11]

Tables 3 and 4 compare ^3J scalar coupling constants and NMR distances with those from CC nearest neighbors in the crystal structure of an A-form duplex.[57] Approximately 85% of the values agree between NMR and A-form. While NMR indicates the nucleosides in C_4 are C3′-endo/anti with NOE peak volumes consistent with base stacking, the moderate agreement between the NMR and A-form distance and ^3J scalar coupling constants suggests that on average C_4 is not exactly A-form, but rather A-form-like (Tables 2−4). For example, C4 is apparently inverted roughly 13% of the time. The NMR of C_4 is consistent with the "unfolded" state having significant structure. The fact that some or all of this structure has to be broken to form a hairpin loop agrees with the weaker folding of HPC_4 relative to HPU_4 because adjacent unpaired U residues do not form structure, even in polyU.[75,76]

The Thermodynamics of Hairpin Folding Is Consistent with U_4 Being Less Structured in the Unfolded State and More Structured in the Hairpin Loop Relative to C_4. In general, folding of U_4 hairpins is associated with a more favorable $\Delta H°$ and a less favorable $\Delta S°$ relative to those of C_4 hairpins (Table 1). NMR reveals that the nucleotides are stacked in C_4, as expected on the basis of optical melting and kinetic experiments with polyC.[30,32] Presumably, U_4 is a

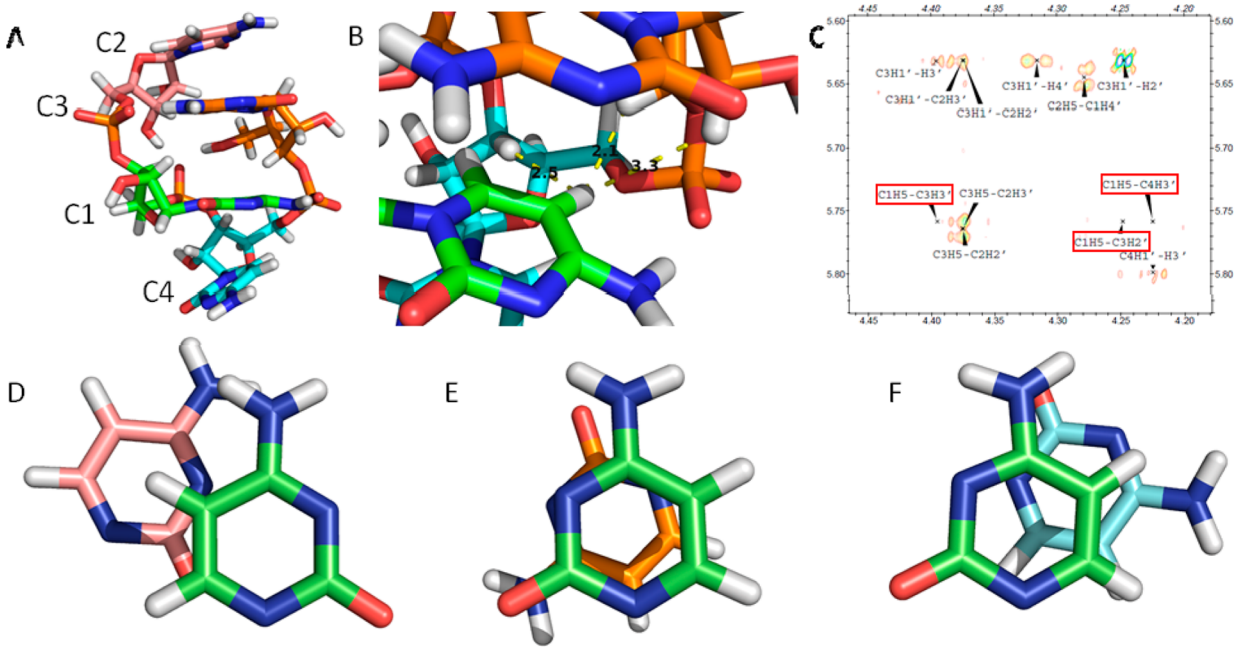


Figure 5. (A) Three-dimensional representation of r(CCCC) when C1 intercalates between C3 and C4 after 770 ns in the MD simulation with A-form (C3′-endo/anti) starting structure with the parm99TOR force field. (B) C1 intercalated between C3 and C4. The distances shown correspond to C1 H5−C4 H3′ (2.5 Å), C1 H5−C3 H3′ (2.1 Å), and C1 H5−C3 H2′ (3.3 Å) distance. (C) 200 ms NOESY spectrum of r(CCCC) showing the absence of the hypothetical H−H cross-peaks (red boxed labels) predicted by the parm99TOR simulation after 770 ns. (D) Typical A-form base stacking between C1 and C2 (from the nucgen structure). (E) Base stacking between C1 and C3 after C1 intercalates between C3 and C4 observed after 770 ns. (F) Base stacking between C1 and C4 after C1 intercalates between C3 and C4 observed after 770 ns. Residues C1−C4 are colored green, pink, orange, and cyan, respectively.

dx.doi.org/10.1021/bi3010347 | Biochemistry 2013, 52, 996−1010

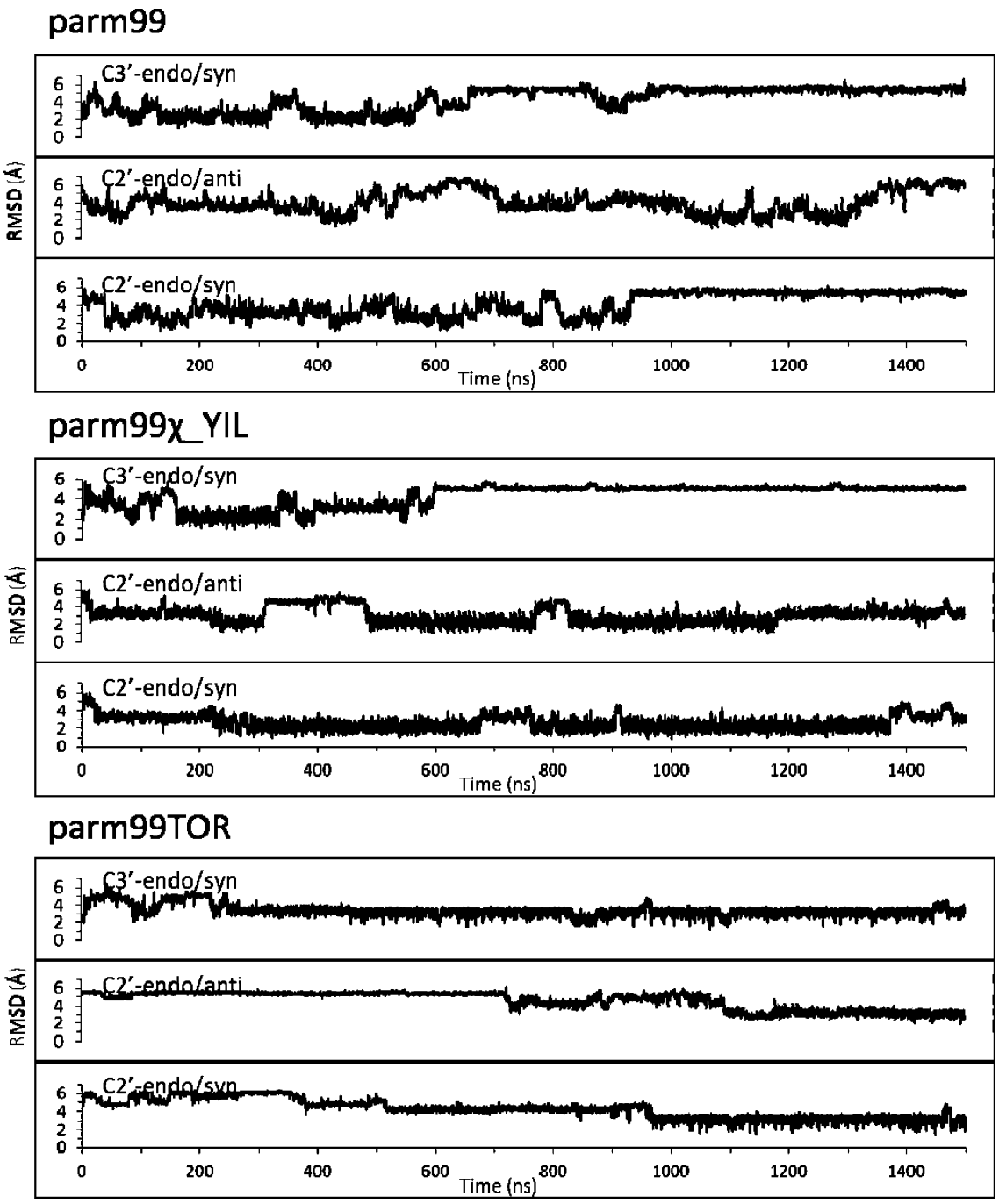


Figure 6. Time evolution (in nanoseconds) of the heavy atom rmsd of the simulations of non-A-form starting structures relative to A-form r(CCCC) for each force field. In both the C3′-endo/syn and C2′-endo/syn starting structures in the parm99 force field, C1 stacks on C4 forming a looplike structure after 656 and 931 ns, respectively. Around 600 ns for the C3′-endo/syn starting structure with the parm99χ_YIL force field, C1 intercalates between C3 and C4.

random coil, as is polyU.[75,76] Thus, at least one stacking interaction must be broken in C_4 to allow the 180° backbone turn required for a C_4 hairpin, but no such interaction is broken in the formation of a U_4 hairpin. Moreover, NMR indicated that a hydrogen-bonded UU pair forms at the base of a U_4 hairpin loop.[83]

parm99χ_YIL and parm99TOR Improve the Agreement between NMR and MD Simulations. Predictions from MD simulations with the parm99 force field show the least agreement with NMR data (Tables 2−4). While the NMR spectra are consistent with an ensemble of A-form-like conformations, the parm99 force field prefers a C2′-endo sugar pucker and a high-anti base orientation.[33,67] Figure 7 shows the performance of force fields in predicting the percent of correct ^3J scalar coupling and distances and the average absolute differences between the MD predictions and NMR data. The parm99 force field shows the poorest performance, and with the non-A-form starting structures, a final A-form-like structure is never reached. parm99χ_YIL and parm99TOR perform similarly to each other (Figure 7).

With the parm99χ_YIL and parm99TOR force fields, all of the non-A-form starting structures eventually reach an A-form-

dx.doi.org/10.1021/bi3010347 | *Biochemistry* 2013, 52, 996−1010

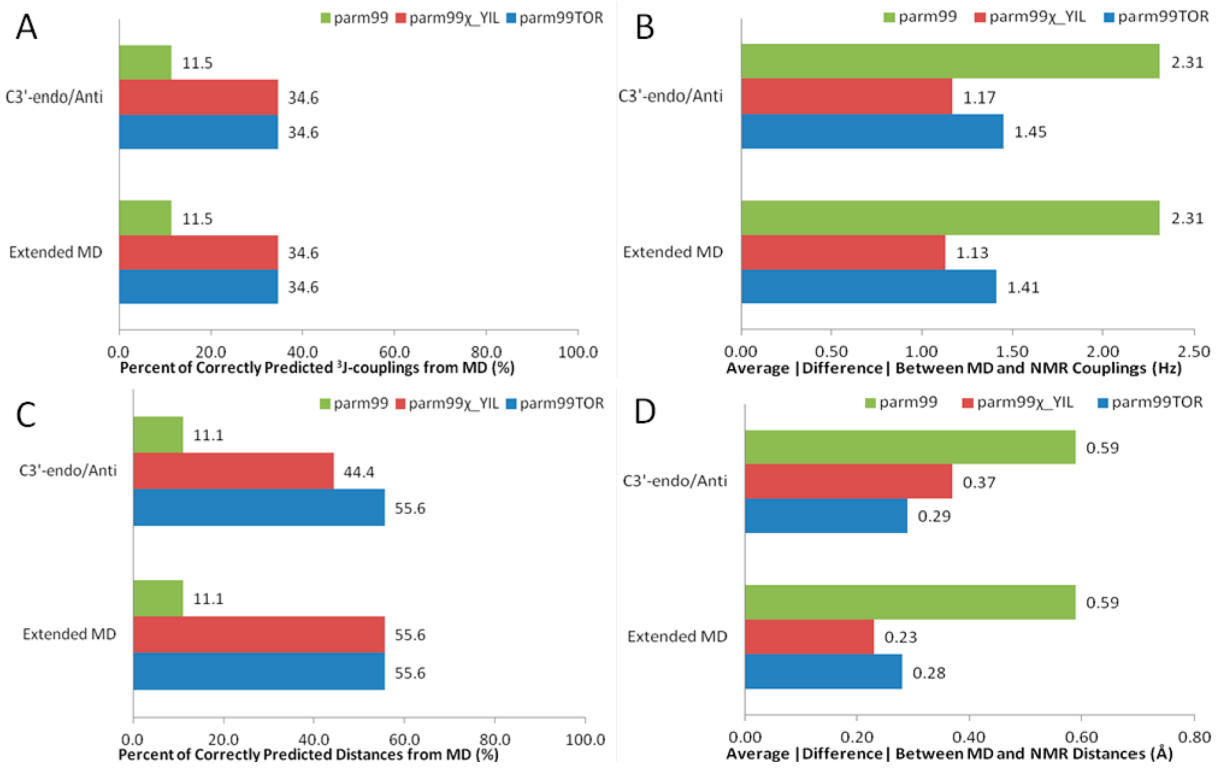


Figure 7. Comparisons between MD-predicted and NMR-measured ^3J couplings (top) and distances (bottom). The top bar graph in each panel summarizes the results from simulations starting from an A-form structure (Tables 3 and 4). The bottom bar graph (Extended MD) in each panel adds to A-form starting structure simulations the results obtained from simulations starting with non-A-form structures in the time range after reaching an A-form-like structure and before any base intercalation. Table 5 details the portion of each simulation used for comparison. For parm99, parm99χ_YIL, and parm99TOR, the total Extended MD time is 1500, 4975, and 2896 ns, respectively. The Extended MD simulation data for parm99 are the same as the C3′-endo/anti simulation data because of the non-A-form starting structure simulations not forming an A-form-like structure. parm99 (green), parm99χ_YIL (red), and parm99TOR (blue) force fields were tested. (A) Percentage of MD ^3J couplings correctly predicted within ±0.5 Hz. (B) Average of the absolute values of differences between MD-predicted and NMR-measured ^3J couplings. (C) Percentage of MD distances correctly predicted between the error limits of the measured NOEs. (D) Average of the absolute values of differences between MD-predicted distances and NMR-measured NOEs. Tables S18−S23 of the Supporting Information show the MD-predicted and NMR-measured values used in these plots.

Table 6. Average Backbone Torsion Angles (degrees) Measured by NMR and Predicted by MD with A-Form Starting Structures

torsion angle	NMR	parm99	parm99χ_YIL	parm99TOR	polyC[f]
β (P−O5′−C5′−C4′)	188[a]	178	173	167	173
γ (O5′−C5′−C4′−C3′)	60[b]	73	69	69	47
δ (C5′−C4′−C3′−O3′)	78−90[c]	99[e]	75[e]	71[e]	78−90[g]
ε (C4′−C3′−O3′−P)	239[d]	219	218	201	231

[a]Calculated from the measured n P−n H5′ and n P−n H5″ ^3J couplings (Table 3 and Table S13 of the Supporting Information) using the relations $\varphi = \beta + 120°$ and $\varphi' = \beta - 120°$, respectively,[85] where φ and φ' are the n P−n H5′ and n P−n H5″ torsions, respectively. An error of ±0.5 Hz was assumed in the ^3J couplings leading to a torsion angle range of 186−192°. [b]Approximated from the n H4′−n H5′ and n H4′−n H5″ torsions having ^3J couplings of ∼2 and ∼1 Hz (Table 3 and Table S13 of the Supporting Information), respectively, corresponding to a gauche$^+$ conformation.[85] An error of ±0.5 Hz was assumed in the ^3J couplings leading to a torsion angle range of 55−65°. [c]The lack of C1−C3 H1′ peak splitting observed in the 1D spectrum (Figure S2 of the Supporting Information) indicates a C3′-endo sugar pucker,[54] so the accepted δ range for C3′-endo is used.[56] [d]Calculated from the measured ^3J couplings (Table 3) using the relation $\psi = \varepsilon + 120°$ for n H3′−n+1 P torsions,[85] where ψ is the n H3′−n+1 P torsion. An error of ±0.5 Hz was assumed in the ^3J couplings leading to a torsion angle range of 237−241°. [e]Calculated directly from the MD δ torsion trajectory for C1−C3. [f]From X-ray fiber studies on polyC.[11] [g]X-ray fiber results indicate a C3′-endo sugar pucker,[11] but no specific δ torsion angle; therefore, an accepted range for the δ torsion in a C3′-endo sugar pucker is used.[56]

like structure in reasonable agreement with NMR spectra. This suggests that the four 1500 ns simulations provide sufficient sampling for comparison of the three force fields. For each non-A-form starting structure, parm99χ_YIL reaches A-form-like significantly faster than parm99TOR. The times taken for non-A-form starting structures to reach A-form-like structure after minimization for parm99χ_YIL are 65, 35, and 22 ns for C3′-endo/syn, C2′-endo/anti, and C2′-endo/syn, respectively, compared to 290, 1093, and 991 ns, respectively, for parm99TOR. A-Form-like conformations are not always stable, however (Figures 4 and 6).

The parm99χ_YIL force field has revised parameters only for the χ dihedral,[33] whereas parm99TOR[34] also adds revised parameters for the β, ε, and ζ torsions along with the parmbsc[66] parameters for the α and γ torsions. Both parm99χ_YIL and parm99TOR are improvements over parm99 for prediction of

dx.doi.org/10.1021/bi3010347 | Biochemistry 2013, 52, 996−1010

NMR data, but the results do not indicate improvement from the additional reparametrizations included in parm99TOR.

parmOL[67] is a recent force field not tested in the experiments reported here. The parmOL force field includes revisions to the χ,[67,84] α, and γ tosions.[66] parmOL and parm99χ_YIL performed similarly when tested with folded RNAs.[67,84] Because parm99χ_YIL and parm99TOR perform similarly on r(CCCC) and tetramer duplexes,[34] revised parameters for α, β, γ, ε, and ζ torsions likely make little difference. Thus, parmOL simulations are likely to be similar to those with parm99χ_YIL and parm99TOR. The results provide consensus that revisions to χ parameters provide major improvements to parm99.[33,34,65,67,84]

Comparison of Backbone Scalar Couplings among the A-Form Reference, NMR, and MD Predictions. NMR scalar couplings can be compared to couplings expected for the standard A-form and to MD predictions for each of the force fields averaged over all β, γ, and ε torsion angles when starting with C3′-endo/anti structure (Table 3). The A-form reference structure[57] has an average β dihedral angle of 165.6°, corresponding to the β trans conformation. This angle predicts average ^3J scalar coupling values of 5.0 and 1.4 Hz for all n P−n H5′ and n P−n H5″ torsions, respectively (Table S13 of the Supporting Information).[54,85] NMR ^3J scalar couplings are an average of 3.8 and 0.9 Hz for n P−n H5′ and n P−n H5″ torsions, respectively, corresponding to a β angle of ∼188° (Table 6 and Table S13 of the Supporting Information).[54,85] Average predicted ^3J scalar couplings are 3.2, 3.7, and 4.5 Hz for n P−n H5′ torsions and 2.9, 2.3, and 2.8 Hz for n P−n H5″ torsions for parm99, parm99χ_YIL, and parm99TOR, respectively (Table 3). These ^3J scalar couplings correspond to β torsions of ∼178°, ∼173°, and ∼167° for parm99, parm99χ_YIL, and parm99TOR, respectively (Table 6 and Table S13 of the Supporting Information).[54,85] Both NMR-measured and MD-predicted β dihedral angles are still within the range associated with the trans conformation.

In the NMR spectra, the ^3J scalar coupling constants for H4′−H5′ and H4′−H5″ cross-peaks, which are representative of the γ torsion, were ∼2 and ∼1 Hz, respectively (Table 3). This corresponds to the γ torsion residing completely in the gauche$^+$ orientation, i.e., a torsion angle of ∼60° (Table 6 and Table S13 of the Supporting Information),[54,85] which is consistent with the A-form reference structure.[57] The average MD predictions from the γ trajectory for parm99, parm99χ_YIL, and parm99TOR are ∼73°, ∼69°, and ∼69°, respectively, for γ (Table 6). MD predictions are averages that result from γ torsions that fall within g$^+$, g$^-$, and trans conformations, but g$^+$ accounts for ∼85% of the angles in the trajectories. The three force fields provide reasonable predictions of γ even though only parm99TOR contains the γ torsion parameters developed by Perez et al.[66] to disfavor the rarely seen trans conformation.

In the NMR spectra, the n H3′−n+1 P ^3J scalar coupling constant has an average of 9.1 Hz (Table 3), which is consistent with an ε torsion angle of ∼239° (Table 6 and Table S13 of the Supporting Information).[54,85] The A-form reference structure[57] has an average ε dihedral angle of ∼210° calculated from four CC doublets, corresponding to an n H3′−n+1 P ^3J scalar coupling constant of 8.2 Hz. Average MD predictions are ∼5.5, ∼8.8, and ∼6.2 Hz for parm99, parm99χ_YIL, and parm99TOR, respectively (Table 3), corresponding to ε angles of ∼219°, ∼218°, and ∼201°, respectively (Table 6 and Table S13 of the Supporting Information).[54,85] An ε dihedral angle of 231° has been observed in X-ray fiber diffraction studies of polyC.[11]

Even though the average MD-predicted backbone torsion angles are in agreement with A-form ranges,[57] the small differences in torsion angles cumulatively lead to a large backbone deformation. Panels A and B of Figure 8 compare

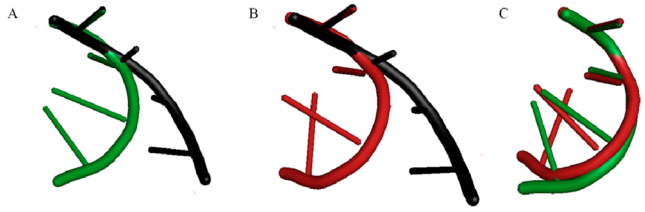


Figure 8. Structures of r(CCCC) representative of those with typical rmsds relative to A-form and aligned with nucgen A-form structure (black) in panels A and B by aligning residue C1 in both structures: (A) parm99χ_YIL (green), (B) parm99TOR (red), and (C) parm99χ_YIL aligned with parm99TOR. These alignments were generated with PyMOL[46] using the align function.

representative parm99χ_YIL and parm99TOR backbones with the nucgen backbone, which is representative of the A-form. This backbone deformation causes the large rmsds of the MD simulations relative to an A-form reference structure.

Areas in Which MD Force Fields Can Be Improved. Comparisons between NMR spectra and MD simulations of unpaired oligonucleotides provide benchmarks for testing force fields.[65,69,70] Comparisons of simulations for r(GACC)[65] and r(CCCC) indicate that revision of torsion angle parameters improves AMBER simulations relative to parm99. Comparisons of thermodynamic integration calculations with measured differences in stabilities of tetramer duplexes with GC or isoGisoC base pairs also showed improvement.[34] Several results, however, suggest that further revisions of the force field are necessary.

A striking difference between NMR spectra and the parm99TOR MD simulation starting with A-form and parm99χ_YIL starting with C3′-endo/syn is the persistence of base intercalation in the simulations. NMR spectra lack NOESY cross-peaks expected for intercalations. In both simulations, C1 intercalated between C3 and C4 after 600−800 ns and remained in this conformation until the end of the 1500 ns simulation (Figures 4−6). A similar stacking is observed in the parm99 force field with C3′-endo/syn and C2′-endo/syn starting structures when a looplike structure is formed by C1 stacking on C4 after 650−950 ns. These unique structures were stabilized by favorable stacking interactions between amino and carbonyl groups (e.g. Figure 5E,F). Experimental studies of a similar stacking interaction between the amino of cytosine and a carbonyl of uracil showed this type of interaction is energetically favorable.[86,87] Base intercalation was also observed in a parm99χ_YIL simulation of the tetramer r(GACC), where G1 intercalated between C3 and C4 and the structure was stable for ∼400 ns.[65] This suggests that energy terms driving stacking may be too favorable, as suggested by Banas et al.[88] for A-tracts in B-DNA. Such terms may include base−base and/or base−water interactions. The NMR results for r(GACC)[65] and r(CCCC) should be good benchmarks for testing future revisions of RNA force fields.

dx.doi.org/10.1021/bi3010347 | Biochemistry 2013, 52, 996−1010

■ ASSOCIATED CONTENT

Supporting Information

Chemical shift assignments, NMR acquisition parameters, NOE peak volumes, NMR ^3J couplings, 1D NMR spectrum, $^1H-^{31}P$ HETCOR, MD starting structure coordinates, MD trajectories, MD-predicted interproton distances, MD-predicted ^3J couplings, non-A-form starting structures, and salt-dependent thermodynamic parameters. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: (585) 275-3207. Fax: (585) 276-0205. E-mail: turner@chem.rochestere.edu.

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

C, cytidine; HPC_4, r(5′GGACCCCCGUCC); HPU_4, r(5′GGACUUUUGUCC); MD, molecular dynamics; MM, molecular mechanics; PAGE, polyacrylamide gel electrophoresis; QM, quantum mechanical; U, uracil.

■ REFERENCES

(1) Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., and Cech, T. R. (1982) Self-splicing RNA: Autoexcision and autocatalyzation of the ribosomal RNA intervening sequence of tetrahymena. *Cell 31*, 147−157.

(2) Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983) The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell 35*, 849−857.

(3) Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell 75*, 843−853.

(4) Ruvkun, G. (2001) Glimpses of a tiny RNA world. *Science 294*, 797−799.

(5) Wang, G. H., Zoulim, F., Leber, E. H., Kitson, J., and Seeger, C. (1994) Role of enzymatic activity of the reverse transcriptase of hepatitis B viruses. *J. Virol. 68*, 8437−8442.

(6) Mathews, D. H., Disney, M. D., Childs, J. L., Schroeder, S. J., Zuker, M., and Turner, D. H. (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. U.S.A. 101*, 7287−7292.

(7) Deigan, K. E., Li, T. W., Mathews, D. H., and Weeks, K. M. (2009) Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci. U.S.A. 106*, 97−102.

(8) Hart, J. M., Kennedy, S. D., Mathews, D. H., and Turner, D. H. (2008) NMR-assisted prediction of RNA secondary structure: Identification of a probable pseudoknot in the coding region of an R2 retrotransposon. *J. Am. Chem. Soc. 130*, 10233−10239.

(9) Groebe, D. R., and Uhlenbeck, O. C. (1988) Characterization of RNA hairpin loop stability. *Nucleic Acids Res. 16*, 11725−11735.

(10) Mathews, D. H., Sabrina, 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.

(11) Arnott, S., Chandrasekaren, R., and Leslie, A. G. W. (1976) Structure of the single-stranded polyribonucleotide polycytidylic acid. *J. Mol. Biol. 106*, 735−748.

(12) Langridge, R., and Rich, A. (1963) Molecular Structure of Helical Polycytidylic Acid. *Nature 198*, 725−728.

(13) Broido, M., and Kearns, D. R. (1982) Proton NMR evidence for a left-handed helical structure of poly(ribocytidylic acid) in neutral solution. *J. Am. Chem. Soc. 104*, 5207−5216.

(14) Brahms, J., Maurizot, J. C., and Michelson, A. M. (1967) Conformation and thermodynamic properties of oligocytidylic acids. *J. Mol. Biol. 25*, 465−480.

(15) Brahms, J. (1963) Optical activity and the conformation of polynucleotides. *J. Am. Chem. Soc. 85*, 3298−3300.

(16) Green, G., and Mahler, H. R. (1970) Comparative study of polyribonucleotides in aqueous and glycerol solutions. *Biochemistry 9*, 368−387.

(17) Causley, G. C., and Johnson, W. C. (1982) Polynucleotide conformation from flow dichroism studies. *Biopolymers 21*, 1763−1780.

(18) Fasman, G. D., Lindblow, C., and Grossman, L. (1964) The helical conformations of polycytidylic acid: Studies on the forces involved. *Biochemistry 3*, 1015−1021.

(19) Alder, A., Grossman, L., and Fasman, G. D. (1967) Single-stranded oligomers and polymers of cytidylic and 2′-deoxycytidylic acids: Comparative optical rotatory studies. *Proc. Natl. Acad. Sci. U.S.A. 57*, 423−430.

(20) Akinrimsi, E. O., Sander, C., and Ts'o, P. O. P. (1963) Properties of helical polycytidylic acid. *Biochemistry 2*, 340−344.

(21) Maleev, V., Semenov, M., Kashpur, V., Bolbukh, T., Shestopalova, A., and Anishchenko, D. (2002) Structure and hydration of polycytidyic acid from the data of infrared spectroscopy, EHF dielectrometry and computer modeling. *J. Mol. Struct. 605*, 51−61.

(22) Chou, C. H., and Thomas, G. J. (1977) Raman spectral studies of nucleic acids. XVI. Structures of polyribocytidylic acid in aqueous solution. *Biopolymers 16*, 765−789.

(23) Florian, J., Baumruk, V., and Leszczynski, J. (1996) IR and Raman spectra, tautomeric stabilities, and scaled quantum mechanical force fields of protonated cytosine. *J. Phys. Chem. 100*, 5578−5589.

(24) O'Connor, T., and Scovell, W. M. (1981) pH-dependant Raman spectra and thermal melting profiles for polycytidylic acid. *Biopolymers 20*, 2351−2367.

(25) Annamalai, A., and Keiderling, T. A. (1987) Vibrational circular dichroism of poly(ribonucleic acids). A comparative study in aqueous solution. *J. Am. Chem. Soc. 109*, 3125.

(26) Xiang, T., Goss, D. J., and Diem, M. (1993) Strategies for the computation of infrared CD and absorption of biological molecules: Ribonucleic acids. *Biophys. J. 65*, 1255−1261.

(27) Self, B. D., and Moore, D. S. (1997) Nucleic acid vibrational circular dichroism, absorption, and linear dichroism spectra. I. A DeVoe theory approach. *Biophys. J. 73*, 339−347.

(28) Petrovic, A. G., and Polavarapu, P. L. (2006) Structural transitions in polycytidylic acid by changes in pH and temperature: Vibrational circular dichroism study and film states. *J. Phys. Chem. B 110*, 22826−22833.

(29) Zarudnaya, M. I., Samijlenko, S. P., Potyahaylo, A. L., and Hovorun, D. M. (2002) Structural transitions in polycytidylic acids: Proton buffer capacity data. *Nucleosides, Nucleotides Nucleic Acids 21*, 125−137.

(30) Freier, S. M., Hill, K. O., Dewey, T. G., Marky, L. A., Breslauer, K. J., and Turner, D. H. (1981) Solvent effects on the kinetics and thermodynamics of stacking in poly(cytidylic acid). *Biochemistry 20*, 1419−1426.

(31) Turner, D. H. (2000) Conformational changes. In *Nucleic Acids: Structures, Properties, and Functions* (Bloomfield, V. A., Crothers, D. M., and Tinoco, J. I., Eds.) pp 259−334, University Science Books, Herndon, VA.

(32) Porchke, D. (1976) The nature of stacking interactions in polynucleotides. Molecular states in oligo- and polyribocytidylic acid by relaxation analysis. *Biochemistry 15*, 1495−1499.

(33) Yildirim, I., Stern, H. A., Kennedy, S. D., Tubbs, J. D., and Turner, D. H. (2010) Reparameteriation of RNA χ torsion parameters for the AMBER force field and comparison to NMR spectra for cytidine and uridine. *J. Chem. Theory Comput. 6*, 1520−1531.

(34) Yildirim, I., Kennedy, S. D., Stern, H. A., Hart, J. M., Kierzek, R., and Turner, D. H. (2012) Revision of AMBER torsional parameters for RNA improves free energy predictions for tetramer duplexes with GC and iGiC base pairs. *J. Chem. Theory Comput. 8*, 172−181.

(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 second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc. 117*, 5179−5197.
(36) Wang, J., Cieplak, P., and Kollman, P. A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem. 21*, 1049−1074.
(37) 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.
(38) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) Nmrpipe: A multidimensional spectral processing system based on Unix Pipes. *J. Biol. NMR 6*, 277−293.
(39) Goddard, T. D., and Kneller, D. G. (2004) *SPARKY 3: NMR assignment and integration software*, University of California, San Francisco.
(40) Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) *AMBER 9*, University of California, San Francisco.
(41) *Avogadro: An open-source molecular builder and visualization tool*, version 1.1.0 (2012) http://avogadro.openmolecules.net/.
(42) Hawkins, G. D., Cramer, C. J., and Truhar, D. G. (1995) Pairwise solute descreening of solute charges from a dielectric medium. *Chem. Phys. Lett. 246*, 122−129.
(43) Hawkins, G. D., Cramer, C. J., and Truhar, D. G. (1996) Parameterized models of aqueous free energy of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *J. Phys. Chem. 100*, 19824−19839.
(44) Tsui, V., and Case, D. A. (2000) Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers 56*, 275−291.
(45) Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., Dinola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys. 81*, 3684−3670.
(46) *The PyMOL Molecular Graphics System*, version 1.3 (2010) Schrödinger, LLC, New York.
(47) Murray, L. J. W., Arendall, W. B., III, Richardson, D. C., and Richardson, J. S. (2003) RNA backbone is rotameric. *Proc. Natl. Acad. Sci. U.S.A. 100*, 13904−13909.
(48) Aqvist, J. (1990) Ion-water interaction potentials derived from free energy perturbation simulations. *J. Phys. Chem. 94*, 8021−8024.
(49) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926−935.
(50) Sagui, C., Pedersen, L. G., and Darden, T. A. (2004) Towards an accurate representation of electrostatics in classical force fields: Efficient implementation of multipolar interactions in biomolecular simulations. *J. Chem. Phys. 120*, 73−87.
(51) Toukmaji, A., Sagui, C., Board, J., and Darden, T. A. (2000) Efficient particle-mesh Ewald based approach to fixed and induced dipolar interactions. *J. Chem. Phys. 113*, 10913−10927.
(52) Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical-Integration of Cartesian equations of motion of a system with constraints: Molecular-Dynamics of N-Alkanes. *J. Comput. Phys. 23*, 327−341.
(53) Cerutti, D. S., Duke, R., Freddolino, P. L., Fan, H., and Lybrand, T. P. (2008) A vulnerability in popular molecular dynamics packages concerning Langevin and Andersen dynamics. *J. Chem. Theory Comput. 4*, 1669−1680.
(54) Marino, J. P., Schwalbe, H., and Griesinger, C. (1999) *J*-Coupling Restraints in RNA structure determination. *Acc. Chem. Res. 32*, 614−623.
(55) Altona, C., and Sunderlingam, M. (1972) Conformational analysis of the sugar ring in nucleosides and nucleotides. A new description using the concept of pseudorotation. *J. Am. Chem. Soc. 94*, 8205−8212.
(56) Richardson, J. S., Schneider, B., Murray, L. W., Kapral, G. J., Immormino, R. M., Headd, J. J., Richardson, D. C., Ham, D., Hershkovits, E., Williams, L. D., Keating, K. S., Pyle, A. M., Micallef, D., Westbrook, J., and Berman, H. M. (2008) RNA backbone: Consensus all-angle conformers and modular string nomenclature (an RNA Ontology Consortium contribution). *RNA 14*, 465−481.
(57) 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.
(58) Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res. 28*, 235−242.
(59) 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.
(60) Sheehy, J. P., Davis, A. R., and Znosko, B. M. (2010) Thermodynamic characterization of naturally occurring RNA tetraloops. *RNA 16*, 417−429.
(61) Schneider, B., Moravek, Z., and Berman, H. M. (2004) RNA conformational classes. *Nucleic Acids Res. 32*, 1666−1677.
(62) Klein, D. J., Schmeing, T. M., Moore, P. B., and Steitz, T. A. (2001) The kink-turn: A new RNA secondary structure motif. *EMBO J. 20*, 4214−4221.
(63) Correll, C. C., Beneken, J., Plantinga, M. J., Lubbers, M., and Chan, Y. L. (2003) The common and distinctive features of the bulged-G motif bases on a 1.04 Å resolution RNA structure. *Nucleic Acids Res. 31*, 6806−6818.
(64) Szeczak, A. A., and Moore, P. B. (1995) The Sarcin/Ricin loop a modular RNA. *J. Mol. Biol. 247*, 81−98.
(65) Yildirim, I., Stern, H. A., Tubbs, J. D., Kennedy, S. D., and Turner, D. H. (2011) Benchmarking AMBER force fields from RNA: Comparisons to NMR spectra for single-stranded r(GACC) are improved by revised χ torsions. *J. Phys. Chem. B 115*, 9261−9270.
(66) Perez, A., Marchan, I., Svozil, D., Sponer, J., Cheathan, T. E., Laughton, C. A., and Orozco, M. (2007) Refinement of AMBER force field for nucleic acids: Improving the description of α/γ conformer. *Biophys. J. 92*, 3817−3829.
(67) Zgarbova, M., Otyepka, M., Sponer, J., Mladek, A., Banas, P., Cheatham, T. E., and Jurecka, P. (2011) Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J. Chem. Theory Comput. 7*, 2886−2902.
(68) Swaminathan, S., Ravishanker, G., and Beveridge, D. L. (1991) Molecular dynamics of B-DNA including water and counterions: A 140-ps trajectory for d(CGCGAATTCGCG) based on the GROMOS force field. *J. Am. Chem. Soc. 113*, 5027−5040.
(69) Vokacova, Z., Budesinsky, M., Rosenberg, I., Schneider, B., Sponer, J., and Sychrovsky, V. (2009) Structure and dynamics of the ApA, ApC, CpA, and CpC RNA dinucleoside monophosphates resolved with NMR scalar spin-spin couplings. *J. Phys. Chem. B 113*, 1182−1191.
(70) Eichhorn, C. D., Feng, J., Suddala, K. C., Walter, N. G., Brooks, C. L., and Al-Hashimi, H. M. (2012) Unraveling the structural complexity in a single-stranded RNA tail: Implications for efficient ligand binding in the prequeuosine riboswitch. *Nucleic Acids Res. 40*, 1345−1355.
(71) Dewey, T. G., and Turner, D. H. (1979) Laser temperature-jump study of stacking in adenylic acid polymers. *Biochemistry 18*, 5757−5762.
(72) Das, R., and Baker, D. (2008) Macromolecular modeling with Rosetta. *Annu. Rev. Biochem. 77*, 363−382.

(73) Bernauer, J., Huang, X., Sim, A. Y. L., and Levitt, M. (2011) Fully differentiable coarse-grained and all-atom knowledge-based potentials for RNA structure evaluation. *RNA 17*, 1066−1075.
(74) Sripakdeevong, P., Kladwang, W., and Das, R. (2011) An enumerative stepwise ansatz enables atomic-accuracy RNA loop modeling. *Proc. Natl. Acad. Sci. U.S.A. 108*, 20573−20578.
(75) Richards, E. G., Flessel, C. P., and Fresco, J. R. (1963) Polynucleotides. VI. Molecular properties of polyribouridylic acid. *Biopolymers 1*, 431−446.
(76) Inners, L. D., and Felsenseld, G. (1970) Conformation of polyuridylic acid in solution. *J. Mol. Biol. 50*, 373−389.
(77) Shu, Z., and Bevilacqua, P. C. (1999) Isolation and characterization of thermodynamically stable and unstable RNA hairpins from a triloop combinatorial library. *Biochemistry 38*, 15369−15379.
(78) Manning, G. S. (1978) The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys. 11*, 179−246.
(79) Fenley, M. O., Manning, G. S., and Olson, W. K. (1990) Approach to the limit of counterion condensation. *Biopolymers 30*, 1191−1203.
(80) Olmsted, M. C., Anderson, C. F., and Record, M. T. (1989) Monte-Carlo description of oligoelectrolyte properties of DNA oligomers: Range of the end effect and the approach of molecular and thermodynamic properties to the polyelectrolyte limits. *Proc. Natl. Acad. Sci. U.S.A. 86*, 7766−7770.
(81) Popenda, M., Szachniuk, M., Blazewicz, M., Wasik, S., Burke, E. K., Blazewiczm, J., and Adamiak, R. W. (2010) RNA FRABASE 2.0: An advanced web-accessible database with the capacity to search the three-dimensional fragments within RNA structures. *BMC Bioinf. 11*, 231.
(82) Vanegas, P. L., Hudson, G. A., Davis, A. R., Kelly, S. C., Kirkpatrick, C. C., and Znosko, B. M. (2011) RNA CoSSMos: Characterization of secondary structure motifs—a searchable database of secondary structure motifs in RNA three-dimensional structures. *Nucleic Acids Res. 40*, D439−D444.
(83) Sashital, D. G., Venditti, V., Angers, C. G., Cornilescu, G., and Butcher, S. E. (2007) Structure and thermodynamics of a conserved U2 snRNA domain from yeast and human. *RNA 13*, 328−338.
(84) Banas, P., Hollas, D., Zgarbova, M., Jurecka, P., Orozco, M., Cheatam, T. E., Sponer, J., and Otyepka, M. (2010) Performance of molecular mechanics force fields for RNA simulations: Stability of UUCG and GNRA hairpins. *J. Chem. Theory Comput. 6*, 3836−3849.
(85) Wijmenga, S. S., and van Buuren, B. N. M. (1998) The use of NMR methods for conformational studies of nucleic acids. *Prog. Nucl. Magn. Reson. Spectrosc. 32*, 287−387.
(86) Burkard, M. E., Kierzek, R., and Turner, D. H. (1999) Thermodynamics of unpaired terminal nucleotides on short RNA helixes correlates with stacking at helix termini in larger RNAs. *J. Mol. Biol. 290*, 967−982.
(87) Blose, J. M., Proctor, D. J., Veeraraghavan, N., Misra, V. K., and Bevilacqua, P. C. (2009) Contribution of closing base pair to the exceptional stability in RNA tetraloops: Roles for molecular mimicry and electrostatic factors. *J. Am. Chem. Soc. 131*, 8474−8484.
(88) Banas, P., Mladek, A., Otyepka, M., Zgarbova, M., Jurecka, P., Svozil, D., Lankas, F., and Sponer, J. (2012) Can we accurately describe the structure of adenine tracts in B-DNA? Reference quantum-chemical computations reveal overstabilization of stacking by molecular mechanics. *J. Chem. Theory Comput. 8*, 2448−2460.