JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Refinement of the Sugar–Phosphate Backbone Torsion Beta for AMBER Force Fields Improves the Description of Z- and B-DNA

Marie Zgarbová,[†] Jiří Šponer,[†,‡] Michal Otyepka,[†] Thomas E. Cheatham, III,[§] Rodrigo Galindo-Murillo,[§] and Petr Jurečka*[,†]

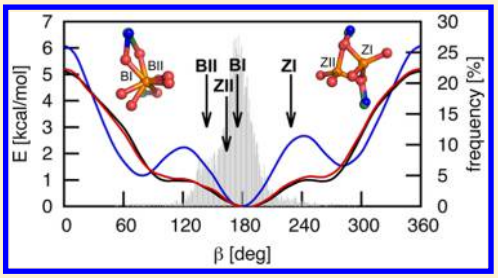
[†]Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacky University, 17. listopadu 12, 77146 Olomouc, Czech Republic

[‡]Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 612 65 Brno, Czech Republic

[§]Department of Medicinal Chemistry, University of Utah, 30 South 2000 East, Skaggs 105, Salt Lake City, Utah 84112, United States

Supporting Information

ABSTRACT: Z-DNA duplexes are a particularly complicated test case for current force fields. We performed a set of explicit solvent molecular dynamics (MD) simulations with various AMBER force field parametrizations including our recent refinements of the ε/ζ and glycosidic torsions. None of these force fields described the ZI/ZII and other backbone substates correctly, and all of them underpredicted the population of the important ZI substate. We show that this underprediction can be attributed to an inaccurate potential for the sugar–phosphate backbone torsion angle β. We suggest a refinement of this potential, β_{OL1}, which was derived using our recently introduced methodology that includes conformation-dependent solvation effects. The new potential significantly increases the stability of the dominant ZI backbone substate and improves the overall description of the Z-DNA backbone. It also has a positive (albeit small) impact on another important DNA form, the antiparallel guanine quadruplex (G-DNA), and improves the description of the canonical B-DNA backbone by increasing the population of BII backbone substates, providing a better agreement with experiment. We recommend using β_{OL1} in combination with our previously introduced corrections, $\varepsilon\zeta_{OL1}$ and χ_{OL4}, (the combination being named OL15) as a possible alternative to the current β torsion potential for more accurate modeling of nucleic acids.



INTRODUCTION

Empirical force fields for nucleic acids are under constant development. Molecular dynamics simulations performed with the parametrizations developed over the last 15 years have provided important insights into the structure and dynamics of nucleic acids,[1−6] but there are still situations that are difficult to model accurately with current force fields. Important examples include some conformational equilibria of DNA such as the BI/BII equilibrium of B-DNA,[7,8] the ZI/ZII equilibrium of Z-DNA,[9] and γ = {gauche+, trans, gauche−} equilibria in guanine quadruplexes.[9,10] In addition, the conformational equilibria of RNA are particularly challenging: it is difficult to reproduce the backbone substates of many important motifs such as the sarcin-ricin loop,[11] some tetraloops,[12] tetranucleotides,[13−15] reversed kink-turns,[16] and ribozymes,[17,18] among others. In this work, we focus on the nonpolarizable force fields of the Cornell et al.[19] family used in the AMBER suite of programs, namely, the ff99 based parametrizations, which are among the most widely used force fields for nucleic acid simulations. Note, however, that other widely used force fields such as the polarizable and nonpolarizable CHARMM are under continuous development, including ongoing efforts to address these problems and related issues.[20,21]

The most frequent targets for force field refinement are torsional parameters. This is no surprise because torsional conformation changes are involved in the transitions between various nucleic acid forms, such as A-DNA, B-DNA, and Z-DNA. Consequently, the accuracy of the torsional parameters has profound effects on the predicted relative stability of these forms. Several torsion reparameterizations have been published[7,9,12,22−27] since the introduction of the ff94[19] force field for nucleic acids based on the work of Cornell et al. For instance, modifications of the sugar pucker and glycosidic torsion parameter χ[22,23] improved the predicted conformational behavior of the deoxyribose rings, the bsc0 modification of the α/γ backbone angles[26] removed the tendency to predict the formation of spurious α/γ backbone substates in long B-DNA simulations, and the χ_{OL3} modification prevented the formation of non-native ladder-like structures in RNA helices.[12,25] The latter modification was derived using a parametrization procedure developed in our laboratory that accounts for certain previously neglected conformation-dependent solvation effects.[28] Since then, we have used this procedure to derive two more modifications: one known as χ_{OL4}[9] that is intended to

Received: July 29, 2015
Published: November 3, 2015

DOI: 10.1021/acs.jctc.5b00716
J. Chem. Theory Comput. 2015, 11, 5723−5736

improve the description of guanines in the χ = *syn* state found, for instance, in DNA quadruplexes and a modification of the ε/ζ sugar−phosphate torsions, $\varepsilon\zeta_{OL1}$, which improved the description of helical twist, groove widths, and the BI/BII populations in B-DNA simulations[7] and helped to improve the modeling of the reactive conformation in the active site of the hairpin ribozyme.[18]

In this work, we derive a modification of the torsion potential for the β angle of the sugar−phosphate nucleic acid backbone, β_{OL1}. As the main test case for the newly developed β_{OL1} potential, we chose the Z-DNA molecule. Z-DNA is particularly interesting for two reasons. First, its backbone assumes noncanonical conformations that make it an excellent probe for the width of applicability of new backbone angle potentials. Second, there is an equilibrium between two well-defined conformational substates of Z-DNA, ZI and ZII,[29] which are rather poorly described by the current AMBER force fields.[7] The ZI and ZII substates appear only in the GpC step and differ mainly in the conformation on the phosphate group, which may either face the minor groove (ZI) or be rotated away from it (ZII; Figure 1). Conformational interchange

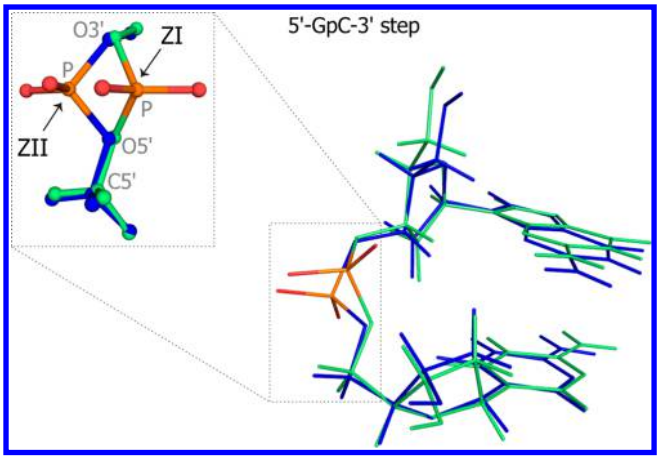


Figure 1. Overlay of the ZI (green) and ZII (blue) conformational substates of the 5′-GpC-3′ dinucleotide step in Z-DNA (different views).

between these substates leaves the overall geometry of the attached G and C nucleosides and the overall Z-DNA fold almost unaffected. The characteristic values of the four most affected backbone angles obtained from X-ray structures are $\varepsilon,\zeta,\alpha,\beta$ = 242°/295°/211°/230° for ZI and 184°/63°/165°/160° for ZII (for variabilities and comparison with other X-ray data,[30] see Table S1 of the Supporting Information). For comparison, the corresponding values for canonical B-DNA are $\varepsilon,\zeta,\alpha,\beta$ = 183°/259°/303°/181°.[30]

ZI is observed as the major conformation in X-ray studies, and ZII is the minor conformation.[29] It has been suggested that the occurrence of ZII in Z-DNA crystals can be linked to the presence of magnesium ions.[29] However, Chatake[31] and Luo et al.[32] showed that both conformations may coexist at the same phosphate in a crystal structure in the absence of divalent metal cations and polyamines. Schneider et al.[33] suggested that the existence of the ZII substate in crystal structures may be attributed to crystal packing effects linked to differences in the hydration patterns of the two substates. Unfortunately, much less is known about the population of the ZII substate in solution. Rauch et al. studied the transition between so-called Z1 and Z2 conformer substates, which they suggested to be related to the ZI/ZII substates observed in single crystals.[34] On the basis of Fourier-transform infrared experiments, they proposed that 39% of the phosphates exist in the Z1 substate at 290 K and 61% exist in the Z2 substate. However, these experiments were performed in hydrated films with a fairly low water content and on very long sequences (ca. 800 nucleotides), so it is not clear whether their conclusions hold for shorter sequences in solution such as those investigated in this work. We therefore assume that the ZI and ZII conformers coexist in solution, but their exact ratio is not known.

We have previously shown that the overall fold of Z-DNA is preserved quite well in molecular dynamics (MD) simulations using various Cornell et al. based force fields and that the duplex is stable on the simulation time scales investigated to date. However, closer inspection shows that almost no pure ZI or ZII states are observed.[7] Instead, simulations performed using the ff99bsc0 force field populate spurious backbone substates that do not correspond to any observed experimental structure. Although the $\varepsilon\zeta_{OL1}$ modification partially improved this situation, the population of the non-native backbone conformations remained too high. In addition, non-native α/γ = trans/g+ substates were observed that further worsened the description of the Z-DNA backbone.[7,9] Thus, none of the current AMBER force fields provides a satisfactory picture of the Z-DNA molecule. Interestingly, the χ_{OL4} correction, which was primarily intended to improve the description of the *syn* region of the χ torsion and which substantially improved the structural description of antiparallel quadruplexes, worsened the description of Z-DNA. This suggests that there is some imbalance between the remaining torsions of Z-DNA that is unmasked by improving the χ potential.

In this work, we derived a modification of the torsion potential for the β angle of the sugar−phosphate nucleic acid backbone, β_{OL1}. We first tested that the newly derived parameters used in combination with the previous $\varepsilon\zeta_{OL1}$ and χ_{OL4} modifications (ff99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$) did not negatively affect simulations of the biologically important B-DNA and guanine DNA quadruplex (G-DNA). Interestingly, we observed subtle improvements, in particular, a better description of some backbone substates of G-DNA and an increase in the population of the BII backbone substate of B-DNA. In addition, the ff99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$ parametrization significantly improved the description of the ZI and ZII substates in Z-DNA. We introduce a shorthand "OL15" for the combination of the β_{OL1}, $\varepsilon\zeta_{OL1}$ and χ_{OL4} dihedral modifications (named after the city of Olomouc, parametrization version 2015). However, we keep the more detailed ff99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$ description in this article to facilitate discussion of contributions from individual dihedral modifications. The parameter files for OL15 modifications are available at http://ffol.upol.cz.

METHODS

Starting Structures and MD Simulations. The starting structure of the Z-DNA hexamer duplex d(CGCGCG)$_2$ was taken from a high-resolution (0.95 Å) X-ray structure (PDB ID 1ICK).[35] Spermine and magnesium ions were not included in the molecular dynamic simulations. Furthermore, 1 μs test simulations were carried out on the Drew−Dickerson (DD) dodecamer d(CGCGAATTCGCG)$_2$ (initial structure, PDB ID 1BNA; resolution, 1.9 Å),[36] antiparallel guanine quadruplex (G-DNA; initial structure, PDB ID 2GWQ; resolution, 2 Å; the

DOI: 10.1021/acs.jctc.5b00716
J. Chem. Theory Comput. 2015, 11, 5723−5736

first molecule in the crystal lattice), 14 bp Jun-Fos oligomer d(GCATTCTGAGTCAG)[37] (initial structure prepared using nucgen module of AMBER[38]), $d(CCCCCGGGGG)_2$ A-DNA duplex (abbreviated C_5G_5; initial structure, PDB ID 1ZF9;[39] resolution, 1.38 Å), short A-tract $d(CGCA_3T_3GCG)_2$ (abbreviated A_3T_3; initial structures, PDB ID 1S2R;[40] resolution, 1.53 Å), A-tract structure $d(CGCGA_6CG)_2$ (abbreviated A_6; initial structure, PDB ID 1D89;[41] resolution, 2.30 Å), and AT-rich sequence $d(CGCATATATGCG)_2$ (abbreviated 3AT; initial structure prepared by nucgen module of AMBER). Short simulations of A-DNA duplex $d(CCCGATCGGG)_2$ (1M77[42]) in 85% ethanol/water solution (ethanol parameters taken from OPLS,[43] and water was TIP3P) were carried out to test the ability of the considered force fields to model A-DNA. In addition, we ran simulations of two short RNA duplexes: a decamer denoted 1QC0′ (500 ns; the decamer $r(GCACC-GUUGG)_2$ was excised from PDB ID 1QC0[44] with a resolution of 1.55 Å) and a tetradecamer duplex $r(U(AU)_6A)$[45] denoted 1RNA (PDB ID 1RNA; 500 ns) as a preliminary test of the applicability of our correction for modeling canonical A-RNA. A list of simulations and simulation conditions is provided in Supporting Information, Table S2.

Simulations were performed with the PMEMD code from the AMBER 12 and 14 program suites[46] under NPT conditions (1 bar, 298 K) with default temperature and pressure settings (tautp = 1.0 ps, taup = 1.0 ps), a 2 fs time step (except for simulations of Jun-Fos, C_5G_5, A_3T_3, A_6, and 3AT duplexes, which were performed with hydrogen mass repartitioning[47,48] and a 4 fs time step), a 9 Å direct space nonbonded cutoff, and SHAKE applied to bonds to hydrogen atoms with the default tolerance (0.00001 Å). The nonbonded pair list was updated every 25 steps. PME was used with default grid settings (1 Å) and default tolerance (10^{-5}). Coordinates were stored every 10 ps. PMEMD for CUDA was used for all simulations.[49] The starting structures were either neutralized with Na^+ ions[50,51] for net-neutral simulations or the ion concentration was adjusted to 2 M for Z-DNA (Z-DNA is experimentally observed at high salt concentrations) with the Joung and Cheatham Na^+ and Cl^- ion parameters.[50,51] Aqvist Na^+ ions[52] were used for G-DNA simulations based on our earlier work.[9] A TIP3P[53] or SPC/E[54] water box was used to solvate the molecular systems (see Supporting Information, Table S2). Cubic boxes were used for G-DNA and Z-DNA molecules, and octahedral boxes, for the remaining simulations.

X-ray Data. Reference X-ray data for DD dodecamer were extracted from the following structures with resolution better than 1.5 Å (regular duplexes, natural basis only, no ligands): 1FQ2, 1JGR, 3U2N, 4C64, 355D. Z-DNA backbone angles (Supporting Information, Table S1) were obtained from the X-ray structures with resolution 1 Å or better with canonical GC pairing only (4R15, 3WBO, 4OCB, 4HIF, 4HIG, 3P4J, 2ELG, 1ICK, 1I0T, 1DJ6, 336D, 292D, 293D, 131D, 1D48, 1DCG, 2DCG). Backbone angles and substate populations of G-DNA were extracted from the X-ray structures with a resolution of 2 Å or better without ligands (1JPQ, 2GWQ, 2HBN, 4R45, 4R47).

AMBER Force Fields. Several Cornell et al.[19] force field variants and combinations were tested: (i) the combination of the ff99 force field with the bsc0 correction for α/γ torsions (ff99bsc0)[26] is available for use in DNA simulations with the AMBER ff10, ff12SB, and ff14SB force fields, (ii) the combination of the ff99bsc0 with the χ_{OL4} correction[9] (ff99bsc0χ_{OL4}), which was introduced to improve the description of *syn* nucleosides, (iii) the combination of the ff99bsc0 with the $\varepsilon\zeta_{OL1}$ correction,[7] which was derived recently to improve the helical parameters of canonical DNA forms, (iv) the combination of the χ_{OL4} and $\varepsilon\zeta_{OL1}$ corrections with the ff99bsc0 force field (ff99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$), which is now the recommended force field for DNA simulations, and (v) the new torsion parameters for the β torsion in combination with the $\varepsilon\zeta_{OL1}$ and χ_{OL4} corrections (ff99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$).

QM and MM Calculations on Z-DNA. Quantum mechanical (QM) calculations in vacuum and in implicit solvent were performed for GpC step fragments of the Z-DNA molecule in the ZI and ZII conformations. Input geometries for QM modeling were obtained from the X-ray structures of Z-DNA: the G2pC3 and G4pC5 steps were used for the ZI and ZII states, respectively (Figure 1). In the QM calculations, the starting structures were optimized at the PBE-D/6-311++G(3df,3pd)[55−58] level with empirical dispersion corrections (1.06−23)[59] using the Turbomole 6.3 code.[60] Cartesian position constraints were applied to three carbon atoms of each nucleic acid base to keep their geometry close to that of the X-ray ZI and ZII substates. No constraints were applied to the sugar−phosphate backbone. The same constraints were used in the molecular mechanics (MM) optimizations, which were performed in the Gaussian 09 software[61] suite using the external function for all of the tested force fields. QM- and MM-optimized geometries were used to compute single-point energies in continuum solvent (water), represented by the COSMO model[62] and a Poisson−Boltzmann (PB)[63,64] model for QM and MM calculations, respectively.

Parameter Development. The β parametrization presented here is based on high-level QM calculations performed on model compounds. The methodology used in its derivation was previously used to derive three other torsion parameters for nucleic acids: χ_{OL3}, χ_{OL4}, and $\varepsilon\zeta_{OL1}$. The main difference between this method of parameter derivation and conventional approaches is the inclusion of certain previously neglected conformation-dependent solvation effects.[28] The solvation correction is based on the difference between QM/COSMO and MM/PB solvation energies and takes into account both the solute−solvent polarization and the conformation dependence of the solvent response to conformation-dependent charges while avoiding double counting of any solvation energy components (eq 1).[28]

$$E_{\mathrm{dih},\beta}^{\mathrm{solv}} = E^{\mathrm{QM/COSMO}} - E_{-\beta}^{\mathrm{MM/PB}} \quad (1)$$

$E^{\mathrm{QM/COSMO}}$ and $E_{-\beta}^{\mathrm{MM/PB}}$ single-point energies were obtained on geometries optimized at the QM/COSMO and MM/PB levels (see refs 25 and 28 for details). A relatively small fragment of the nucleic acid backbone was used as a model structure (Figure 2) to avoid extensive steric clashes with the sugar ring.

The starting structures were optimized at the PBE/6-311++G(3df,3pd) (PBE/LP hereafter) level with the COSMO[62] solvent model (using default scaled Bondi radii, i.e., a scaling factor of 1.17) and with ε_r = 78.4. QM single-point calculations were then performed at the RI-MP2/CBS level. Complete basis set (CBS) extrapolations were obtained according to the scheme of Helgaker and Halkier[65,66] (HF and MP2 energies were extrapolated separately) using the cc-pVTZ and cc-pVQZ basis sets. The in vacuo MP2/CBS energy was combined with the solvation energy obtained from PBE/LP/COSMO calculations. All QM calculations were performed using

DOI: 10.1021/acs.jctc.5b00716
J. Chem. Theory Comput. 2015, 11, 5723−5736

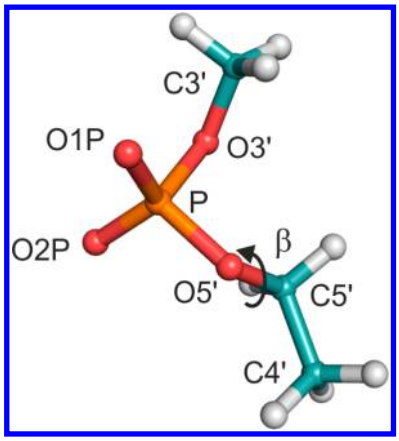


Figure 2. Model structure used for derivation of the β parameters.

TurboMole 5.10.[60,67] MM optimizations were performed with the default parm99 parameters and atomic charges in Gaussian 03[61] using the external function and an in-house script linking Gaussian to the sander module of AMBER 9.[46] A Poisson−Boltzmann (PB)[63,64] continuum solvent was used in the sander calculations. The grid spacing was set to 0.1 Å, and the calculation was performed using the same radii as that for the QM calculation,[28] with ε_r = 78.4. The C5′−O5′−P−O3′ torsion (α) was constrained at its canonical value (−60°), and the β torsion (C4′−C5′−O5′−P) was scanned in 10° increments in both the MM and QM relaxed scans. The dihedral terms (eq 1) were fit with double weights assigned to all points in proximity to the canonical region (β = 130−230°). The final parameters are listed in Table S3 of the Supporting Information.

RESULTS AND DISCUSSION

β_{OL1} Torsion Parameters. The derived torsion parameters and corresponding torsion profiles are shown in Figure 3. All profiles include the relevant solvation energy (QM/COSMO or MM/PB). The most important change with respect to the original ff99 potential is the stabilization of the regions below and above the canonical trans conformation, i.e., around 120° and 240°. As can be seen from the distribution of β values from the X-ray DNA database of Svozil et al.[68] (gray histogram), both of these regions are populated in real structures. The shoulder on the right side of the main distribution corresponds to the ZI substate of Z-DNA (around 230°) and one of the backbone substates in the G-DNA quadruplex (around 260°; see below), whereas the shoulder on the left side corresponds to the BII backbone substate of B-DNA and the ZII substate of Z-DNA. β_{OL1} stabilizes both of these regions significantly, by about 2 kcal/mol compared to the original force field, leading to much shallower wells. Thus, stabilization of the above-discussed regions and larger fluctuations in β can be expected when using β_{OL1} in MD simulations. Note the that the original force field's rigidity with respect to the β angle has previously been remarked upon in a paper that compared MM results to advanced QM calculations on guanine quadruplex stems.[69] It is also clear that the β_{OL1} potential reduces the barrier around 0°. The following sections discuss the effects of these changes on simulations of selected DNA and RNA molecules.

MD Simulations of the DD Dodecamer. The DD dodecamer is one of the most extensively experimentally studied B-DNA duplexes. Although current force fields describe the B-DNA form fairly well, they have some important shortcomings. Notably, the ff99bsc0 force field underestimates this form's helical twist, which proved to be a serious limitation in situations where accurate description of twisting is critical, such as supercoiling[70] or DNA origami.[71] Another shortcoming is an underprediction of the populations of BII backbone substates (see below). These two problems were the main targets of our previous study, in which we introduced the $\varepsilon\zeta_{OL1}$ correction.[7] Compared to simulations using the unmodified ff99bsc0 force field, this correction increased the predicted overall helical twist of B-DNA, improved its sequence dependence, increased the population of the BII substates, and increased the BI/BII interconversion rate in B-DNA duplexes. The $\varepsilon\zeta_{OL1}$ parameters also improved the description of groove widths and substantially reduced the RMSD of the sugar−phosphate backbone compared to that achieved with ff99bsc0 MD simulations.[7] Unfortunately, despite these pronounced improvements, the populations of the BII substates continued to be somewhat underpredicted by the ff99bsc0$\varepsilon\zeta_{OL1}$ force field variant. Here, we present the effects of the more

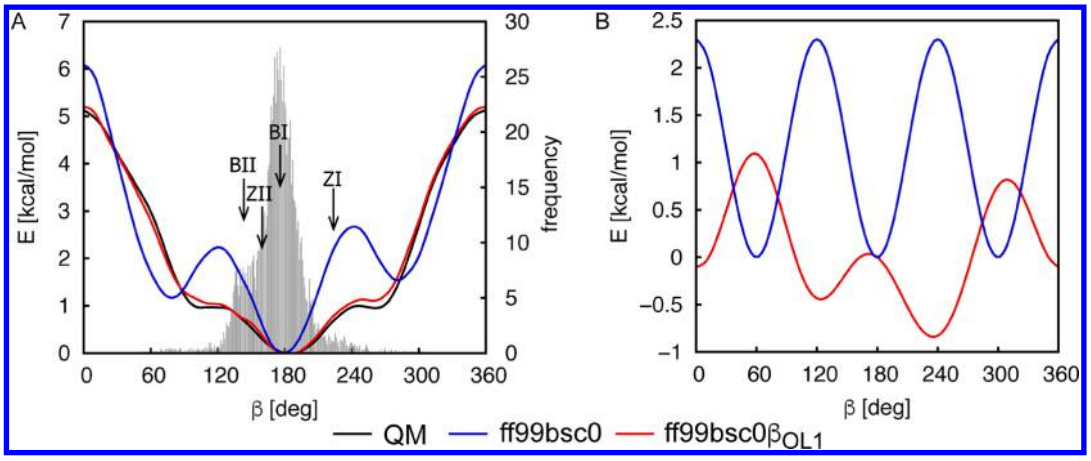


Figure 3. (A) β_{OL1} torsion profile (red) compared to the reference MP2/CBS curve (black) and the ff99bsc0 curve (blue). Solvation energies are included in both the QM and MM results (PBE/LP/COSMO and PB, respectively), and profiles are offset to the canonical minimum (180°). The β distribution (gray) is based on the DNA X-ray data of Svozil et al.[68] (B) MM dihedral terms for β_{OL1} (red) and the original force field term (ff99bsc0, blue), offset to the canonical minimum (180°).

DOI: 10.1021/acs.jctc.5b00716
J. Chem. Theory Comput. 2015, 11, 5723−5736

Table 1. Average Structural Parameters (Last 800 ns of 1 μs Simulations) for the Drew−Dickerson Dodecamer[a]

	NMR[b]	X-ray[c]	99bsc0	99bsc0$\varepsilon\zeta_{OL1}$	99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$	99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$
minor groove/Å	10.5	9.9 ± 1.1	11.6	11.1	11.2	11.0 ± 1.1
major groove/Å	17.6	17.6 ± 1.0	19.0	18.6	17.9	17.8 ± 1.7
χ/deg	−110.8	−112.4 ± 13.8	−120.0	−115.4	−109.7	−108.6 ± 15.8
α/deg	−61.6	−61.3 ± 6.6	−69.9	−68.6	−69.3	−69.7 ± 11.5
β/deg	173.5	176.6 ± 8.3	174.2	176.0	176.4	179.2 ± 13.9
γ/deg	50.3	52.6 ± 5.0	56.1	55.3	53.8	53.4 ± 9.7
δ/deg	126.7	122.4 ± 18.9	119.6	125.5	130.2	131.5 ± 14.4
ε/deg	−171.6	−180.0 ± 11.7	−173.6	−176.1	−175.9	−175.7 ± 14.6
ζ/deg	−101.3	−91.4 ± 11.2	−89.4	−92.6	−94.5	−93.7 ± 15.2
P/deg	136.8	129.8 ± 36.5	127.5	138.0	146.1	147.3 ± 26.2
T_m/deg	33.4	36.6 ± 4.4	37.4	37.4	37.2	37.7 ± 7.0
shift/Å	0.0	0.0 ± 0.3	0.0	0.0	0.0	0.0 ± 0.7
slide/Å	−0.2	0.0 ± 0.5	−0.5	−0.3	−0.1	0.0 ± 0.6
rise/Å	3.2	3.3 ± 0.2	3.3	3.3	3.3	3.3 ± 0.3
tilt/deg	0.0	−0.2 ± 1.7	0.1	0.0	0.0	−0.1 ± 4.7
roll/deg	3.0	2.0 ± 4.0	3.7	2.9	3.7	3.1 ± 5.4
twist/deg	35.7	33.6 ± 5.9	32.7	34.3	34.8	34.9 ± 5.6
shear/Å	0.0	0.0 ± 0.2	0.0	0.0	0.0	0.0 ± 0.3
buckle/deg	0.0	0.9 ± 6.8	−0.2	0.1	0.1	−0.1 ± 10.0
stretch/Å	−0.3	−0.1 ± 0.1	0.0	0.0	0.0	0.0 ± 0.1
propeller/deg	−17.5	−11.8 ± 3.7	−12.7	−12.0	−12.2	−12.1 ± 7.8
stagger/Å	−0.1	0.1 ± 0.1	0.1	0.0	0.0	0.0 ± 0.4
opening/deg	−1.1	2.1 ± 2.4	0.2	0.1	0.1	0.1 ± 4.2
X-displacement/Å	−0.8	−0.5 ± 0.6	−1.6	−1.1	−0.8	−0.6 ± 1.5
Y-displacement/Å	0.0	−0.1 ± 1.0	0.0	0.0	0.0	0.0 ± 1.2
helical rise/Å	3.2	3.3 ± 0.2	3.2	3.3	3.3	3.3 ± 0.4
inclination/deg	5.0	4.3 ± 8.4	7.2	5.2	6.3	5.5 ± 9.4
tip/deg	0.0	0.5 ± 3.4	−0.1	0.0	0.0	0.2 ± 7.8
helical twist/deg	36.1	34.0 ± 5.4	33.9	35.3	35.9	35.9 ± 5.4
%BII			8.1	12.4	17.7	23.4
RMSD[d] (all)/Å			1.65	1.43	1.42	1.37
RMSD[d] (backbone)/Å			1.87	1.58	1.57	1.51

[a]The two terminal base pairs at each end of the dodecamer were excluded from the analysis. Standard deviations are shown only for one simulation for reference, and they are very similar for other force fields. Analyzed using the X3DNA code.[72] [b]Average values from the 1NAJ structure. [c]Average values from well resolved crystal structures (see Methods). [d]RMSD values (mass weighted) for all atoms and the backbone were calculated against the 1FQ2 structure.

recent χ_{OL4}[9] and β_{OL1} modifications on these important features of B-DNA.

Table 1 compares the experimental (X-ray and NMR) structural parameters for the Drew−Dickerson dodecamer to the ff99bsc0, ff99bsc0$\varepsilon\zeta_{OL1}$, ff99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$, and ff99bsc0-$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$ results. All simulations were stable on the simulation time scale (1 μs) and yielded structural parameters in relatively good agreement with the experimental data. Note that this time scale has been shown to provide essentially converged results for this duplex[73] and also other duplexes,[74] except for two to three terminal steps that are affected by base pair fraying.[75] The tested force field modifications provided very similar results: the ff99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$ and ff99bsc0-$\varepsilon\zeta_{OL1}\chi_{OL4}$ parameters produced marginal shifts of the average helical twist and inclination that were somewhat closer to the experimental reference value than were obtained with other parameter sets. The β_{OL1} correction did not significantly affect any of the helical parameters, all of which took values close to the experimental results. Distributions of selected base pairs (those that were most strongly affected by the parameter changes) and base pair step parameters are shown in the Supporting Information (Figure S1).

Backbone Substates in the DD Dodecamer. There are two important conformational equilibria in the B-DNA sugar phosphate backbone. First is the undesirable concerted flip of the α and γ backbone angles from the native combination, α(g−)/γ(g+) to the noncanonical α(g+)/γ(t) substate (see the Introduction), which is a force field artifact that was examined in detail in earlier MD studies. Overpopulation of the noncanonical α(g+)/γ(t) substate was the main problem of the ff99 parametrization, which was corrected in 2007 by the bsc0 correction to the α/γ torsion potentials.[26] Noncanonical α(g+)/γ(t) flips are very rare in ff99bsc0 simulations, appearing for less than 1% of the simulation time, which is acceptable. The population of these noncanonical substates remained low when the $\varepsilon\zeta_{OL1}$ correction was introduced,[7] and it was found to be smaller than 1% also in the simulations reported herein when the last two terminal base pair steps at each end were excluded from the analysis in each case. Note that some of the α(g+)/γ(t) flips may have long lifetimes (tens of nanoseconds[76]), so the populations obtained for 1 μs simulations may not be fully converged, although they should be very good estimates. No accumulation of noncanonical substates over time was observed with either force field. Thus, the new force

DOI: 10.1021/acs.jctc.5b00716
J. Chem. Theory Comput. 2015, 11, 5723−5736

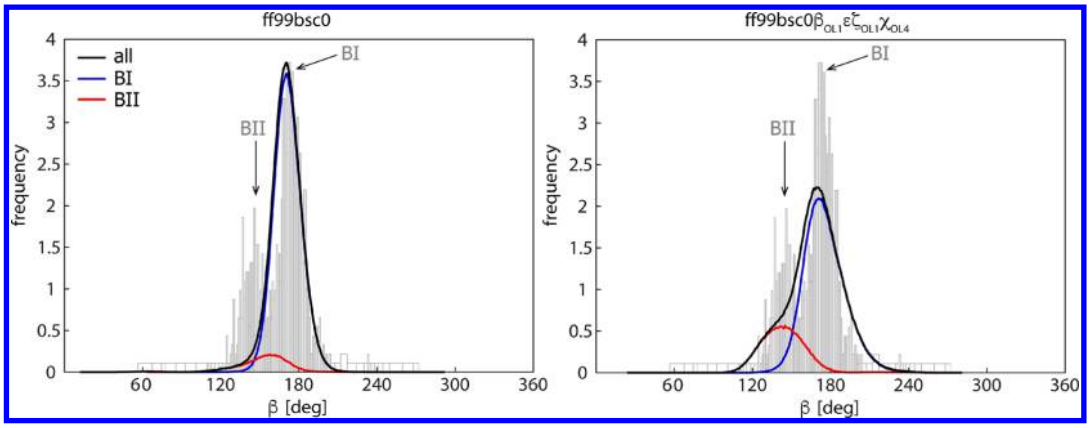


Figure 4. Distribution of β angles (black) for the ff99bsc0 (left) and ff99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$ (right) parametrizations. The distributions of the β angle for the BI state alone (blue) and the BII state alone (red) are also shown. Histograms from MD simulations (the last 800 ns of 1 μs simulations) are compared to X-ray data for B-DNA (gray) reported by Svozil et al.[30] The two terminal base pairs at each end of the studied sequence were excluded from the analysis.

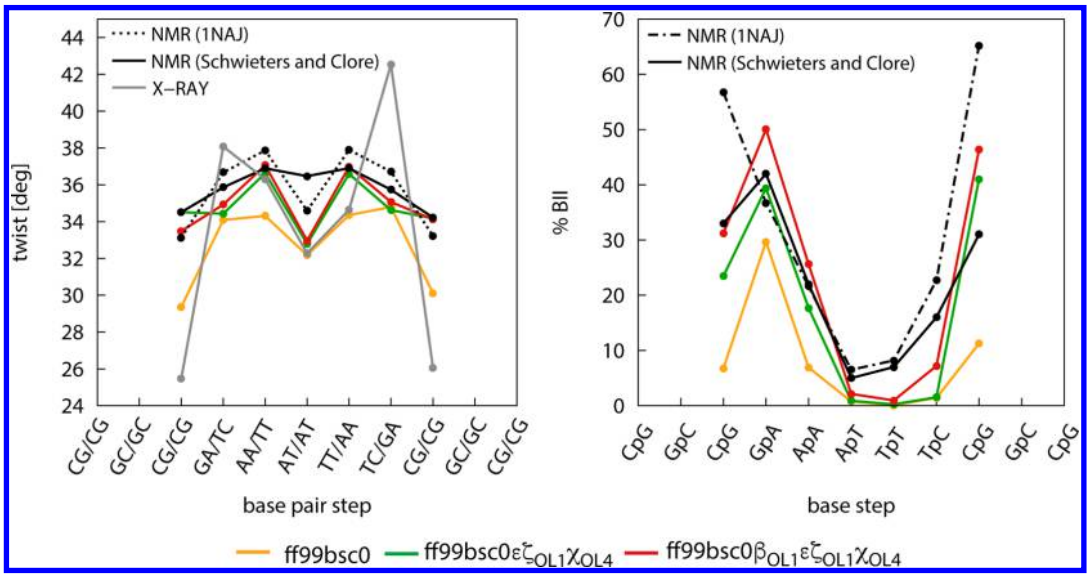


Figure 5. Twist and BII percentage values for individual base pair steps in the DD dodecamer. MD simulations with ff99bsc0, ff99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$, and ff99bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}\beta_{OL1}$ are compared to ensemble-refined NMR data,[78] NMR data with standard refinement[81] (averaged based on the 1NAJ structure), average X-ray data (see Methods), and NMR data from Tian et al.[82] The last 800 ns of the 1 μs simulations were averaged, and the two terminal base pairs at each end of the sequence were excluded from the analysis.

field modifications provide essentially the same description of the α/γ substates as the original ff99bsc0.

The second important equilibrium is between the BI and BII backbone conformations, defined by ε/ζ = t/g− (BI) and ε/ζ = g−/t (BII). The BI conformation is dominant for phosphates in most base pair steps, but some (especially pyrimidine−purine steps) favor the BII state.[77] The BII state is hypothesized to play an important role in sequence-specific recognition of DNA by proteins.[68] Unfortunately, its population seems to be underestimated by some force field parametrizations, including ff99bsc0.[8] This problem was targeted by our recent reparameterization of the ε/ζ torsion potential, $\varepsilon\zeta_{OL1}$, which increased the BII population and improved its sequence dependence.[7] Table 1 shows the influence of the χ_{OL4} and β_{OL1} corrections on the BI/BII equilibrium. Both parametrizations increase the predicted BII content relative to the $\varepsilon\zeta_{OL1}$ version, χ_{OL4} by about 5.3% and β_{OL1} by a further 5.7%, to a final value of 23.4% BII. This is significantly more than that predicted with the $\varepsilon\zeta_{OL1}$ correction and is in much better agreement with experiment.[78,79] The significant increase in the BII population observed in the β_{OL1} simulation is probably due to the stabilization of the $\beta \sim 150°$ region that is characteristic of the BII substate, as shown in Figure 4.

Sequence Dependence of Helical Parameters. Sequence-dependent local conformational variability is a very important feature of B-DNA double helices. Among the most studied helical parameters is helical twist, but sequence dependence of propeller twist has also been discussed, especially in A-tracts.[80] As shown in Figure 5, all of our corrections provide sequence-dependent twists in good agreement with ensemble-refined NMR and large-angle X-ray scattering data of Schwieters and Clore[78] and also with the NMR data with standard refinement.[81] Also shown are X-ray averages obtained from high-resolution X-ray structures (see Methods). Note that the X-ray data are probably affected by crystal packing, which causes notable asymmetry of the twist values (the DD sequence is a palindrome and should be symmetric around the central step). The BII populations

DOI: 10.1021/acs.jctc.5b00716
J. Chem. Theory Comput. 2015, 11, 5723−5736

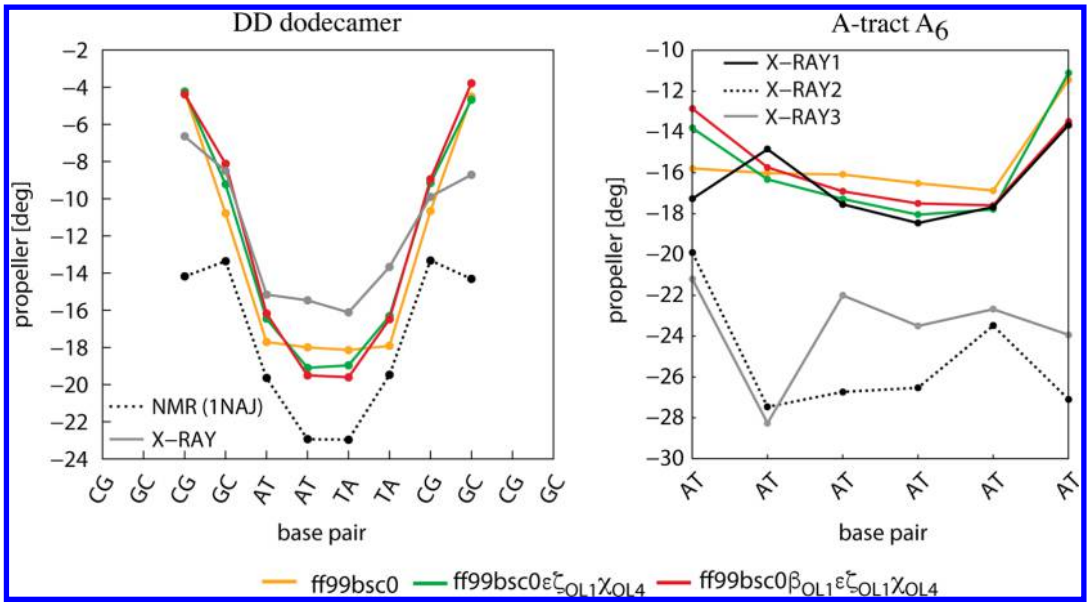


Figure 6. Propeller twist values for individual base pair steps in the DD dodecamer and short A-tract A_6. MD simulations with ff99bsc0, ff99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$, and ff99bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}\beta_{OL1}$ are compared to NMR data (averaged based on the 1NAJ structure[81]) and averaged X-ray data (see Methods) for DD and X-ray data[41] forA_6. The last 800 ns of the 1 μs simulations were averaged, and the two terminal base pairs at each end of the sequence were excluded from the analysis.

(Figure 5) are also in good agreement with the ensemble-derived data and in reasonable agreement with data derived from the ^{31}P NMR chemical shifts.[82] Note, however, that there is some controversy regarding the interpretation of the ^{31}P chemical shifts from NMR experiments[83] and that ensemble-based analysis may, in principle, be a more correct approach than NMR analysis, albeit one associated with larger error bars. We do not compare our BII populations to X-ray data[77] because we did not attempt crystal simulations due to the problematic treatment of crystal waters.[84] The β_{OL1} correction increased the propeller twist only slightly, which seems to be in acceptable agreement with experimental data for DD, but it somewhat overestimated it in the short A-tract A_6 (Figure 6). For more detailed discussion on propeller twist in A-tracts, see, e.g., ref 80. Further testing on five additional duplexes (Jun-Fos, C_5G_5, A_3T_3, A_6, and 3AT; compare also ref 7) and comparison with available experiments[40,77,85] showed similar trends, confirming that the β_{OL1} correction impacts mainly the population of the BII state, whereas its effect on helical twist is smaller (Figure S2 in Supporting Information). Thus, the β_{OL1} modification in connection with the earlier corrections χ_{OL4} and $\varepsilon\zeta_{OL1}$ seems to perform well for simulations of the canonical B-DNA form.

A/B Equilibrium in DNA. The equilibrium between the A- and B-DNA forms is a very important topic in molecular dynamics simulations of nucleic acids.[1,86−88] It has been shown that the older ff99 force field[23] was able to provide stable A-DNA trajectories in concentrated ethanol solution.[1,86,87] Although sequences crystallizing as A-DNA often turn to the B form when immersed in water solution, some GC-rich sequences exist in the A form in 85% ethanol solution.[89,90] In our simulations, the C_5G_5 duplex that crystallizas as A-DNA (1ZF9) quickly (during the first nanoseconds of the simulation) undergoes a transition to the B-DNA structure in water solution with all force feilds considered. In order to test the ability of our force feilds to capture the A-DNA form, we performed additional simulations of $d(CCCGATCGGG)_2$ duplex (1M77) in 85% ethanol solution because this sequence is expected to exist in the A form under low water activity conditions. However, none of the tested force fields (ff99bsc0, ff99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$, and ff99bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}\beta_{OL1}$) was able to preserve the C3′-endo sugar conformation that is characteristic of A-DNA (see Supporting Information, Figure S3). The β torsion angle does not seem to be coupled with this equilibrium because it does not significantly influence population of the C3′-endo sugar pucker. Thus, our β_{OL3} modification cannot improve the unsatisfactory description of A-DNA. Further force field adjustments will be needed to solve this important issue.

Antiparallel Guanine DNA Quadruplex. Guanine quadruplex DNA (G-DNA) is a biologically important class of noncanonical four-stranded DNA molecules[91,92] that has been extensively studied by MD simulations.[93,94] G-DNA molecules are extremely topologically variable, and many G-DNA structures, including antiparallel ones, contain unusual backbone topologies.[95] As such, they are useful test cases for force field refinement. Their G-tetrad stems exist in a complicated equilibrium of α/γ conformational substates and also often have alternating *syn* and *anti* nucleotides as well as unusual values of the β dihedrals. Furthermore, the stems are complemented by single-stranded loops whose conformations are known to be sensitive to force field accuracy.[10] We have already shown that the ff99bsc0χ_{OL4} and ff99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$ force field variants provide notably improved descriptions of the antiparallel G-DNA stem; a detailed discussion of the population of various α/γ substates of *syn*-χ guanines is provided in our previous publications.[7,9] The best (albeit still imperfect) results were achieved with the ff99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$ combination.

Here, we test whether incorporating the β_{OL1} torsion parameters has any effect on the peculiar structure of the dimeric *Oxytricha nova* $d(G_4T_4G_4)_2$ guanine quadruplex. $d(G_4T_4G_4)_2$ forms an antiparallel structure with two equivalent four-thymidine diagonal loops (Figure 7). Due to the availability of several well-resolved X-ray structures of d-

DOI: 10.1021/acs.jctc.5b00716
J. Chem. Theory Comput. 2015, 11, 5723−5736

$(G_4T_4G_4)_2$ (see below), it has been suggested to be an ideal system for testing force fields.[10]

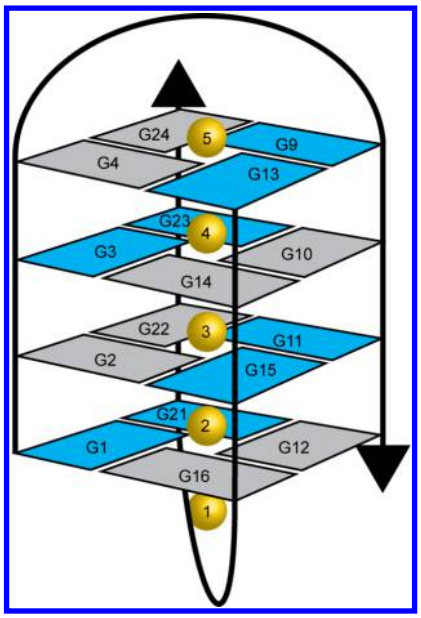


Figure 7. Numbering of the $d(G_4T_4G_4)$ antiparallel DNA quadruplex. Guanine residues with *syn* and *anti* glycosidic conformations are shown in blue and gray, respectively. Monovalent ions observed in the X-ray structures are numbered from 1 to 5.

An important feature of the $d(G_4T_4G_4)_2$ quadruplex is the equilibrium of the α/γ backbone substates in its four-tetrad G-stem. MD trajectories obtained with the ff99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$ and ff99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$ force field combinations were stable on a time scale of 1 μs, and the RMSD values of the quadruplex stems with respect to the X-ray structure (2GWQ) were also very similar: 0.71 Å for ff99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$ and 0.74 Å for ff99bsc0$\varepsilon\zeta_{OL1}\beta_{OL1}\chi_{OL4}$. However, the tested force fields differed in the description of the backbone substates. Table 2 shows the populations of various backbone substates for individual *syn* guanines. There are two distinct types of *syn* guanines in this quadruplex: guanines of the outer tetrads, G9 and G21, which assume the α/γ = g+/t backbone conformation (the outer tetrad guanines G1 and G13 are not shown because they lack phosphates) and guanines of the inner tetrads, G3, G11, G15, and G23, which exist in two α/γ conformations in the X-ray structures, g+/g− and g−/g−. The α/γ = g+/g− substate is more heavily populated in the X-ray structures, being found in 78% of the inner tetrad guanines for high-resolution structures without proteins or ligands acquired at resolutions of 2.0 Å or below (see Methods). As shown in Table 2, when the β_{OL1} correction is applied, the populations of the backbone substates are similar to those predicted by the ff99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$ force field except that the population of the α/γ = g−/g− substates of the inner tetrad guanines is predicted to be more than twice as high. The increase in the g−/g− population of the inner guanines can be explained by the coupling of this state with the $\beta \sim 260°$ region (Figure 8 and ref 9). Because this region is stabilized by the β_{OL1} correction, the stability of the α/γ = g−/g− substate also increases (Figure 8); this is consistent with the relatively high population of this substate in the X-ray structures (ca. 22%; see above). A similar coupling of the β dihedral to the equilibrium among α/γ = g+/t, g+/g−, and other substates also exists in the *syn* guanines of the outer tetrads (see Supporting Information, Figure S4). The β_{OL1} correction slightly increases the population of the native state in

Table 2. Population (%) of Various α/γ Substates of *syn*-χ Guanines in the Stem of the Antiparallel $d(G_4T_4G_4)$ G-DNA Quadruplex[a]

	syn guanines	α/γ g-/g+	g+/g-	g+/t	g-/g-
ffbsc0	G3 (inner)	92	7	0	1
	G9 (outer)	1	0	94	0
	G11 (inner)	78	20	0	2
	G15 (inner)	89	10	0	1
	G21 (outer)	7	0	80	0
	G23 (inner)	61	29	0	10
ff99bsc0 $\varepsilon\zeta_{OL1}\chi_{OL4}$	G3 (inner)	20	76	0	3
	G9 (outer)	10	0	85	0
	G11 (inner)	10	82	0	7
	G15 (inner)	18	78	0	3
	G21 (outer)	26	0	68	0
	G23 (inner)	8	87	0	6
ff99bsc0 $\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$	G3 (inner)	28	62	0	10
	G9 (outer)	20	0	77	0
	G11 (inner)	9	78	0	13
	G15 (inner)	27	61	0	10
	G21 (outer)	19	0	79	0
	G23 (inner)	9	77	0	13
X-ray	G9,21 (outer)	0	0	100	0
	G3,11,15,23 (inner)	0	78	0	22

[a]Gauche+ , gauche−, and trans states of the α and γ dihedrals are denoted g+, g−, and t, respectively. Green backgrounds indicate native states, white backgrounds indicate non-native states, and light green indicates states that are less populated in the X-ray structures (see text). Populations do not sum to 100% because some rare substates are not shown in the table.

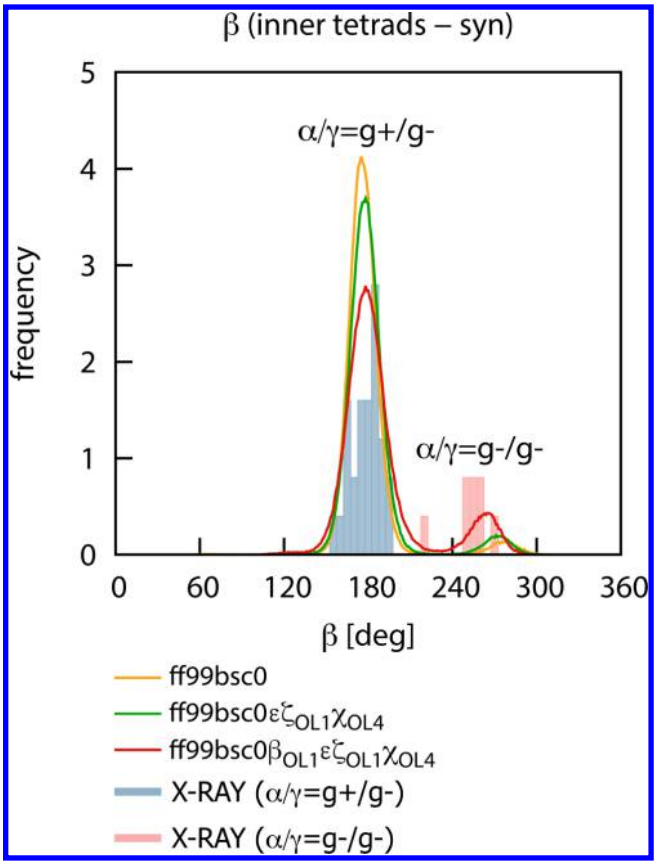


Figure 8. Coupling of the high β ($\beta \sim 260°$) distribution with the minor α/γ = g−/g− backbone substate for inner *syn* guanosine residues in the antiparallel G-DNA quadruplex.

DOI: 10.1021/acs.jctc.5b00716
J. Chem. Theory Comput. 2015, 11, 5723−5736

this case as well. The distributions of the β dihedrals for *anti* guanines and glycosidic torsions are not significantly influenced by the β_{OL1} potential (Figure S4). Note, however, the overall positive effect of the $\varepsilon\zeta_{OL1}$ and χ_{OL4} corrections on the positions of these distributions (relative to the ff99bsc0 results; Figure S4).

Interestingly, both the earlier ff99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$ and the ff99bsc0$\varepsilon\zeta_{OL1}\beta_{OL1}\chi_{OL4}$ force fields provide more stable trajectories for the d$(G_4T_4G_4)_2$ loops than ff99bsc0. Loop reconformation was observed after 150 ns in the ff99bsc0 simulation, after which the loop RMSD increased slowly to about 2.5 Å and remained high for the rest of the simulation (see Supporting Information, Figure S5). Similar behavior of ff99bsc0 force field was observed before by Fadrna et al.,[10] who reported complete irreversible restructuring of the loop regions with loss of the original hydrogen bonding and stacking interactions. Some reconformation was also observed in the ff99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$ simulation, where RMSD started increasing only after about 450 ns and remained somewhat smaller than with ff99bsc0 for most of the simulation time (Figure S5). Finally, reconformation of one loop was also observed in the ff99bsc0$\varepsilon\zeta_{OL1}\beta_{OL1}\chi_{OL4}$ simulation. However, it lasted only for about 100 ns and was reversible, i.e., the loop returned to its crystallographic geometry and was retained for the rest of our simulation (Figure S5). This indicates that the stabilization of the loop by our β modification is not due to artificially rigidifying the structure: the molecule can sample alternative conformations without losing the ability to return to its original state.

The retention of the diagonal loop in the ff99bsc0-$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$ simulation represents significant improvement over ff99bsc0. Note that the loop geometries were preserved even though two monovalent ions (here sodium ions) located directly above and below the terminal tetrads (ions 1 and 5 in Figure 7) escaped from their crystallographic positions within a few tens of nanoseconds of the start in all of our simulations. These experimentally well-established ion positions were never subsequently fully reoccupied by the ions. The loss of the terminal cations has been observed also in MD simulations carried out with various parametrizations of K^+ ions.[9] This is a reminder that, despite many improvements, we remain far from having a perfect force field. The inability of the simulations to reproduce the ion binding at the stem–loop junctions of the d$(G_4T_4G_4)_2$ guanine quadruplex may be a consequence of neglecting explicit polarization effects, which causes a strong overestimation of intercation repulsion inside the quadruplex stems.[96] It is unlikely that such artifacts could be corrected by modifying the dihedral potentials, and it is possible that a polarizable force field for nucleic acids would have to be used, such as that developed in the MacKerell group.[88] Because the two departed ions interact directly with two thymines in the loops, it is understandable that the loop region is destabilized by their escape and becomes more flexible. Thus, short departures from the crystallographic structure are not surprising.

In conclusion, the overall structure of the d$(G_4T_4G_4)_2$ guanine quadruplex is described quite well by both the ff99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$ and ff99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$ force fields, but the ff99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$ combination is recommended because it better describes the frequently occurring α/γ = g−/g− backbone substates associated with the high β ($\beta \sim 260°$) values in G-DNA. The improvement with respect to the ff99bsc0 benchmark is notable also for the loop region.

MD Simulations of Z-DNA. Z-DNA is a very challenging molecule for force field modeling due to its complex topology with alternating *anti* and *syn* nucleotides (Figure 9). Its *syn*

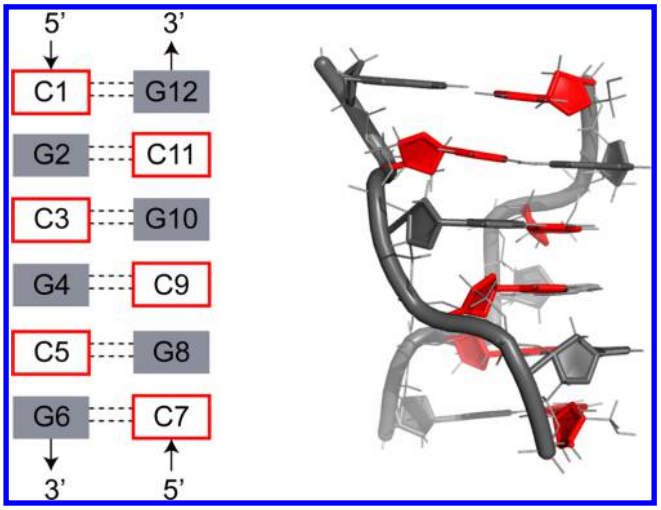


Figure 9. Numbering of the d$(CGCGCG)_2$ Z-DNA hexamer, which has alternating *syn* (cytosines) and *anti* (guanines) glycosidic conformations.

guanines are accompanied by multiple noncanonical values of the sugar–phosphate backbone dihedrals (Table 3). In our previous works, we tested the ability of the ff99bsc0, ff99bsc0χ_{OL4}, ff99bsc0$\varepsilon\zeta_{OL1}$, and ff99bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ force fields to describe the structure and behavior of the Z-DNA hexamer shown in Figure 9,[7,9] and we showed that none of them could correctly reproduce the populations of its backbone conformational substates. The main problems were poor descriptions of the ZI state and the ZI/ZII equilibrium and an undesirable tendency to populate non-native α/γ = trans/g+ substates (α(G) = trans and γ(G) = g+; see also refs 2 and 9). Here, we performed more detailed analyses of the previously published simulations and compared them to a simulation including the β_{OL1} parameters.

In all of our simulations, the overall Z-DNA structure was stable and no major structural changes were observed on a time scale of 1 μs. However, histograms of several sugar–phosphate backbone angles show that multiple non-native substates are populated in all of the simulations with the different tested force fields (Figure 10). We found that our trajectories could be decomposed (clustered) into six distinct conformational substates of the sugar–phosphate backbone (Table 4). The first two substates correspond to the equilibrium between the native α(G)/γ(G) = g+/t and non-native α(G)/γ(G) = t/g+ conformations described earlier.[7,9] The non-native conformation was significantly populated with all tested force field combinations. The remaining four conformational substates emerge around phosphate groups in the GpC steps: these are the native ZI and ZII states and two additional clusters denoted ZI′ and ZII′ (Table 4). It should be noted that we did not observe ZI′ and ZII′ substates in our analysis of highly resolved X-ray structures of Z-DNA; thus, they are probably artifacts created by the used force fields. Histograms of the backbone torsions involved in the ZI/ZII equilibrium with some non-native substates (mainly non-native α/γ = t/g+) filtered off for clarity are shown in Figure S6 in the Supporting Information, separately for pure ZI and ZII substates and also for a mixture

DOI: 10.1021/acs.jctc.5b00716
J. Chem. Theory Comput. 2015, 11, 5723−5736

Table 3. Dihedral Angles (Degrees) in Canonical B-DNA[30] and Z-DNA (This Work) X-ray Structures[a]

B-form[30]		Z-form (5′-CpG-3′ steps)			ZI-form (5′-GpC-3′ steps)		ZII-form (5′-GpC-3′ steps)	
α	303	χ(C)	207	χ(G)		60		
β	181	δ(C)	144	δ(G)		96		
γ	44	ε(C)	265	ε(G)	242		184	
δ	136	ζ(C)	74	ζ(G)	295		63	
ε	183	α(G)	70	α(C)	211		165	
ζ	259	β(G)	185	β(C)	230		160	
χ	252	γ(G)	182	γ(C)	54		50	
		δ(G)	96	δ(C)		144		
		χ(G)	60	χ(C)		207		

[a]Although the 5′-CpG-3′ dinucleotide step exists only in one possible conformation (Z-form), two forms (ZI and ZII) are possible for the 5′-GpC-3′ dinucleotide step.

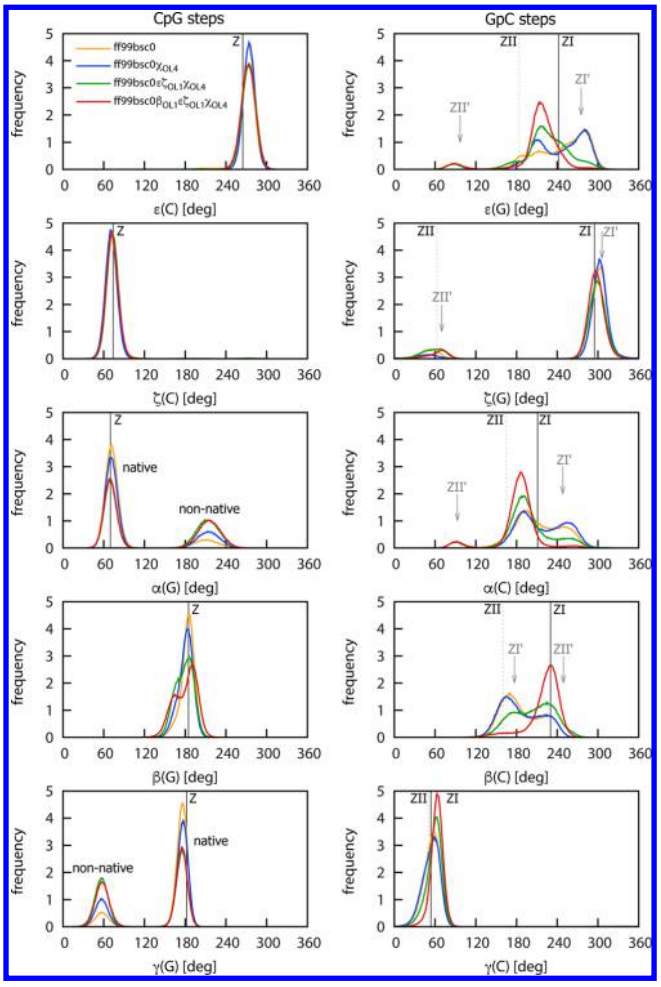


Figure 10. Distributions of Z-DNA backbone dihedrals from the ff99bsc0 (orange), ff99bsc0χ_{OL4} (blue), ff99bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ (green), and ff99bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}\beta_{OL1}$ (red) MD simulations (1 μs). Histograms (averaged over all CpG, respectively GpC, steps) are compared to X-ray averages (solid vertical line, ZI state; dashed vertical line, ZII state). Native (α/γ = g+/t) and non-native (α/γ = t/g+) states are indicated in CpG steps, and non-native ZI′ and ZII′ states are indicated in GpC steps.

of ZI and ZI′ substates. The time evolution of individual backbone angles (Supporting Information, Figure S7) shows that interconversions between the ZI/ZII/ZI′/ZII″ backbone substates are quite frequent on the 1 μs time scale (note that the initial structure 1ICK contains both ZI and ZII backbone substates).

Table 4. Conformational Substates of GpC Z-DNA Base Pair Steps

	X-ray (this work)		MD (ff99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$)					
backbone	ZI	ZII	native α/γ	non-native α/γ	ZI	ZI′	ZII	ZII′
ε(G)	242	187			226	254	178	97
ζ(G)	295	63			299	306	51	70
α(C)	211	169			186	229	171	93
β(C)	230	162			229	177	164	249
γ(C)	54	44			64	57	54	72
α(G)	70		71	216				
γ(G)	182		175	58				

The population of the α/γ and ZI/ZII/ZI′/Z″ substates was clearly force field-dependent, as can be seen from Table 5. In

Table 5. Overall Population of Z-DNA Backbone Substates (% of Simulation Time) during 1 μs MD Simulations Using Different Parameter Sets (Averages of All GpC Steps)

	RMSD [Å]	non-native α/γ [%]	ZI [%]	ZII [%]	ZI′ [%]	ZII′ [%]
99bsc0	1.05	13	29	12	59	0
99bsc0χ_{OL4}	1.01	24	31	6	63	0
99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$	1.07	44	47	11	37	5
99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$	1.08	42	79	4	11	6

the ff99bsc0 simulation, only around 30% of the DNA existed in the ZI state and 12% existed in the ZII state; the rest existed mainly in the non-native ZI′ and ZII″ substates and a relatively small percentage of the α/γ = t/g+ substates. The ff99bsc0 parameter set thus described the Z-DNA conformational equilibrium quite poorly. When the χ_{OL4} correction was applied, the main consequence was an increase in the abundance of the non-native α/γ = t/g+ substates (to 24%), whereas the relative stability of the ZI/ZII/ZI′/ZII′ conformations remained similar. Thus, the glycosidic angle was coupled with the relatively remote α/γ backbone conformations, but this was not so much the case with the ZI/ZII/ZI′/ZII′ substates. As concluded in our previous work, although the χ_{OL4} correction improved the description of the χ = *syn* states of guanine in G-DNA, it somewhat worsened the backbone behavior near the *syn* guanines in Z-DNA. This may reflect an unbalanced description of the α/γ potential in this force field.

Introducing the $\varepsilon\zeta_{OL1}$ correction (99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$ combination) influenced the ZI/ZII equilibrium more directly,

DOI: 10.1021/acs.jctc.5b00716
J. Chem. Theory Comput. 2015, 11, 5723−5736

Table 6. QM and MM Relative Energies of ZI/ZII Substates ($E_{ZII}-E_{ZI}$, kcal/mol)

	QM	99bsc0	99bsc0χ_{OL4}	99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$	99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$
ZII−ZI (vac)	−0.51	−3.78	−3.47	−3.21	−1.40
ZII−ZI (wat)	−0.17	−3.02	−2.74	−2.62	−0.80

increasing the total ZI population to about 47%, which is a significant improvement. Unfortunately, $\varepsilon\zeta_{OL1}$ also further increased the population of the non-native α/γ = t/g+ substates and introduced the non-native ZII′ state, whose population was 5.4% in the 99bsc0$\varepsilon\zeta_{OL1}\chi_{OL4}$ simulation. Thus, although this correction provided some important improvements, it also created additional problems.

The most important effect observed in these tests was that of the β_{OL1} correction, whose inclusion strongly reduced the population of the non-native ZI′ substate and increased that of the native ZI state to almost 79%, leaving 4% of the minor ZII conformation. This is a major improvement compared to that of all of the other tested force field variants. In addition, the 99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$ result was very close to what one would expect for Z-DNA in solution, i.e., the ZI conformation was abundant, although it should be noted that accurate experimental data on this equilibrium in solution are not available (see the discussion in the Introduction and ref 7). The observed changes can be directly related to the modifications of the β potential since this potential is clearly coupled with the ZI/ZII equilibrium (see Figures S9 and S10 in the Supporting Information). We note that the β_{OL1} correction also affected the ZII state (see Figure S6 in the Supporting Information), further increasing the population of the non-native ZII′ conformation such that the ZII′/ZII ratio reached almost 2:1. Although this is clearly undesirable, the total population of the ZII′ state is relatively small (about 6%), so it contributes very little to the total conformational ensemble of Z-DNA. Because of its small absolute population, non-native ZII′ states are unlikely to play an important role in most Z-DNA simulations. Because the Z-DNA conformational ensemble is dominated by the ZI conformation, which is strongly improved by the β_{OL1} correction, we can conclude that the β_{OL1} correction represents an important improvement in the overall description of Z-DNA.

The relative energies of the ZI/ZII states determined in the MM calculations can be compared to results obtained in QM calculations. Although direct comparisons of QM calculations to results obtained using a prepolarized fixed-point charge force field are not straightforward, trends in the relative energies of the two substates may be informative. To perform such a comparison, we chose a small model of the GpC step containing two neighboring bases connected by a phosphate (see Figure 1). The initial geometries of the ZI and ZII conformations were taken from X-ray structures and relaxed such that only three atoms of each nucleobase were fixed at their Cartesian positions in space. This forced the overall geometry of the GpC step to remain essentially identical to that in the crystal while allowing all of the backbone's internal coordinates to relax separately for the ZI and ZII substates. Both ZI and ZII were found to be local minima at our level of theory (PBE-D/LP) and to be relatively close in energy; ZII was slightly more stable, by about 0.5 kcal/mol in vacuum, according to the QM calculations (Table 6). Because the relative energy in vacuum can be different from that in solution, rough estimates of the solvation energy were obtained using the COSMO solvation model in the QM calculations and the GB model in the MM calculations (Table 6, second row). Although approximate implicit solvation reduced the energy gap between the ZI and ZII states, ZII remained more stable than ZI, in contrast to what is observed in MD simulations. This may be explained by inaccuracies of the solvation models (implicit vs explicit) or entropic effects due to conformation space sampling in MD simulations. Because the results obtained in vacuum exhibited the same trends as those obtained in continuum solvent, we will discuss only the latter in the following. In the ff99bsc0 parametrization, the ZII state was much more stable than in the reference QM calculation, by about 3.0 kcal/mol. This discrepancy was slightly reduced by the χ_{OL4} correction (to 2.7 kcal/mol) and further by the $\varepsilon\zeta_{OL1}$ correction (to 2.6 kcal/mol). Nevertheless, the largest improvement came from the β_{OL1} correction, which reduced the difference with respect to QM calculations to only 0.8 kcal/mol. We note that we do not expect a perfect match of the QM and MM results because the force field is not parametrized for use in vacuum. Nevertheless, the fact that the modified β_{OL1} torsion brought the MM values closer to the QM reference is encouraging.

Although the β_{OL1} modification brought important improvements in G-DNA, B-DNA, and mainly Z-DNA structures, there are still problems to be solved. One of them is inaccurate description of the α/γ backbone substates in Z-DNA discussed above (Table 5). Too high population of the non-native α/γ = t/g+ substates may indicate need for further refinement of these torsions, possibly the γ torsion potential, which may overly destabilize the γ = trans state. Note, however, that population of the non-native α/γ substates depends also on water model used (see below). Furthermore, non-native α/γ populations are clearly coupled with other torsions (Table 5), and further modifications of other dihedrals (and possibly also other force field components) may be necessary to solve this problem. Unfortunately, such extensive couplings between different torsions obscure situation and make stepwise refinements of parametrizations of heterogeneous origin rather difficult. We hope that our procedure including conformation-dependent solvation effects will provide a better reference and help to identify where further improvements are needed. Efforts in this direction are underway.

Effect of the Water Model. Water models are known to influence the properties of simulated nucleic acids in MD simulations.[97,98] The two most widely used models are TIP3P and SPC/E.[54] In the case of the B-DNA DD dodecamer modeled with the ff99bsc0$\varepsilon\zeta_{OL1}$ force field, we showed that changing the water had only very small effects on helical parameters but caused significant increases in the population of the BII substate, from 12.3% in TIP3P to 15.6% in SPC/E.[7] We also tested the effect of SPC/E water on Z-DNA simulation. Interestingly, we observed several important changes. First, the SPC/E model strongly reduced the population of the non-native α/γ = t/g+ backbone substates in the CpG steps, from 42% in TIP3P to 22% in SPC/E (see also Supporting Information, Figure S8). Second, it substantially increased the population of the native ZI substate, from 79% in TIP3P to 89% in SPC/E. Finally, the SPC/E model shifted the distributions of the ε(G) and ζ(G) of the ZI substate closer

DOI: 10.1021/acs.jctc.5b00716
J. Chem. Theory Comput. 2015, 11, 5723−5736

to the X-ray reference values (Figure S8). The SPC/E water model thus seems to improve the description of the Z-DNA structure.

Preliminary RNA Simulations. Although the β_{OL1} parameters were developed for DNA simulations, we also carried out test MD simulations on the RNA duplexes 1QC0′ (r(GCACCGUUGG)$_2$) and 1RNA (r(U(AU)$_6$A); see Methods). Tables S4 and S5 in Supporting Information compare the X-ray structural parameters for 1QC0′ and 1RNA with the results of ff99bsc0χ_{OL3}, ff99bsc0$\beta_{OL1}\chi_{OL3}$, and ff99bsc0-$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL3}$ simulations. All simulations were stable on the simulation time scale (500 ns) and provided very similar results, suggesting that the β_{OL1} correction might also be suitable for RNA simulations. However, due to complexity of RNA structures and the numerous noncanonical motives they contain, extensive MD simulations on a multitude of RNA structures would be necessary to assess the reliability and usefulness of the newly derived parameters for RNA MD simulations. Work in this direction is underway.

CONCLUSIONS

This study reports a reparametrization of the β torsion parameters of the Cornell et al. force field, β_{OL1}. β_{OL1} was derived based on high-level QM calculations with inclusion of conformation-dependent solvation effects[7] that are not covered by the standard explicit solvent models in MD simulations. The refined parameters were tested by using them in extended MD simulations of several representative DNA systems, namely, B-DNA duplexes, an antiparallel guanine quadruplex, and a Z-DNA molecule. β_{OL1} is intended mainly for simulations of DNA and was tested in combination with our previously derived torsion modifications for DNA, $\varepsilon\zeta_{OL1}$[7] and χ_{OL4}[9] (ff99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$).

β_{OL1} correction significantly improved the modeling of the Z-DNA molecule, which is a particularly complicated test case for current force fields. It significantly influenced the relative stability of the native and non-native sugar−phosphate backbone substates in Z-DNA and improved the structural description of the ZI substate. In particular, the ff99bsc0-$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$ force field combination increased the ZI population from 29% (with ff99bsc0) or 47% (with ff99bsc0-$\varepsilon\zeta_{OL1}\chi_{OL4}$) to 79%, in better agreement with experiment. Even better results were obtained when the SPC/E water model was used instead of TIP3P. However, further improvements of the B-DNA force field may be needed because the population of some non-native α/γ backbone substates in Z-DNA was still relatively high and did not seem to be influenced by the β potential.

The description of a second benchmark system, the antiparallel guanine quadruplex, was improved only slightly by the β_{OL1} potential. In particular, our β correction modestly improved the agreement of the populations of some α/γ backbone substates with experiment and provided more stable trajectory of the loop region than the ff99bsc0 and ff99bsc0-$\varepsilon\zeta_{OL1}\chi_{OL4}$ simulations.

Significant improvements were achieved by using the β_{OL1} potential in simulations of the B-DNA DD dodecamer. Specifically, the β_{OL1} correction increased the average population of the BII substates from 17.7% with ff99bsc0-$\varepsilon\zeta_{OL1}\chi_{OL4}$ to 23.4% with the ff99bsc0$\beta_{OL1}\varepsilon\zeta_{OL1}\chi_{OL4}$ force field combination; the latter value is in better agreement with experiment.[79] The predicted sequence dependence of BII propensity was also in good agreement with experiment. Furthermore, β_{OL1} increased the helical twist of some base pair steps while reducing it in others; overall, the simulation results remained in good agreement with the experimentally observed sequence dependence of the helical twist in the DD dodecamer and other tested duplexes. Other helical and geometrical parameters are almost unaffected by our β correction.

Preliminary tests on canonical duplexes did not show large effects, either positive or negative, on the RNA simulations. However, we do not recommend β_{OL1} for RNA simulations without completing more extensive tests.

In conclusion, the β_{OL1} correction significantly improved the overall description of sugar−phosphate backbone equilibria in B-DNA, G-DNA, and, in particular, Z-DNA molecules. We recommend a combination of β_{OL1} with our previously derived $\varepsilon\zeta_{OL1}$ and χ_{OL4} parameters (the combination being named OL15) for DNA simulations. Although β_{OL1} is in better agreement with QM reference than the original β potential, one should keep in mind that the original force field was in a sense balanced and errors in one component were more or less well compensated by the other components. Thus, new imbalances may arise for untested systems, and long-term experience will be necessary before the β_{OL1} modification can be routinely used with confidence for DNA simulations. One should also keep in mind that not all problems can be addressed by torsion refinement and that other force field components such as vdW parameters or point charges may need adjustment. Also, additive force fields have inherent limitations in reproducing polarization effects and conformation-dependent electronic structure effects.[99] Thus, some artifacts cannot be corrected within the additive force field approximation.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00716.

β_{OL1} torsion parameters, list of MD simulations, and time series and distributions of selected structural parameters in MD simulations and X-ray data (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: petr.jurecka@upol.cz.

Funding

This work was supported by grant 14-29874P (M.Z.) from the Grant Agency of the Czech Republic. Further funding was provided by project LO1305 of the Ministry of Education, Youth and Sports of the Czech Republic.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Cheatham, T. E., III; Crowley, M. F.; Fox, T.; Kollman, P. A. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (18), 9626−9630.

(2) Cheatham, T. E., III; Case, D. A. *Biopolymers* **2013**, *99* (12), 969−977.

(3) Sponer, J.; Banas, P.; Jurecka, P.; Zgarbova, M.; Kuhrova, P.; Havrila, M.; Krepl, M.; Stadlbauer, P.; Otyepka, M. *J. Phys. Chem. Lett.* **2014**, *5* (10), 1771−1782.

(4) Maffeo, C.; Yoo, J.; Comer, J.; Wells, D. B.; Luan, B.; Aksimentiev, A. *J. Phys.: Condens. Matter* **2014**, *26* (41), 413101.

(5) Laughton, C. A.; Harris, S. A. *Wires Comput. Mol. Sci.* **2011**, *1* (4), 590−600.

DOI: 10.1021/acs.jctc.5b00716
J. Chem. Theory Comput. 2015, 11, 5723−5736

(6) Vargiu, A. V.; Magistrato, A. *ChemMedChem* **2014**, *9* (9), 1966−1981.
(7) Zgarbova, M.; Luque, F. J.; Sponer, J.; Cheatham, T. E., III; Otyepka, M.; Jurecka, P. *J. Chem. Theory Comput.* **2013**, *9* (5), 2339−2354.
(8) Hart, K.; Foloppe, N.; Baker, C. M.; Denning, E. J.; Nilsson, L.; MacKerell, A. D. *J. Chem. Theory Comput.* **2012**, *8* (1), 348−362.
(9) Krepl, M.; Zgarbova, M.; Stadlbauer, P.; Otyepka, M.; Banas, P.; Koca, J.; Cheatham, T. E., III; Jurecka, P.; Sponer, J. *J. Chem. Theory Comput.* **2012**, *8* (7), 2506−2520.
(10) Fadrna, E.; Spackova, N.; Sarzynska, J.; Koca, J.; Orozco, M.; Cheatham, T. E., III; Kulinski, T.; Sponer, J. *J. Chem. Theory Comput.* **2009**, *5* (9), 2514−2530.
(11) Kruse, H.; Havrila, M.; Sponer, J. *J. Chem. Theory Comput.* **2014**, *10* (6), 2615−2629.
(12) Banas, P.; Hollas, D.; Zgarbova, M.; Jurecka, P.; Orozco, M.; Cheatham, T. E., III; Sponer, J.; Otyepka, M. *J. Chem. Theory Comput.* **2010**, *6* (12), 3836−3849.
(13) Bergonzo, C.; Henriksen, N. M.; Roe, D. R.; Swails, J. M.; Roitberg, A. E.; Cheatham, T. E., III *J. Chem. Theory Comput.* **2014**, *10* (1), 492−499.
(14) Roe, D. R.; Bergonzo, C.; Cheatham, T. E., III *J. Phys. Chem. B* **2014**, *118* (13), 3543−3552.
(15) Condon, D. E.; Kennedy, S. D.; Mort, B. C.; Kierzek, R.; Yildirim, I.; Turner, D. H. *J. Chem. Theory Comput.* **2015**, *11* (6), 2729−2742.
(16) Sklenovsky, P.; Florova, P.; Banas, P.; Reblova, K.; Lankas, F.; Otyepka, M.; Sponer, J. *J. Chem. Theory Comput.* **2011**, *7* (9), 2963−2980.
(17) Mlynsky, V.; Banas, P.; Hollas, D.; Reblova, K.; Walter, N. G.; Sponer, J.; Otyepka, M. *J. Phys. Chem. B* **2010**, *114* (19), 6642−6652.
(18) Mlynsky, V.; Kuhrova, P.; Zgarbova, M.; Jurecka, P.; Walter, N. G.; Otyepka, M.; Sponer, J.; Banas, P. *J. Phys. Chem. B* **2015**, *119* (11), 4220−4229.
(19) 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.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117* (19), 5179−5197.
(20) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30* (10), 1545−1614.
(21) Baker, C. M.; Anisimov, V. M.; MacKerell, A. D. *J. Phys. Chem. B* **2011**, *115* (3), 580−596.
(22) Cheatham, T. E., III; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16* (4), 845−862.
(23) Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21* (12), 1049−1074.
(24) Yildirim, I.; Stern, H. A.; Kennedy, S. D.; Tubbs, J. D.; Turner, D. H. *J. Chem. Theory Comput.* **2010**, *6* (5), 1520−1531.
(25) Zgarbova, M.; Otyepka, M.; Sponer, J.; Mladek, A.; Banas, P.; Cheatham, T. E., III; Jurecka, P. *J. Chem. Theory Comput.* **2011**, *7* (9), 2886−2902.
(26) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., III; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92* (11), 3817−3829.
(27) Yildirim, I.; Kennedy, S. D.; Stern, H. A.; Hart, J. M.; Kierzek, R.; Turner, D. H. *J. Chem. Theory Comput.* **2012**, *8* (1), 172−181.
(28) Zgarbova, M.; Luque, F. J.; Sponer, J.; Otyepka, M.; Jurecka, P. *J. Chem. Theory Comput.* **2012**, *8* (9), 3232−3242.
(29) Wang, A. H. J.; Quigley, G. J.; Kolpak, F. J.; Vandermarel, G.; Vanboom, J. H.; Rich, A. *Science* **1981**, *211* (4478), 171−176.
(30) Cech, P.; Kukal, J.; Cerny, J.; Schneider, B.; Svozil, D. *BMC Bioinf.* **2013**, *14*, 205.
(31) Chatake, T. *J. Synchrotron Radiat.* **2013**, *20*, 864−868.
(32) Luo, Z. P.; Dauter, M.; Dauter, Z. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2014**, *70*, 1790−1800.
(33) Schneider, B.; Ginell, S. L.; Jones, R.; Gaffney, B.; Berman, H. M. *Biochemistry* **1992**, *31* (40), 9622−9628.
(34) Rauch, C.; Pichler, A.; Trieb, M.; Wellenzohn, B.; Liedl, K. R.; Mayer, E. *J. Biomol. Struct. Dyn.* **2005**, *22* (5), 595−614.
(35) Dauter, Z.; Adamiak, D. A. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2001**, *57*, 990−995.
(36) Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U. S. A.* **1981**, *78* (4), 2179−2183.
(37) Heddi, B.; Foloppe, N.; Oguey, C.; Hartmann, B. *J. Mol. Biol.* **2008**, *382* (4), 956−970.
(38) Bansal, M.; Bhattacharyya, D.; Ravi, B. *Bioinformatics* **1995**, *11* (3), 281−287.
(39) Hays, F. A.; Teegarden, A.; Jones, Z. J. R.; Harms, M.; Raup, D.; Watson, J.; Cavaliere, E.; Ho, P. S. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (20), 7157−7162.
(40) Woods, K. K.; Maehigashi, T.; Howerton, S. B.; Sines, C. C.; Tannenbaum, S.; Williams, L. D. *J. Am. Chem. Soc.* **2004**, *126* (47), 15330−15331.
(41) Digabriele, A. D.; Steitz, T. A. *J. Mol. Biol.* **1993**, *231* (4), 1024−1039.
(42) Ramakrishnan, B.; Sekharudu, C.; Pan, B. C.; Sundaralingam, M. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003**, *59*, 67−72.
(43) Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. *J. Am. Chem. Soc.* **1996**, *118* (45), 11225−11236.
(44) Klosterman, P. S.; Shah, S. A.; Steitz, T. A. *Biochemistry* **1999**, *38* (45), 14784−14792.
(45) Dockbregeon, A. C.; Chevrier, B.; Podjarny, A.; Johnson, J.; Debear, J. S.; Gough, G. R.; Gilham, P. T.; Moras, D. *J. Mol. Biol.* **1989**, *209* (3), 459−474.
(46) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26* (16), 1668−1688.
(47) Feenstra, K. A.; Hess, B.; Berendsen, H. J. C. *J. Comput. Chem.* **1999**, *20* (8), 786−798.
(48) Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. *J. Chem. Theory Comput.* **2015**, *11* (4), 1864−1874.
(49) Gotz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. *J. Chem. Theory Comput.* **2012**, *8* (5), 1542−1555.
(50) Joung, I. S.; Cheatham, T. E., III *J. Phys. Chem. B* **2008**, *112* (30), 9020−9041.
(51) Joung, I. S.; Cheatham, T. E., III *J. Phys. Chem. B* **2009**, *113* (40), 13279−13290.
(52) Aaqvist, J. *J. Phys. Chem.* **1990**, *94* (21), 8021−8024.
(53) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79* (2), 926−935.
(54) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91* (24), 6269−6271.
(55) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. *J. Comput. Chem.* **1983**, *4* (3), 294−301.
(56) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80* (7), 3265−3269.
(57) Gill, P. M. W.; Johnson, B. G.; Pople, J. A.; Taylor, S. W. *J. Chem. Phys.* **1992**, *96* (9), 7178−7179.
(58) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72* (1), 650−654.
(59) Jurecka, P.; Cerny, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.* **2007**, *28* (2), 555−569.
(60) Ahlrichs, R.; Bar, M.; Haser, M.; Horn, H.; Kolmel, C. *Chem. Phys. Lett.* **1989**, *162* (3), 165−169.
(61) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; 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.;

DOI: 10.1021/acs.jctc.5b00716
J. Chem. Theory Comput. 2015, 11, 5723−5736

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 D.02; Gaussian, Inc.: Wallingford, CT, 2003.

(62) Klamt, A.; Schuurmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, No. 5, 799−805.

(63) Lu, Q.; Luo, R. *J. Chem. Phys.* **2003**, *119* (21), 11035−11047.

(64) Luo, R.; David, L.; Gilson, M. K. *J. Comput. Chem.* **2002**, *23* (13), 1244−1253.

(65) Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286* (3−4), 243−252.

(66) Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Olsen, J. *Chem. Phys. Lett.* **1999**, *302* (5−6), 437−446.

(67) Weigend, F.; Häser, M. *Theor. Chem. Acc.* **1997**, *97* (1−4), 331−340.

(68) Svozil, D.; Kalina, J.; Omelka, M.; Schneider, B. *Nucleic Acids Res.* **2008**, *36* (11), 3690−3706.

(69) Sponer, J.; Mladek, A.; Spackova, N.; Cang, X. H.; Cheatham, T. E., III; Grimme, S. *J. Am. Chem. Soc.* **2013**, *135* (26), 9785−9796.

(70) D'Annessa, I.; Coletta, A.; Sutthibutpong, T.; Mitchell, J.; Chillemi, G.; Harris, S.; Desideri, A. *Nucleic Acids Res.* **2014**, *42* (14), 9304−9312.

(71) Yoo, J.; Aksimentiev, A. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (50), 20099−20104.

(72) Lu, X. J.; Olson, W. K. *Nucleic Acids Res.* **2003**, *31* (17), 5108−5121.

(73) Galindo-Murillo, R.; Roe, D. R.; Cheatham, T. E., III *Biochim. Biophys. Acta, Gen. Subj.* **2015**, *1850* (5), 1041−1058.

(74) Dans, P. D.; Perez, A.; Faustino, I.; Lavery, R.; Orozco, M. *Nucleic Acids Res.* **2012**, *40* (21), 10668−10678.

(75) Zgarbova, M.; Otyepka, M.; Sponer, J.; Lankas, F.; Jurecka, P. *J. Chem. Theory Comput.* **2014**, *10* (8), 3177−3189.

(76) Perez, A.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2007**, *129* (47), 14739−14745.

(77) Heddi, B.; Foloppe, N.; Bouchemal, N.; Hantz, E.; Hartmann, B. *J. Am. Chem. Soc.* **2006**, *128* (28), 9170−9177.

(78) Schwieters, C. D.; Clore, G. M. *Biochemistry* **2007**, *46* (5), 1152−66.

(79) Djuranovic, D.; Hartmann, B. *J. Biomol. Struct. Dyn.* **2003**, *20* (6), 771−788.

(80) Banas, P.; Mladek, A.; Otyepka, M.; Zgarbova, M.; Jurecka, P.; Svozil, D.; Lankas, F.; Sponer, J. *J. Chem. Theory Comput.* **2012**, *8* (7), 2448−2460.

(81) Tjandra, N.; Tate, S.; Ono, A.; Kainosho, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122* (26), 6190−6200.

(82) Tian, Y.; Kayatta, M.; Shultis, K.; Gonzalez, A.; Mueller, L. J.; Hatcher, M. E. *J. Phys. Chem. B* **2009**, *113* (9), 2596−2603.

(83) Precechtelova, J.; Munzarova, M. L.; Vaara, J.; Novotny, J.; Dracinsky, M.; Sklenar, V. *J. Chem. Theory Comput.* **2013**, *9* (3), 1641−1656.

(84) Liu, C. M.; Janowski, P. A.; Case, D. A. *Biochim. Biophys. Acta, Gen. Subj.* **2015**, *1850* (5), 1059−1071.

(85) Stefl, R.; Wu, H. H.; Ravindranathan, S.; Sklenar, V.; Feigon, J. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (5), 1177−1182.

(86) Sprous, D.; Young, M. A.; Beveridge, D. L. *J. Phys. Chem. B* **1998**, *102* (23), 4658−4667.

(87) Noy, A.; Perez, A.; Laughton, C. A.; Orozco, M. *Nucleic Acids Res.* **2007**, *35* (10), 3330−3338.

(88) Savelyev, A.; MacKerell, A. D. *J. Comput. Chem.* **2014**, *35* (16), 1219−1239.

(89) Ivanov, V. I.; Minchenkova, L. E.; Minyat, E. E.; Frank-Kamenetskii, M. D.; Schyolkina, A. K. *J. Mol. Biol.* **1974**, *87* (4), 817−833.

(90) Tolstorukov, M. Y.; Ivanov, V. I.; Malenkov, G. G.; Jernigan, R. L.; Zhurkin, V. B. *Biophys. J.* **2001**, *81* (6), 3409−3421.

(91) Balasubramanian, S.; Hurley, L. H.; Neidle, S. *Nat. Rev. Drug Discovery* **2011**, *10* (4), 261−275.

(92) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res.* **2006**, *34* (19), 5402−5415.

(93) Sponer, J.; Cang, X. H.; Cheatham, T. E., III *Methods* **2012**, *57* (1), 25−39.

(94) Cang, X. H.; Sponer, J.; Cheatham, T. E., III *Nucleic Acids Res.* **2011**, *39* (10), 4499−4512.

(95) Karsisiotis, A. I.; O'Kane, C.; da Silva, M. W. *Methods* **2013**, *64* (1), 28−35.

(96) Gkionis, K.; Kruse, H.; Platts, J. A.; Mladek, A.; Koca, J.; Sponer, J. *J. Chem. Theory Comput.* **2014**, *10* (3), 1326−1340.

(97) Besseova, I.; Banas, P.; Kuhrova, P.; Kosinova, P.; Otyepka, M.; Sponer, J. *J. Phys. Chem. B* **2012**, *116* (33), 9899−9916.

(98) Kuehrova, P.; Otyepka, M.; Sponer, J.; Banas, P. *J. Chem. Theory Comput.* **2014**, *10* (1), 401−411.

(99) Mladek, A.; Krepl, M.; Svozil, D.; Cech, P.; Otyepka, M.; Banas, P.; Zgarbova, M.; Jurecka, P.; Sponer, J. *Phys. Chem. Chem. Phys.* **2013**, *15* (19), 7295−310.

DOI: 10.1021/acs.jctc.5b00716
J. Chem. Theory Comput. 2015, 11, 5723−5736