JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Revision of AMBER Torsional Parameters for RNA Improves Free Energy Predictions for Tetramer Duplexes with GC and iGiC Base Pairs

Ilyas Yildirim,[†] Scott D. Kennedy,[‡] Harry A. Stern,[†] James M. Hart,[†] Ryszard Kierzek,[§] and Douglas H. Turner*[,†]

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

[‡]Department of Biochemistry and Biophysics, School of Medicine and Dentistry, University of Rochester, Rochester, New York 14642, United States

[§]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 60-714 Poznan, Poland

Supporting Information

ABSTRACT: All-atom force fields are important for predicting thermodynamic, structural, and dynamic properties of RNA. In this paper, results are reported for thermodynamic integration calculations of free energy differences of duplex formation when CG pairs in the RNA duplexes r(CCGG)$_2$, r(GGCC)$_2$, r(GCGC)$_2$, and r(CGCG)$_2$ are replaced by isocytidine−isoguanosine (iCiG) pairs. Agreement with experiment was improved when ε/ζ, α/γ, β, and χ torsional parameters in the AMBER99 force field were revised on the basis of quantum mechanical calculations. The revised force field, AMBER99TOR, brings free energy difference predictions to within 1.3, 1.4, 2.3, and 2.6 kcal/mol at 300 K, respectively, compared to experimental results for the thermodynamic cycles of CCGG → iCiCiGiG, GGCC → iGiGiCiC, GCGC → iGiCiGiC, and CGCG → iCiGiCiG. In contrast, unmodified AMBER99 predictions for GGCC → iGiGiCiC and GCGC → iGiCiGiC differ from experiment by 11.7 and 12.6 kcal/mol, respectively. In order to test the dynamic stability of the above duplexes with AMBER99TOR, four individual 50 ns molecular dynamics (MD) simulations in explicit solvent were run. All except r(CCGG)$_2$ retained A-form conformation for ≥82% of the time. This is consistent with NMR spectra of r(iGiGiCiC)$_2$, which reveal an A-form conformation. In MD simulations, r(CCGG)$_2$ retained A-form conformation 52% of the time, suggesting that its terminal base pairs may fray. The results indicate that revised backbone parameters improve predictions of RNA properties and that comparisons to measured sequence dependent thermodynamics provide useful benchmarks for testing force fields and computational methods.

1. INTRODUCTION

RNA has a wide variety of biological roles in cells.[1] The genome of some human viruses, such as hepatitis papilloma virus (HPV), human immunodeficiency virus (HIV), smallpox and influenza viruses, is RNA. Messenger RNA (mRNA) carries the code for protein synthesis. Transfer RNAs (tRNA) bring specific amino acids to ribosomes for protein synthesis. Some RNAs, including ribosomal RNA (rRNA), are catalysts.[2−4] MicroRNAs (miRNA) regulate gene expression.[5,6] More functions of RNA are still being discovered.

The ability of theoretical and computational approaches to reproduce experimental results provides a test of our understanding of the interactions that shape RNA.[7] Molecular dynamics (MD) simulations and quantum mechanical (QM) calculations are used to provide insight into biological processes, including folding and dynamics of RNA.[8−13] The quality of the MD simulations, however, depends on the parametrization of the force fields.

Force fields can be benchmarked against various types of experimental results.[7] For example, revisions for χ torsional parameters have improved structural predictions of cytidine and uridine,[14] of tetraloop hairpins,[15] and of single-stranded r(GACC).[16] Here, revisions of various torsional parameters are presented and tested against structural and thermodynamic data for duplexes of RNA tetramers.

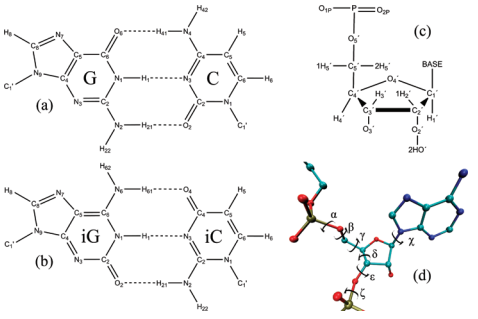

Figure 1. Schematic representations of (a) GC and (b) iGiC base pairs, with atom notations used for (a) guanine (G), cytidine (C), (b) isoguanine (iG), isocytosine (iC), (c) ribose and phosphate, and (d) torsions of nucleic acids.

The unnatural bases isoguanosine (iG) and isocytidine (iC) are similar to the natural bases of guanosine and cytidine except that the amino and carbonyl groups are transposed. They form Watson−Crick-like iGiC base pairs in RNA (Figure 1).[17,18] UV melting experiments show that the free energies of duplex formation at 300 K of structures with iGiC base pairs are more

Received: August 9, 2011
Published: December 01, 2011

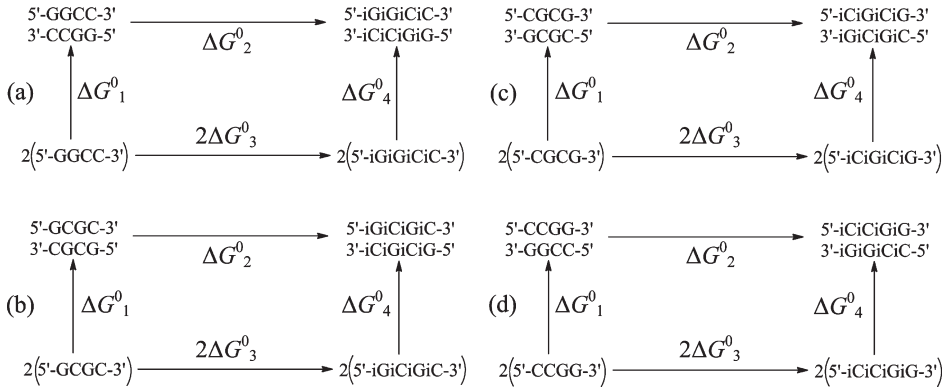

Figure 2. Thermodynamic cycles of (a) GGCC → iGiGiCiC, (b) GCGC → iGiCiGiC, (c) CGCG → iCiGiCiG, and (d) CCGG → iCiCiGiG. ΔG^0_2 and ΔG^0_3 represent alchemical transformations of the duplex and single strand, respectively, while ΔG^0_1 and ΔG^0_4 represent duplex formations. Each cycle satisfies the equation of $\Delta G^0_1 + \Delta G^0_2 = 2\Delta G^0_3 + \Delta G^0_4$, where ΔG^0_1 and ΔG^0_4 are experimental values and ΔG^0_2 and ΔG^0_3 are calculated with the TI approach.

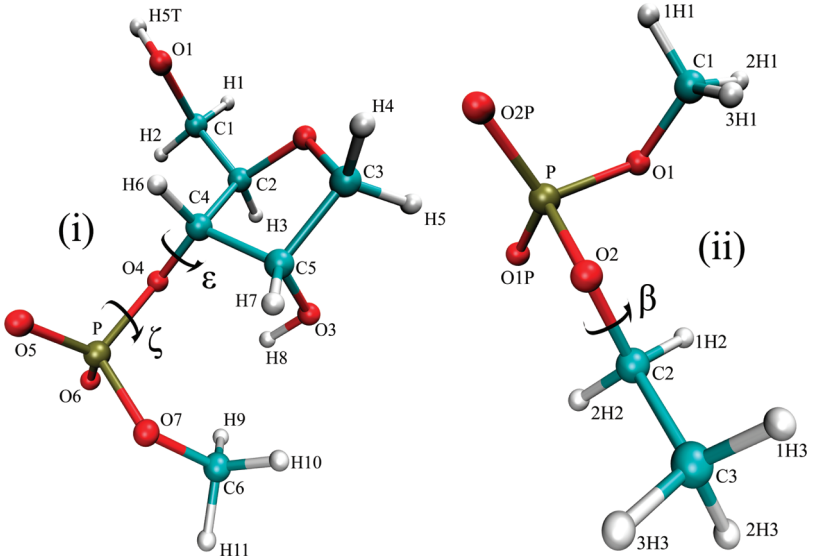

Figure 3. Model systems used to reparameterize torsions ε (C2−C4−O4−P), ζ (C4−O4−P−O7), and β (P−O2−C2−C3). 2D and 1D Potential Energy Surface (PES) scans were done to reparameterize ε and ζ, and β using model systems (i) and (ii), respectively.

favorable than the structures with GC base pairs.[18] In this paper, r(iGiGiCiC)$_2$ is shown by NMR to have an A-form conformation. Previous NMR and optical melting studies of r(CCGGp)$_2$, where p represents a terminal phosphate, are also consistent with an A-form duplex conformation.[19−21] These results were generalized for modeling the duplexes of r(iCiCiGiG)$_2$, r(iCiGiCiG)$_2$, r(iGiCiGiC)$_2$, r(CCGG)$_2$, r(CGCG)$_2$, r(GCGC)$_2$, and r(GGCC)$_2$ as A-form structures to allow free energy calculations using the thermodynamic integration (TI)[22] approach with thermodynamic cycles shown in Figure 2. The torsional parameters for ε, ζ, and β were reparameterized and free energy calculations were made with AMBER99[23] modified with various combinations of parameters for the α/γ,[24] β, ε/ζ, and χ[14] torsions. The version of the AMBER99 force field with revised parameters for all six torsions, which we call AMBER99TOR, improves predictions of differences in experimental free energy changes of duplex formation when iGiC pairs replace GC pairs.

2. METHODS

2.1. Synthesis and Purification of iGiGiCiC. Phosphoramidites for iC and iG were prepared as described previously.[18] The oligoribonucleotide, iGiGiCiC, was synthesized on an Applied Biosystems DNA/RNA synthesizer, using β-cyanoethyl phosphoramidite chemistry.[25,26] Thin-layer chromatography (TLC) purification of iGiGiCiC was carried out on Merck 60 F254 TLC plates with the mixture 1-propanol/aqueous ammonia/water = 55:35:10 (v/v/v). The details of deprotection and purification of oligoribonucleotides have been described previously.[27,28]

2.2. NMR. The concentration of the sample was measured with a NanoDrop 2000 Micro-Volume UV−vis spectrophotometer. The NMR sample had 1.65 mM iGiGiCiC in 80 mM NaCl, 10 mM sodium phosphate, and 0.5 mM disodium EDTA at pH 7.0. For spectra in D_2O, two lyophilizations were performed on the sample, reconstituting each time with 99.9% D_2O (Cambridge Isotopes Laboratories), followed by a third lyophilization and reconstitution in 99.990% D_2O (Sigma Aldrich).

dx.doi.org/10.1021/ct200557r | J. Chem. Theory Comput. 2012, 8, 172–181

Table 1. Conformations Used in 2D ε/ζ PES Scan of Model System (i) in Figure 3

conformation	sugar pucker	H5T−O1−C1−C2 (deg)	O1−C1−C2−C4 (deg)	O4−P−O7−C6 (deg)	C3−C5−O3−H8 (deg)
(i)	C2′-endo	174	54	60	−61
(ii)	C2′-endo	174	54	180	−61
(iii)	C2′-endo	174	54	300	−61
(iv)	C3′-endo	174	54	60	−153
(v)	C3′-endo	174	54	180	−153
(vi)	C3′-endo	174	54	300	−153

All spectra were acquired on Varian Inova 500 or 600 MHz NMR spectrometers. Resonances were assigned by standard procedures[29,30] from NOESY, Watergate NOESY, $^{1}H-^{31}P$ HETCOR, DQF-COSY, and TOCSY at 0, 20, and 35 °C (see Supporting Information). NOESY spectra were recorded with mixing times (τ_m) of 100, 150, 200, and 400 ms.

2.3. Parametrization. RESP charges for C, G, iC, and iG were calculated as previously described (see Supporting Information).[31] For C and G, the revised χ torsion parameters of AMBER99χ were used.[14] The same methodology using Gaussian03[32] was applied to reparameterize the χ torsions of iC and iG (see Supporting Information for the parameters). The ε and ζ torsions were reparameterized on the basis of 2D potential energy surface (PES) scans on six conformations of model system (i) (Figure 3), defined in Table 1. For each conformation, ε and ζ torsions were rotated with increments of 10°, yielding $6 \times (36 \times 36) = 7776$ data points for ε/ζ reparameterization. Model system (ii) (Figure 3) was used to reparameterize the β torsion. β torsions were rotated with increments of 10°, yielding 36 data points for β reparameterization. For each conformation in the PES scan, the structures were first optimized with HF/6-31G* level of theory. Then, QM energies were calculated with MP2/6-31G* level of theory. Comparisons between the AMBER99 and revised ε/ζ and β torsional parameters are shown in Table 2. For α and γ, torsional parameters of the parmbsc[24] force field were used. AMBER99TOR is defined as the AMBER99 force field including all the revised parameters for α/γ,[24] β, ε/ζ, and χ[14] torsions.

2.4. Thermodynamic Cycles. Thermodynamic cycles of CCGG → iCiCiGiG, CGCG → iCiGiCiG, GCGC → iGiCiGiC, and GGCC → iGiGiCiC (Figure 2) were used to test free energy calculations with the TI approach. In Figure 2, ΔG^0_1 and ΔG^0_4 are the experimental free energies of duplex formation with GC and iGiC base pairs, respectively. ΔG^0_2 and ΔG^0_3 are the free energies of the alchemical transformations of duplexes and single strands, respectively, from G and C to iG and iC bases. The TI approach with the new mixing rule described previously[31] was used to calculate ΔG^0_2 and ΔG^0_3. Each cycle satisfies $\Delta G^0_1 + \Delta G^0_2 = 2\Delta G^0_3 + \Delta G^0_4$.

2.5. Explicit Solvent Simulations. All structures were created with the nucgen module of AMBER9. Structures were solvated with TIP3P[33] water molecules in a truncated octahedral box. In each $S_{G/C} \rightarrow S_{iG/iC}$ alchemical transformation, where $S_{G/C}$ and $S_{iG/iC}$ represent states with G and C and iG or iC bases, respectively, each state had the same number of water molecules (see Supporting Information for the number of water molecules used in each calculation). A total of six and three Na^+ ions were used to neutralize the duplex and single-stranded RNA systems, respectively. The parameter/topology files for each $S_{G/C} \rightarrow S_{iG/iC}$ transformation were created with the xleap module.[23]

Table 2. Comparison of Revised $\varepsilon, \zeta, \beta$ Torsional Parameters with AMBER99 Counterparts

torsion	n[a]	AMBER99 $V_{n,i}$[a]	AMBER99TOR $V_{n,i}$[a]
ε	1	0.000	−1.494
	2	0.000	−0.714
	3	0.383	−0.161
	4	0.000	0.121
ζ	1	0.000	−0.561
	2	1.200	0.575
	3	0.250	−0.997
	4	0.000	−0.078
β	1	0.000	−2.598
	2	0.000	0.011
	3	0.383	−0.322
	4	0.000	−0.082

[a] Torsional potential energy in AMBER force field is calculated as $E_{MM,tor}(\phi) = \Sigma_{i=1}^{4} V_{n,i}\,(1 + \cos(n\phi - \gamma))$ where $V_{n,i}$ is the relative potential energy barrier, ϕ is the dihedral angle, γ is the phase shift, and n is the periodicity. For ε, ζ, and β torsions, $\gamma = 0$ (see Supporting Information for the modified force field file).

2.5.1. Minimization. Structures were minimized in two steps. For each system, the same protocol was used: (1) RNA structures were held fixed with a restraint force of 500 kcal/mol Å^2. Steepest descent minimization of 1000 steps was followed by a conjugate gradient minimization of 1500 steps. The long-range cutoff for nonbonded interactions during minimizations was 8.0 or 10.0 Å. (2) The whole system was minimized without any restraints. Steepest descent minimization of 1000 steps was followed by a conjugate gradient minimization of 1500 steps.

2.5.2. Pressure Regulation. After the minimization, two steps of pressure equilibration were done on each system: (1) RNA structures were held fixed with a restraint force of 10 kcal/mol Å^2. Constant volume dynamics with a cutoff of 8.0 or 10.0 Å was used. SHAKE[34] was turned on for bonds involving hydrogen atoms, except for the amino hydrogen and dummy atoms. Temperature was raised from 0 to 300 K in 20 ps. Langevin dynamics with a collision frequency of 1 ps^{-1} was used. Ten thousand MD steps were run with a 2 fs time step, yielding a total of 20 ps of MD. (2) The same conditions as above were chosen, except that no restraints on the structures and constant pressure dynamics were used. Reference pressure was set to 1 atm with a pressure relaxation time of 2 ps. A total of 100 ps of MD was run with a 2 fs time step. The final restart file was used as the starting structure

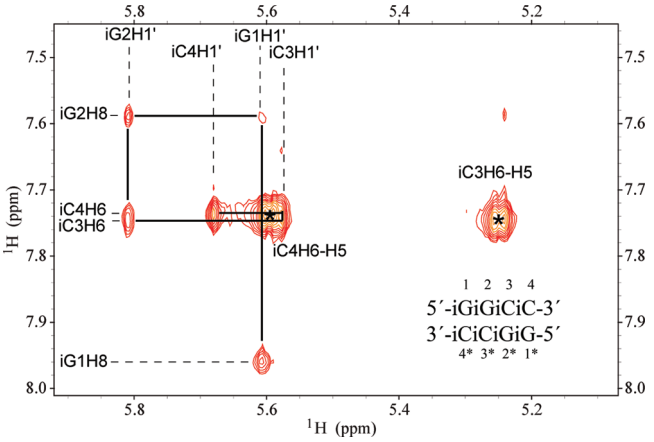

Figure 4. NOESY walk of $(iGiGiCiC)_2$ from 200 ms mixing time NOESY experiment at 20 °C. Because the sequence is symmetric, the cross-peaks from each stand overlap. Residue numberings are shown in the bottom right corner.

for the λ simulations. In the AMBER99 force field simulations, constant volume dynamics were used.

2.5.3. λ Simulations. Nineteen λ values were used; $\lambda = 0.05$ to $\lambda = 0.95$, with an increment of 0.05. The new mixing rule of TI approach was used in all λ simulations.[31] For each λ simulation, the last structure of pressure regulation was taken as the initial structure. The production run was similar to the second step of the pressure equilibration described above. Duplex and single-strand MD simulations, respectively, were run for 2 and 3 ns with 1 fs time steps.

2.5.4. Restrained λ Simulations. Additional MD simulations used dihedral restraints to restrict the sampling space to A-form conformations (see Supporting Information). A total of 1 ns of MD was run with a 1 fs time step for both the duplex and single-strand simulations. Further calculations used positional restraints with weight of 10 kcal/mol Å^2 on the backbone heavy atoms in single-strand simulations.

2.5.5. Dynamic Stabilities of RNA Duplexes with AMBER99-TOR Force Field. In order to analyze the dynamic stability of each duplex and single-strand with the AMBER99TOR force field, MD simulations in explicit solvent were run. Systems were prepared similar to part 2.5. Six and three Na^+ ions were used to neutralize duplex and single-strand systems, respectively. Duplex and single-strand systems were solvated with 1786 and 1345 TIP3P[33] water molecules, respectively, in a truncated octahedral box. The systems were minimized and pressure regulated as described above. Each production run included 50 ns of MD with 1 fs time step at 300 K. Trajectory files were written at each 500 fs time step. Four individual simulations were run for each system, yielding a total of 200 ns of MD.

2.6. Analysis. *Free Energy Calculations Using TI Approach.* The first 250 ps of each λ simulation were omitted from the calculations. For each λ simulation, $\langle \partial E/\partial\lambda \rangle_\lambda$ was calculated. The trapezoidal rule was used to numerically integrate $\langle \partial E/\partial\lambda \rangle_\lambda$ vs λ curves to get ΔG^0. Multiple transformations were done to calculate the means and standard deviations (see Supporting Information).

Stability Analysis of Duplex Simulations with AMBER99TOR Force Field. The combined 200 ns of MD simulations were analyzed to test the dynamic stability of each duplex and single strand (see Supporting Information). All the trajectory data were rmsd fitted to the initial A-form starting structure. For each simulation, the ptraj module of AMBER 9[23] was used to calculate the percentage of structures having an all-atom rmsd less than 1.5 and 3.0 Å. Qualitatively, A-form and A-form-like structures are defined as conformations with rmsd less than 1.5 Å and 3.0 Å, respectively. Total overlap area of the stacked base pairs were calculated with 3DNA[35] using snapshots extracted from the trajectories at intervals of 0.5 ns.

3. RESULTS

3.1. Conformation of $r(iGiGiCiC)_2$. Because the electronic structure of iCiG base pairs differs from CG,[7] NMR was used to test the expectation that $r(iGiGiCiC)_2$ has an A-form conformation. Figure 4 shows the NOESY walk region of $(iGiGiCiC)_2$ from a 200 ms mixing time NOESY experiment in D_2O at 20 °C. NMR distance limits were extracted from NOESY spectra at 20 and 35 °C with 200 ms mixing time using intranucleotide H1′/H2′ cross-peaks as reference NOEs (see Supporting Information).

At 1.65 mM iGiGiCiC, there is an iG1H1′/ iC4H2′ cross-peak. This cross-peak disappeared when the iGiGiCiC concentration was diluted to 0.17 mM (see Supporting Information) and is due to coaxial stacking of duplexes. The rest of the spectrum was essentially unchanged at the lower concentration.

The NMR spectra of iGiGiCiC are consistent with an A-form duplex conformation. All the residues prefer C3′-endo sugar pucker as evidenced by $^3J_{H1'-H2'}$ couplings of less than 2 Hz as estimated from peak splittings. Intranucleotide iGH8/H1′ and iCH6/H1′ cross-peaks have volumes indicating anti conformations. A Watergate NOESY spectrum at 0 °C with 150 ms mixing time showed iG1H61−iC3H5 and iG2H1−iC4H1′ cross-peaks consistent with a duplex structure in which these peaks are actually iG1H61−iC3*H5, iG1*H61−iC3H5, iG2H1−iC4*H1′, and iG2*H1−iC4H1′, where an asterisk represents the opposite RNA strand (see Supporting Information). The chemical shifts of the iG imino protons (Supporting Information) are consistent with hydrogen bonding, as observed for other iCiG pairs.[17,18] Separate resonances are seen for the two protons of the iG amino groups with one of the shifts consistent with hydrogen bonding.[17,18] Another expectation for a "Watson−Crick" iGiC pair is a cross-peak from iGH1 to both protons of an iC amino group, and iG2 shows such cross-peaks to two broad resonances. A HETCOR spectrum showed phosphorus shifts within 0.3 ppm, implying regular A-form conformation (see Supporting Information). The HETCOR spectrum also showed strong (n)P−(n-1)H3′ and weak (n)P−(n)H5′/5″ scalar coupling typical of A-form ε and β conformations and weak H4′−H5′/5″ scalar coupling consistent with A-form γ conformation. Distances measured for the nucgen model of $(iGiGiCiC)_2$ are consistent with the distance limits calculated from NOEs, with the exception of a 0.15 Å difference for iG2*H3′−iG2*H8 (see Supporting Information). On the basis of the NMR spectra for $r(iGiGiCiC)_2$, A-form conformations were also modeled for $r(iCiCiGiG)_2$, $r(iCiGiCiG)_2$, and $r(iGiCiGiC)_2$.

3.2. Comparisons of Molecular Mechanics (MM) to QM Energies before and after ε/ζ and β Reparameterizations. Model systems (i) and (ii) (Figure 3) were used to reparameterize the ε/ζ and β torsional parameters, respectively. A force field with the new ε/ζ parameters is called AMBER99EZ. Figure 5 shows comparisons of the 2D potential energy surfaces approximated by AMBER99EZ (see Supporting Information for the definitions) and AMBER99 force fields with the QM potential energy surfaces for six conformations (see Supporting Information).

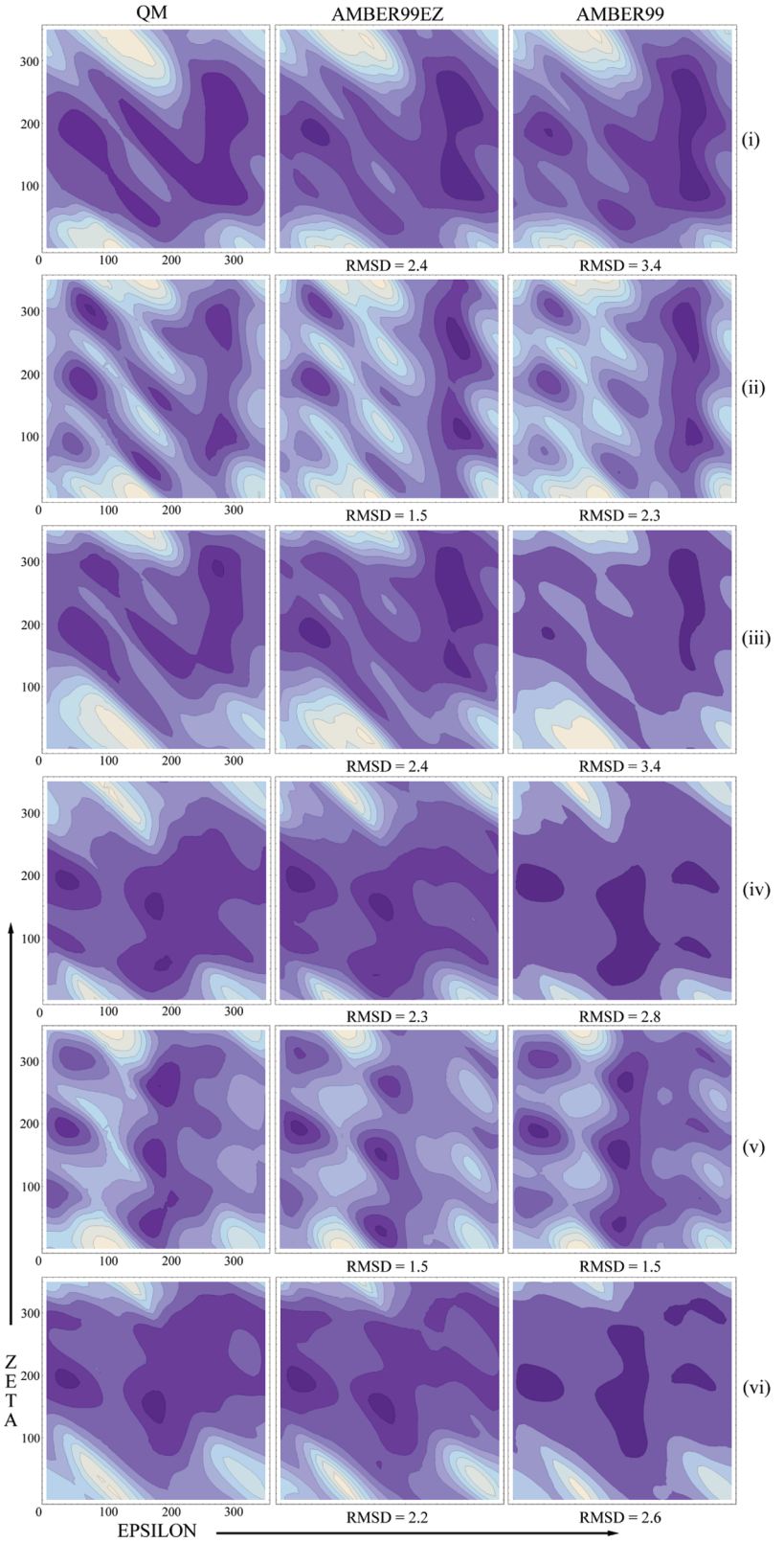

Figure 5. Two-dimensional PES scans of ε (x-axis) and ζ (y-axis) for six conformations described in Table 1. QM, AMBER99, and AMBER99EZ stand for PES scans of the conformations (i−vi) using quantum mechanics, AMBER99, and reparameterized ε/ζ torsional set of AMBER99 force field, respectively. rmsd values (kcal/mol) under each PES scan are with respect to QM. The darker the violet color, the lower the energy value.

Figure 6 shows the 1D potential energy surfaces for model system (ii) of Figure 3 as calculated by QM, AMBER99 with revised β torsional parameters, and AMBER99. It is clear from Figures 5 and 6 that revisions of the torsional parameters improve the

dx.doi.org/10.1021/ct200557r |J. Chem. Theory Comput. 2012, 8, 172–181

approximations of the QM potential energy surfaces for these model systems. The revision of the AMBER99 force field uses four cosine terms to describe each of the torsional energy profiles of ε/ζ and β, while the original AMBER99 force field uses one cosine term each for ε and β, and two cosine terms for ζ to describe the torsional energy profiles (Table 2). Increasing the number of cosine terms evidently improves the predictions of QM PES profiles.

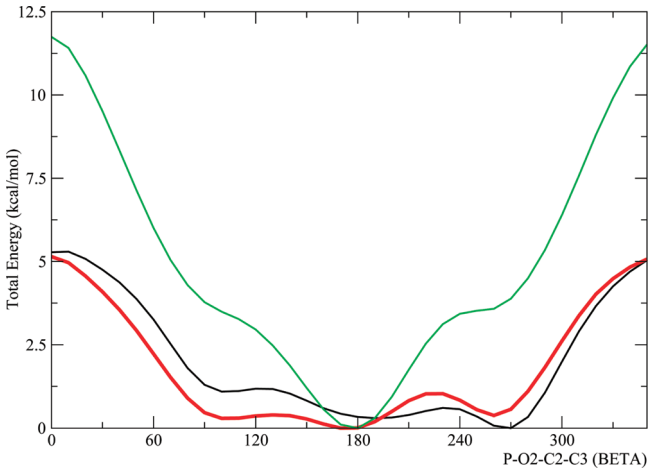

Figure 6. Potential energy (kcal/mol) vs β dihedral angle (P−O2−C2−C3) of model system (ii) (Figure 3) with QM (black), AMBER99 with revised β torsion parameter (red), and AMBER99 (green). For visualization purposes, minimum energies are set to zero.

3.3. Comparisons between Measured and Calculated Changes in Free Energies of Duplex Formation by CG and iCiG Sequences to Benchmark the Effects of Revising Torsional Parameters. As illustrated in Figure 2, experimental measurements[7,18,19,28,36,37] provide values for the free energy changes, ΔG^0_1 and ΔG^0_4, of formation of duplexes with CG and iCiG pairs, respectively. The values of ΔG^0_2 and ΔG^0_3 for the alchemical transitions in Figure 2 can be calculated with the TI approach,[22,31] where ΔG^0_3 represents the change for one single strand morphing from C and G to iC and iG nucleotides. From the thermodynamic cycles in Figure 2, $\Delta G^0_4 - \Delta G^0_1 = \Delta G^0_2 - 2\Delta G^0_3$. Thus comparisons of experimental values for $\Delta G^0_4 - \Delta G^0_1$ with calculated values for $\Delta G^0_2 - 2\Delta G^0_3$ provide benchmarks for the effects of revising force field parameters. This is a rather stringent test because the individual values for ΔG^0_2 and $2\Delta G^0_3$ are on the order of 200 kcal/mol (see Supporting Information), while the experimental values for $\Delta G^0_4 - \Delta G^0_1$ are on the order of a few kcal/mol (Table 3).

Table 3 shows results from unrestrained and restrained simulations with several combinations of torsion parameters. Initial results for unrestrained simulations with AMBER99 on the GCGC → iGiCiGiC and GGCC → iGiGiCiC cycles gave values for $\Delta G^0_2 - 2\Delta G^0_3$ of 8.7 and 10.5 kcal/mol, respectively, whereas the experimental values for $\Delta G^0_4 - \Delta G^0_1$ are −3.0 and −2.1 kcal/mol, respectively. In contrast, AMBER99 simulations with the backbone torsions restrained to A-form gave $\Delta G^0_2 - 2\Delta G^0_3$ values of −1.4 and −1.5 kcal/mol, respectively. When the positions of backbone heavy atoms were restrained, the values were −3.0 and −2.4 kcal/mol, respectively, close to the experimental values. The lack of agreement for unrestrained AMBER99

Table 3. Free Energy Results (kcal/mol at 300 K) of Unrestrained and Restrained Simulations for the Thermodynamic Cycles of GCGC → iGiCiGiC, GGCC → iGiGiCiC, CGCG → iCiGiCiG, and CCGG → iCiCiGiG Using TI Approach with Revised Torsions for AMBER99 Force Field

	$\Delta G^0_2 - 2\Delta G^0_3$					
	revised torsions					
thermodynamic cycle	none[b]	χ[b]	$\chi\alpha\gamma$[b]	$\chi\alpha\gamma\varepsilon\zeta$[b]	$\chi\alpha\gamma\varepsilon\zeta\beta$[b]	$\Delta G^0_4 - \Delta G^0_1$[a]
			Unrestrained Simulations			
GCGC → iGiCiGiC	8.7	−1.2 ± 0.4	−0.7 ± 1.2	−1.3 ± 1.0	−0.7 ± 1.0	−3.0 ± 0.4
GGCC → iGiGiCiC	10.5	−1.3 ± 0.6	0.1 ± 2.4	−6.2 ± 2.4	−3.5 ± 2.4	−2.1 ± 0.4
CGCG → iCiGiCiG	–	−0.8 ± 1.6	0.0 ± 0.4	0.2 ± 1.9	0.4 ± 1.0	−2.2 ± 0.4
CCGG → iCiCiGiG	–	4.5 ± 1.6	1.7 ± 1.0	2.1 ± 3.4	0.9 ± 2.4	−0.4 ± 0.3
rmsd[c]	12.2	2.7	2.2	2.8	2.0	–
			Restrained Simulations[d]			
GCGC → iGiCiGiC	−1.4 (−3.0)	–	–	−1.8 ± 0.4	−1.5 ± 0.2	−3.0 ± 0.4
GGCC → iGiGiCiC	−1.5 (−2.4)	–	–	−1.4 ± 0.2	−1.4 ± 0.2	−2.1 ± 0.4
CGCG → iCiGiCiG	−1.1	–	–	−1.2 ± 0.4	−1.0 ± 0.2	−2.2 ± 0.4
CCGG → iCiCiGiG	−0.4	–	–	−0.5 ± 0.4	−0.5 ± 0.4	−0.4 ± 0.3
rmsd[c]	1.0			0.9	1.0	–

[a] These values are experimental results at 300 K.[18] Error limits assume ±4% error for each ΔG^0.[28] [b] none = AMBER99, χ = AMBER99χ,[14] $\chi\alpha\gamma$ = AMBER99χ[14] + parmbsc,[24] $\chi\alpha\gamma\varepsilon\zeta$ = AMBER99χ + parmbsc + AMBER99EZ, $\chi\alpha\gamma\varepsilon\zeta\beta$ = AMBER99TOR (see Supporting Information for the definitions of these force fields). [c] rmsd = $({}^1/_4\Sigma_{i=1}^4(\Delta G^0_{\text{calculated},i} - \Delta G^0_{\text{measured},i})^2)^{1/2}$ where $\Delta G^0_{\text{calculated}} = \Delta G^0_2 - 2\Delta G^0_3$, $\Delta G^0_{\text{measured}} = \Delta G^0_4 - \Delta G^0_1$, and i stands for results of each thermodynamic cycle. [d] Values not in parentheses are for simulations with dihedral restraints (see Supporting Information). Values in parentheses are simulations with positional restraints. Restraint weight of 10 kcal/mol Å^2 was applied to backbone heavy atoms in single-strand MD simulations to force them to sample around A-form conformations. No restraints were used in duplex simulations for these calculations.

dx.doi.org/10.1021/ct200557r | J. Chem. Theory Comput. 2012, 8, 172–181

Table 4. All-Atom RMSD (Å) Results and Total Overlap Area of the Stacked Base Pairs for Duplex Simulations of $(CCGG)_2$, $(CGCG)_2$, $(GCGC)_2$, $(GGCC)_2$, $(iCiCiGiG)_2$, $(iCiGiCiG)_2$, $(iGiCiGiC)_2$, and $(iGiGiCiC)_2$ with AMBER99TOR[a]

duplex	≤1.5 (%)	≤3.0 (%)	overlap area[b] (Å^2)	duplex	≤1.5 (%)	≤3.0 (%)	overlap area[b] (Å^2)	$\Delta G^0_4 - \Delta G^0_1$[c]
$(GCGC)_2$	39 ± 29	98 ± 5	10.6 ± 0.3	$(iGiCiGiC)_2$	64 ± 18	100 ± 0	10.4 ± 0.1	−3.0
$(GGCC)_2$	47 ± 16	100 ± 0	7.1 ± 0.1	$(iGiGiCiC)_2$	85 ± 6	100 ± 0	6.7 ± 0.0	−2.1
$(CGCG)_2$	28 ± 15	95 ± 6	7.0 ± 0.0	$(iCiGiCiG)_2$	22 ± 12	82 ± 28	6.5 ± 0.1	−2.2
$(CCGG)_2$	15 ± 6	51 ± 33	4.5 ± 0.0	$(iCiCiGiG)_2$	29 ± 16	95 ± 10	4.8 ± 0.1	−0.4

[a] For each duplex, four individual MD simulations of 50 ns were run at 300 K, yielding a total of 200 ns. Structures were saved every 0.5 ps for rmsd analysis. [b] Total overlap area of the stacked base pairs excluding exocyclic groups was calculated with 3DNA[35] using all the snapshots extracted from the trajectories at intervals of 0.5 ns. [c] ΔG^0_1 and ΔG^0_4 are duplex formation free energies of structures with GC and iGiC base pairs, respectively, at 300 K.[18]

calculations suggests poor sampling of the conformational space. These results suggested that revisions of torsional parameters could improve agreement between calculations and experiments.

To test the effects of adding revised torsional parameters, unrestrained TI calculations were done on all four systems shown in Figure 2. Specifically, unrestrained calculations were done with AMBER99χ[14], AMBER99χ with α/γ parameters from parmbsc[24] ($\chi\alpha\gamma$), and with further revision of parameters for ε/ζ ($\chi\alpha\gamma\varepsilon\zeta$), or ε/ζ and β ($\chi\alpha\gamma\varepsilon\zeta\beta$, AMBER99TOR) as developed here (Table 3).

As shown in Table 3, all revisions improved agreement between predictions and experiments relative to AMBER99 calculations for GCGC → iGiCiGiC and GGCC → iGiGiCiC. Revision of χ parameters provided a large improvement of 10−12 kcal/mol at 300 K. Revisions for other dihedral parameters were tested against experimental results for all four thermodynamic cycles shown in Figure 2. AMBER99TOR gave the best rmsd of 2.0 kcal/mol between predictions and experiment, but AMBER99χ mixed with α/γ parameters taken from parmbsc was similar with an rmsd of 2.2 kcal/mol (Table 3).

Relatively large error limits of 2.4 kcal/mol were found for AMBER99TOR calculations of GGCC → iGiGiCiC and CCGG → iCiCiGiG (Table 3). Nevertheless, the calculated values of $\Delta G^0_2 - 2\Delta G^0_3$ for these transformations are within 1.4 kcal/mol of the experimental $\Delta G^0_4 - \Delta G^0_1$, well within experimental error. In contrast, the calculated values for GCGC → iGiCiGiC and CGCG → iCiGiCiG have error limits of 1.0 kcal/mol, but values of $\Delta G^0_2 - 2\Delta G^0_3$ differ from $\Delta G^0_4 - \Delta G^0_1$ by 2.3 and 2.6 kcal/mol, respectively. These comparisons suggest a difference in the behavior of sequences with 5′GG/3′CC and 5′iGiG/3′iCiC nearest neighbors and those without adjacent G's. Root-mean-square deviation analysis of each λ simulation with AMBER99TOR showed that all the duplex transformations (corresponding to ΔG^0_2 in Figure 2) sample pure A-form conformations over 80% of the time while the single-strand transformations (corresponding to ΔG^0_3 in Figure 2) behave differently for sequences with 5′GG/3′CC nearest neighbors (see Supporting Information). The single-strand transformations of CGCG → iCiGiCiG and GCGC → iGiCiGiC sample A-form conformations 47% of the time on average, whereas the single-strand transformations of CCGG → iCiCiGiG and GGCC → iGiGiCiC sample A-form only 21% of the time on average (see Supporting Information). As a result, errors for the thermodynamic cycles of CCGG → iCiCiGiG and GGCC → iGiGiCiC are 2.4 kcal/mol while they are 1.0 kcal/mol for the cycles of CGCG → iCiGiCiG and GCGC → iGiCiGiC (Table 3). The more the single-strands sample A-form conformations in a thermodynamic cycle, the lower the error.

A study of the ability of AMBER99 to predict experimentally observed[17] relative populations of sheared and imino hydrogen bonded GA pairs in RNA duplexes found that the best agreement required restraining backbones to be similar to those known for the NMR structures.[31] As described above, restraining backbones to A-form conformations dramatically increased agreement of AMBER99 calculations with experiment. To test the effects of dihedral restraints on revised versions of AMBER99, calculations were done with AMBER99TOR and AMBER99χ with α/γ and ε/ζ revisions (Table 3). In both cases, agreement with experiment was improved with RMSDs of 1.0 and 0.9 kcal/mol, respectively, and error limits were reduced to 0.4 or fewer kcal/mol. Even with dihedral restraints, however, experimental and calculated values sometimes differ beyond error limits. The results suggest that the force field can be refined further. It is encouraging, however, that the CCGG → iCiCiGiG transformation is predicted to have the smallest $\Delta G^0_2 - 2\Delta G^0_3$, which agrees with experiment.

3.4. Predicted Dynamic Stabilities with AMBER99TOR Force Field. The predicted dynamic stability of duplexes provides another test of force fields. For each duplex, four individual MD simulations of 50 ns each were run with the AMBER99TOR force field and then combined for an rmsd analysis. The percentage of structures having all-atom rmsd less than 1.5 or 3.0 Å were calculated (Table 4). rmsd ≤1.5 Å and rmsd ≤3.0 Å represent, respectively, essentially the starting A-form conformation and A-form-like conformations.

Table 4 shows the percentage of structures in A-form and A-form-like conformations for the duplexes over 200 ns of unrestrained MD with the AMBER99TOR force field. All duplexes except $r(CCGG)_2$ spent more than four-fifths of time in A-form and A-form-like conformations. The results suggest that $r(CCGG)_2$ may have unusual dynamics for its terminal base pairs, e.g., "fraying". The MD simulations for the other sequences indicate that A-form-like conformations are essentially stable for at least 50 ns with the AMBER99TOR force field. MD simulations of single-strands show that A-form conformations are sampled rarely while even A-form-like conformations are sampled less than 60% of time (see Supporting Information).

3.5. Comparison between Predicted Values of β and Those Observed in Crystal Structures. QM calculations on model system (ii) in Figure 3 predict a shallow dependence of energy on the β dihedral angle (Figure 6). Crystal structures of RNA, however, show a strong preference for $\beta \sim 180°$,[38,39] as does the AMBER99 potential (Figure 6). Histogram analysis of unrestrained 50 ns MD simulations with AMBER99TOR that correspond to a total of 3.2 μs simulation time shows two populations preferred by β (Figure 7). The dominant region has β around 180°, while the minor region is around 80°. While low

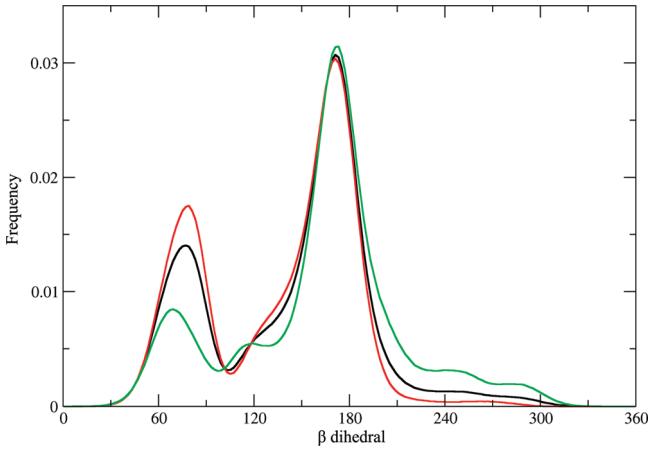

Figure 7. Population distribution analysis for β torsion of MD simulations with AMBER99TOR. Black, red, and green curves represent results including all, duplex, and single-strand MD simulations, respectively. Terminal β torsions (free 5′OH) were omitted from calculations.

values of β are rarely seen in RNA crystals, they are seen in RNA S-motifs.[38]

The β dihedral is coupled with α and γ torsions. Cluster analysis of the unrestrained MD simulations showed that there are three (α, β, γ) populations; 66% in (300°, 180°, 60°), 17% in (180°, 80°, 180°), and 13% in (300°, 80°, 180°). Even though there are three regions preferred by (α, β, γ), 3D structures created by these combinations are similar (see Supporting Information). Crystal structures analyzed by others[38,39] are much bigger systems compared to tetramers discussed here. This might explain why we see 30% of structures with low β values. It is also possible that this low β region might have importance for backbone dynamics such as in base unstacking and base pair opening.

3.6. Comparison between Sequence Dependence of Free Energy Differences and Overlap of Bases. There is a parallelism between $\Delta G^0_4 - \Delta G^0_1$ values for the cycles shown in Figure 2 and total overlap areas of stacked base pairs for the duplexes studied (Table 4). Total overlap areas of stacked base pairs for $(GCGC)_2$ and $(iGiCiGiC)_2$, $(GGCC)_2$ and $(iGiGiCiC)_2$, $(CGCG)_2$ and $(iCiGiCiG)_2$, $(CCGG)_2$ and $(iCiCiGiGi)_2$ are around 10.5, 6.9, 6.8, and 4.7 Å^2, respectively, in parallel with experimental $\Delta G^0_4 - \Delta G^0_1$ results for the corresponding thermodynamic cycles (Table 4). This suggests that the thermodynamic differences between duplexes with CG and iCiG pairs may be primarily due to differences in stacking interactions, which result from different electron distributions in the ring systems.

4. DISCUSSION

Force fields for proteins have proven to be extremely useful for providing insight into protein folding, function, and design.[40−43] Much less effort, however, has been applied to development and testing of force fields for RNA. The emerging recognition that RNA has many different cellular functions[1−6] and that many RNAs are potential therapeutic targets[27,44−47] increases the importance of force fields for RNA.

A key aspect of RNA is base orientation with respect to sugar. This orientation is controlled by the χ torsions, which define whether a nucleotide is in syn or anti conformation. Revision of χ torsion parameters to give the AMBER99χ force field improves structural and thermodynamic predictions for cytidine and uridine[14] and structural predictions for tetraloop hairpins.[15] For example, AMBER99 prefers syn base orientation for pyrimidines, while AMBER99χ prefers anti base orientation.[14] Additionally, AMBER99 prefers either syn or high-anti base orientations for purines, while AMBER99χ prefers syn or anti base orientations.[14] Structural analysis of single-stranded r(GACC) with NMR showed that it prefers A-form-like conformations. AMBER99, however, rapidly generates random coil conformations for r(GACC) while AMBER99χ prefers A-form-like conformations for most of the first 700 ns of an unrestrained MD simulation.[16] After 700 ns, however, a stable conformation and random coil ensemble were generated that are inconsistent with NMR spectra.

A different reparameterization of χ[48] has been tested along with AMBER99χ for ability to maintain known RNA structures during unrestrained MD simulations.[15,48] Both revisions performed better than AMBER99 and similarly to each other even though there were many differences in the details of methods used for parametrization.[14,48] Thus, there is consensus that parameters for χ are important for accurately modeling RNA.

Here, various versions of AMBER99 are developed and benchmarked against measured differences in the free energies of duplex formation by tetramers with GC or iGiC base pairs.[18,19,36,37] Comparisons between measurements and computations are based on the thermodynamic cycles shown in Figure 2, and the results are listed in Table 3.

Relative to AMBER99, AMBER99χ improves agreement between experiments and predictions for the GCGC → iGiCiGiC and GGCC → iGiGiCiC cycles by about 11 kcal/mol at 300 K, corresponding to a 10^8 improvement in prediction of relative equilibrium constants. When AMBER99χ is tested against further refined force fields, the best agreement with experiment for all four cycles shown in Figure 2 is found with AMBER99-TOR, which includes α/γ parameters from the parmbsc force field,[24] along with new parameters developed here for ε/ζ and β (Tables 2 and 3). These additional parameters improve the rmsd comparison by 0.7 kcal/mol at 300 K, corresponding to a 3-fold improvement in prediction of relative equilibrium constants. The largest improvement, however, is 3.6 kcal/mol for CCGG → iCiCiGiG at 300 K, corresponding to a 400-fold improvement in relative equilibrium constants.

TI calculations without restraints or with dihedral restraints did not predict the magnitudes of all the experimental results within error limits (Table 3). This implies that approximations can be improved. The free energy differences for formation of duplexes from single strands depend on many interactions, including stacking, hydrogen bonding, and solvent interactions in both single strands and duplexes. For example, treatment of van der Waals interactions may need revision for RNA force fields to better predict experimental results. Free energy difference calculations provide useful benchmarks for testing such force field revisions.

The ability of force fields to maintain known 3D structures is another test. Here, NMR spectra (Figure 4 and Supporting Information) show that $(iGiGiCiC)_2$ is an A-form duplex. All the other duplexes are also expected to be A-form except for occasional fraying of terminal base pairs. As shown in Table 4, results from four unrestrained 50 ns MD simulations for each of the eight duplexes studied are consistent with this expectation. It is encouraging that AMBER99TOR appears to provide reasonable results for both free energy calculations and dynamics of tetramer duplexes containing either GC or iGiC pairs.

The results presented here show that AMBER99χ and AMBER99TOR improve predictions of the sequence dependence of thermodynamics for several tetramer duplexes and that MD simulations with AMBER99TOR usually retain A-form like structure for at least 50 ns. The revised force fields have not been tested on larger RNAs. It would be surprising, however, if they did not also work well for larger RNAs where the accessible folding space is more limited by volume exclusion.

5. CONCLUSION

Differences in stabilities of short RNA duplexes provide tests of computational methods and force fields. The tests are especially stringent because the calculations include single strands, which have conformational flexibility without much restriction from volume exclusion or hydrogen bonding. Comparisons between measured and predicted stabilities of tetramer duplexes with either GC or iGiC base pairs reveal that reparameterization of torsions can improve agreement between experiment and computations by roughly 10 kcal/mol at 300 K, corresponding to an improvement of about 10^7 in relative equilibrium constant. Most of the improvement relies on new parameters for the χ torsion. The new parameters also largely retain A-form like structures in 50 ns long MD simulations. The revised parameters should improve computations of properties for RNA loops. Loops are often important for function and have weaker interactions and more dynamics than stems. The results also indicate, however, that computations can be further improved.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information. Description of RESP charge calculations of model systems and definition of each torsional revision; atom names/types/charges of model systems, and nucleotides G, iG, C, and iC; modified force field files for revised torsional parameters of χ, α/γ, ε/ζ, β, and nucleotides iC and iG; TI results of all the thermodynamic cycles with different revisions of the torsions (restrained/unrestrained calculations); dihedral restraints used in restrained TI calculations; rmsd analysis of each λ simulation of all the thermodynamic cycles with AMBER99TOR; number of water molecules used to solvate the duplex and single-strand structures; NMR distance, Watson−Crick base pairing, dihedral, and chirality restraints used in simulated annealing of iGiGiCiC; NMR chemical shift assignments of r(iGiGiCiC)$_2$; NOESY walk, ^{1}H−^{31}P HETCOR and NOESY spectra of iGiGiCiC; population distribution analysis of torsions; overlap of stable three (α, β, γ) combinations seen in AMBER99TOR MD simulations. 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. Email: turner@chem.rochester.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENT

We are grateful to the Center for Research Computing at the University of Rochester for providing the necessary computing systems (BlueHive cluster) and personnel to enable the research presented in this manuscript, and Jason D. Tubbs for his help with the NanoDrop 2000 UV−vis spectrophotometer. Dr. Ilyas Yildirim thanks the Faculty of Science, Akdeniz University, Antalya, Turkey, for providing a working environment while he was in Turkey. We also thank the AMBER community and mailing list for all their help and support. This work was supported by NIH grant GM22939 (D.H.T.).

■ ABBREVIATIONS USED

C, cytidine; G, guanosine; iC, isocytidine; G, isoguanosine; QM, quantum mechanics; MM, molecular mechanics; MD, molecular dynamics; TI approach, Thermodynamic Integration approach

■ REFERENCES

(1) Atkins, J. F.; Gesteland, R. F.; Cech, T. R. *RNA Worlds: From Life's Origins to Diversity in Gene Regulation*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2010.
(2) Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. *Cell* **1983**, *35*, 849.
(3) Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. *Cell* **1982**, *31*, 147.
(4) Nissen, P.; Hansen, J.; Ban, N.; Moore, P. B.; Steitz, T. A. *Science* **2000**, *289*, 920.
(5) Lee, R. C.; Feinbaum, R. L.; Ambros, V. *Cell* **1993**, *75*, 843.
(6) Ruvkun, G. *Science* **2001**, *294*, 797.
(7) Yildirim, I.; Turner, D. H. *Biochemistry* **2005**, *44*, 13225.
(8) Auffinger, P.; Hashem, Y. *Curr. Opin. Struct. Biol.* **2007**, *17*, 325.
(9) Csaszar, K.; Spackova, N.; Stefl, R.; Sponer, J.; Leontis, N. B. *J. Mol. Biol.* **2001**, *313*, 1073.
(10) Krasovska, M. V.; Sefcikova, J.; Spackova, N.; Sponer, J.; Walter, N. G. *J. Mol. Biol.* **2005**, *351*, 731.
(11) McDowell, S. E.; Spackova, N.; Sponer, J.; Walter, N. G. *Biopolymers* **2007**, *85*, 169.
(12) Ditzler, M. A.; Otyepka, M.; Sponer, J.; Walter, N. G. *Acc. Chem. Res.* **2010**, *43*, 40.
(13) Denning, E. J.; Priyakumar, U. D.; Nilsson, L.; Mackerell, A. D. *J. Comput. Chem.* **2011**, *32*, 1929.
(14) Yildirim, I.; Stern, H. A.; Kennedy, S. D.; Tubbs, J. D.; Turner, D. H. *J. Chem. Theory Comput.* **2010**, *6*, 1520.
(15) Banas, P.; Hollas, D.; Zgarbova, M.; Jurecka, P.; Orozco, M.; Cheatham, T. E.; Sponer, J.; Otyepka, M. *J. Chem. Theory Comput.* **2010**, *6*, 3836.
(16) Yildirim, I.; Stern, H. A.; Tubbs, J. D.; Kennedy, S. D.; Turner, D. H. *J. Phys. Chem. B* **2011**, *115*, 9261.
(17) Chen, G.; Kierzek, R.; Yildirim, I.; Krugh, T. R.; Turner, D. H.; Kennedy, S. D. *J. Phys. Chem. B* **2007**, *111*, 6718.
(18) Chen, X.; Kierzek, R.; Turner, D. H. *J. Am. Chem. Soc.* **2001**, *123*, 1267.
(19) Petersheim, M.; Turner, D. H. *Biochemistry* **1983**, *22*, 256.
(20) Petersheim, M.; Turner, D. H. *Biochemistry* **1983**, *22*, 264.
(21) Petersheim, M.; Turner, D. H. *Biochemistry* **1983**, *22*, 269.
(22) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395.
(23) Case, D. A.; Darden, T. A.; Cheatham, T. E. I.; 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.; Kollman, P. A. *AMBER 9*; University of California San Francisco: San Francisco, CA, 2006.
(24) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817.
(25) Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981**, *22*, 1859.

(26) Kierzek, R.; Caruthers, M. H.; Longfellow, C. E.; Swinton, D.; Turner, D. H.; Freier, S. M. *Biochemistry* **1986**, *25*, 7840.
(27) Testa, S. M.; Disney, M. D.; Turner, D. H.; Kierzek, R. *Biochemistry* **1999**, *38*, 16655.
(28) Xia, T.; SantaLucia, J., Jr.; Burkard, M. E.; Kierzek, R.; Schroeder, S. J.; Jiao, X.; Cox, C.; Turner, D. H. *Biochemistry* **1998**, *37*, 14719.
(29) Varani, G.; Aboulela, F.; Allain, F. H. T. *Prog. Nucl. Magn. Reson. Spectrosc.* **1996**, *29*, 51.
(30) Varani, G.; Tinoco, I. *Q. Rev. Biophys.* **1991**, *24*, 479.
(31) Yildirim, I.; Stern, H. A.; Sponer, J.; Spackova, N.; Turner, D. H. *J. Chem. Theory Comput.* **2009**, *5*, 2088.
(32) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(33) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(34) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(35) Lu, X. J.; Olson, W. K. *Nature Protoc.* **2008**, *3*, 1213.
(36) Freier, S. M.; Burger, B. J.; Alkema, D.; Neilson, T.; Turner, D. H. *Biochemistry* **1983**, *22*, 6198.
(37) Freier, S. M.; Sinclair, A.; Neilson, T.; Turner, D. H. *J. Mol. Biol.* **1985**, *185*, 645.
(38) 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.; Berman, H. M. *RNA* **2008**, *14*, 465.
(39) Schneider, B.; Moravek, Z.; Berman, H. M. *Nucleic Acids Res.* **2004**, *32*, 1666.
(40) Baker, D. *Protein Sci.* **2010**, *19*, 1817.
(41) Ponder, J. W.; Case, D. A. In *Protein Simulations*; Elsevier Academic Press: San Diego, Ca, **2003**; Vol. *66*, p 27.
(42) Salsbury, F. R. *Curr. Opin. Pharmacol.* **2010**, *10*, 738.
(43) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y. B.; Wriggers, W. *Science* **2010**, *330*, 341.
(44) Disney, M. D.; Childs, J. L.; Turner, D. H. *Biopolymers* **2004**, *73*, 151.
(45) Childs-Disney, J. L.; Wu, M. L.; Pushechnikov, A.; Aminova, O.; Disney, M. D. *ACS Chem. Biol.* **2007**, *2*, 745.
(46) Disney, M. D.; Labuda, L. P.; Paul, D. J.; Poplawski, S. G.; Pushechnikov, A.; Tran, T.; Velagapudi, S. P.; Wu, M.; Childs-Disney, J. L. *J. Am. Chem. Soc.* **2008**, *130*, 11185.
(47) Pushechnikov, A.; Lee, M. M.; Childs-Disney, J. L.; Sobczak, K.; French, J. M.; Thornton, C. A.; Disney, M. D. *J. Am. Chem. Soc.* **2009**, *131*, 9767.
(48) Zgarbova, M.; Otyepka, M.; Sponer, J.; Mladek, A.; Banas, P.; Cheatham, T. E.; Jurecka, P. *J. Chem. Theory Comput.* **2011**, *7*, 2886.