Toward Reproducing Sequence Trends in Phosphorus Chemical Shifts for Nucleic Acids by MD/DFT Calculations

Jana Přecechtělová,[†,‡] Markéta L. Munzarová,*[,†,‡] Juha Vaara,[$] Jan Novotný,[†,‡] Martin Dračínský,[∥,#] and Vladimír Sklenář[†,‡]

[†]National Centre for Biomolecular Research, Faculty of Science and [‡]Central European Institute of Technology (CEITEC), Masaryk University, Kamenice 5/A4, CZ-62500 Brno, Czech Republic

[$]NMR Research Group, Department of Physics, University of Oulu, P.O. Box 3000, FIN-90014 Oulu, Finland

[∥]NMR Laboratory, Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, v. v. i., Flemingovo nám. 2, CZ-16610 Prague 6, Czech Republic

[#]Department of Chemistry, Durham University, Durham, DH13LE, United Kingdom

S *Supporting Information*

ABSTRACT: This work addresses the question of the ability of the molecular dynamics−density functional theory (MD/DFT) approach to reproduce sequence trend in ^{31}P chemical shifts (δP) in the backbone of nucleic acids. δP for [d(CGCGAATTCGCG)]$_2$, a canonical B-DNA, have been computed using density functional theory calculations on model compounds with geometries cut out of snapshots of classical molecular dynamics (MD) simulations. The values of ^{31}P chemical shifts for two distinct B-DNA subfamilies BI and BII, δP/BI and δP/BII, have been determined as averages over the BI and BII subparts of the MD trajectory. This has been done for various samplings of MD trajectory and for two sizes of both the model and the solvent embedding. For all of the

combinations of trajectory sampling, model size, and embedding size, sequence dependence of δP/BI in the order of 0.4−0.5 ppm has been obtained. Weighted averages for individual ^{31}P nuclei in the studied DNA double-helix have been calculated from δP/BI and δP/BII using BI and BII percentages from free MD simulations as well as from approaches employing NMR structural restraints. A good qualitative agreement is found between experimental sequence trends in δP and theoretical δP employing short (24 ns) MD run and BI, BII percentages determined by Hartmann et al. or via MD with the inclusion of NMR structural restraints. Theoretical δP exhibit a systematic offset of ca. 11 ppm and overestimation of trends by a factor of ca. 1.7. When scaled accordingly, theoretical δP/BI and δP/BII can be used to determine the expected percentage of BII to match the experimental value of δP. As evidenced by the calculations on snapshots from Car−Parrinello molecular dynamics, the systematic offsets of the theoretical δP obtained by MD/DFT approach result primarily from the unrealistic bond lengths employed by classical MD. The findings made in this work provide structure−δP relationships for possible use as NMR restraints and suggest that NMR calculations on MD snapshots can be in the future employed for the validation of newly developed force fields.
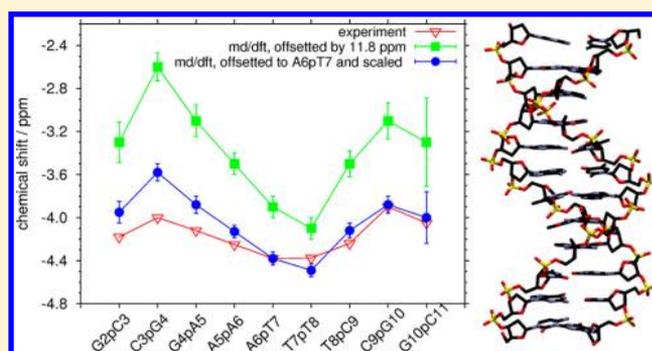
1. INTRODUCTION

The NMR chemical shift represents an extremely sensitive probe of the structure in the vicinity of the magnetic nuclei. In the case of the ^{31}P nucleus in the backbone of nucleic acids, the chemical shift (δP) is being frequently related to the value of the torsion angle ζ, which in the B-type of nucleic acids can be found either in the −*gauche* (−*g*) region (approximately −60°, BI class) or in the *trans* (*t*) region (approximately 180°, BII class), cf. Figure 1. The experimentally estimated difference in δP between the pure BI and pure BII forms amounts to 1.6 ppm,[1,2] and δP for real systems undergoing dynamical BI↔BII transitions have been interpreted in terms of the BI/BII ratio.[3]

However, a number of theoretical studies went beyond the simple BI/BII interpretation of ^{31}P chemical shift in phosphate

esters. As early as in 1975, Gorenstein and Kar[4] considered the dependence of δP in dimethyl phosphate on torsion angles α and ζ around the P−O bonds, exploring a continuous set of torsion angles. Although the adopted computational approach has been extremely approximate from today's point of view, their results are in a qualitative agreement (after transferring from the older τ-scale to the currently used δ-scale of shifts) with our recent calculations.[5] In 1984, Giessner-Prettre et al.[6] analyzed δP in terms of both the OPOC and the CCOP torsion angles and concluded that both of these conformational parameters substantially contribute to the value of δP.
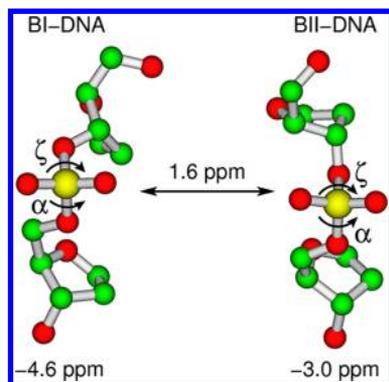
Figure 1. BI and BII classes of B-type nucleic acids: backbone conformation and phosphorus chemical shift.

Recently, we and Benda et al. have published several DFT studies of ^{31}P chemical shift going from static calculations on models with freely rotatable bonds[7−9] toward calculations on molecular dynamics (MD) snapshots of models whose torsion angle spaces have been fixed by the bulk of the nucleic acid.[5] For the system under study, namely the steps G4pA5, T7pT8, and C9pG10 of the Drew-Dickerson dodecamer [d-(CGCGAATTCGCG)]₂ (DD), our results showed that the MD/DFT approach[5] provides a good agreement with the experimentally estimated difference between the BI and BII chemical shifts. At the same time, we found out that both torsion angles ζ and α should be considered as continuous variables. We also concluded that, for the canonical values of angles β and ε, implicit treatment (through the correlated value of the neighboring angle α or ζ) is sufficient.

Our work in ref 5 focused on the torsion angle dependence of δP for data accumulated on three steps of DD. The only sequence dependent property studied was the chemical shift difference between BI and BII states. There, for each of the three steps, we divided the MD trajectory into its BI ($\zeta > 210°$) and BII ($\zeta \leq 210°$) regions, and we computed separately average δP for the BI subregion (δP/BI) and δP for the BII subregion (δP/BII) and its error bars. We noticed that δP/BI and δP/BII depended on the step in question, but considering the error bars, no statistically significant conclusions could have been made about the potential sequence dependence of δP/BI and δP/BII.

The potential sequence dependence of δP/BI and δP/BII is, however, interesting, also from the practical point of view of molecular dynamics force field validations, where the proper description of BI and BII population is still an open, thoroughly discussed question. A current standard employed for force fields is based on the work of the Hartmann group, which suggested that experimental δP can be translated into BI/BII ratios using a simple empirical relationship.[3] This relationship has been derived in the following way. First, the empirical correlation has been established between the sequential interproton distances H2′$_i$−H6/8$_{i+1}$, H2″$_i$−H6/8$_{i+1}$, H6/8$_i$−H6/8$_{i+1}$ from NMR and the difference of torsion angles ε and ζ, $\varepsilon-\zeta$, measured from X-ray structures. Second, empirical correlation has been established between the same interproton distances from NMR and δP from NMR. Finally, the two correlations have been combined to find correlation between δP and $\varepsilon-\zeta$, and the value of $\varepsilon-\zeta$ was associated with the percentage of BII structures. This NMR-determined relative population of BII has been applied to validate major FFs,[10,11] and the translation of δP into distance restraints has been

employed for structure refinement.[12] It has been concluded that the Parmbsc0 FF (and likewise the CHARMM27 FF) suppresses the BII conformation, especially in CG steps.[10,13,14]

All of the latter applications have been built on the premise that δP can be translated into BI/BII ratios without any additional information. That is, within the experimental accuracy, there is assumed to be one δP/BI value and one δP/BII value, and only the relative occurrences of BI and BII determine the final δP. In this paper, we study in detail the sequence dependence of δP/BI and δP/BII in order to judge the suitability of Hartmann's relationship. For this purpose, we employed the same system (DD) and approach (MD/DFT) as in ref 5, but we performed a number of additional calculations. First, the data of reference 5 have been extended from 3 steps to all 9 inner sugar−phosphate steps of DD. One outermost step at each end (C1pG2 and C11pG12) has been excluded from the calculations since the corresponding base pairs reveal increased mobility known as "fraying", observed in X-ray,[15] NMR,[16−18] and computer experiments.[19−21] Second, we have tested several MD trajectories of various lengths and sizes of steps for creating snapshots that entered DFT calculations of δP/BI and δP/BII. Third, we explored the influence of extending the dimethyl phosphate (DMP) model with the first solvation shell to the isopropyl ethyl phosphate (IPEP) with the first solvation shell, cf. Figure 2, and of increasing its
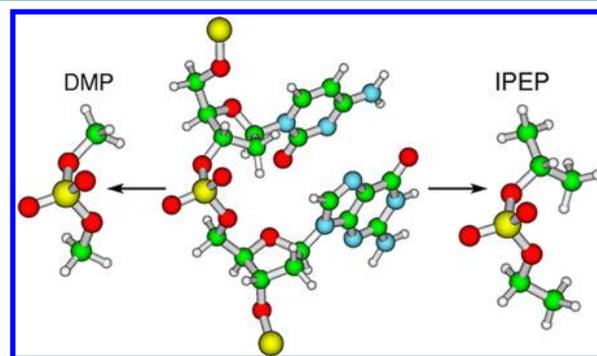
Figure 2. Models DMP and IPEP employed in DFT calculations. DMP and IPEP have in all calculations been surrounded by their first solvation shell, defined as described in the Computational Details section. In part of the calculations on DMP, the solvation was extended to 2nd solvation sphere and closest Na$^+$ inclusion, as is shown in Figure 3.

solvent embedding to the second solvation sphere and one closest sodium ion per each snapshot, cf. Figure 3. In parallel, we checked whether the accuracy of the MD/DFT approach is high enough to reproduce the experimentally determined sequence trends in total δP.

The possibility to calculate δP by quantum chemistry tools on geometries extracted from the MD snapshots evokes the idea of comparing the theoretical chemical shifts to experiment and thus validating the force field (FF) differently than by Hartmann's relationship. For this purpose, we considered our shortest (24 ns) and our longest (264 ns) trajectory. By comparing quantum-chemically determined δP/BI and δP/BII with experimental δP, we estimate the BI/BII ratios and compare these with BI/BII ratios determined from NMR-restrained MD,[22] from NMR experiments using the Hartmann's relationship,[3] and from free molecular dynamics simulations in this and other[23] studies. We also present 2D maps of δP as a function of the torsion angles α and ζ
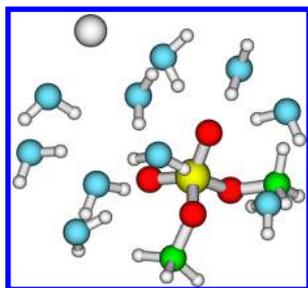
Figure 3. Model DMP with the 1st and 2nd solvation shells and closest Na^+ ion. Color scheme adopted: yellow, phosphorus; red, phosphate oxygens; green, carbons; small white spheres, hydrogens; blue, water oxygens; big white sphere, sodium.

computed by the MD/DFT approach for the DMP and IPEP models. Additionally, we performed a Car−Parrinello molecular dynamics (CPMD) simulation of a dimethyl phosphate in water and again calculated on the CPMD snapshots the values of δP, presented in the form of 2D maps of δP as a function of the torsion angles α and ζ. Our motivation for obtaining these was a search for possible reasons of the necessity to offset theoretical values and scale theoretical trends in order to obtain an agreement with experiment. The chemical shift map for CPMD snapshots on free DMP covers a much wider torsion-angle space than the map of chemical shielding for the DD system. In order to explore the chemical shift throughout the full conformational space of 360° for α and ζ, we have performed a series of DFT calculations on a grid of α and ζ. Our paper is concluded with an interpretation of the dependence of ^{31}P shielding on either α or ζ in terms of orbital interactions.

2. METHODS AND COMPUTATIONAL DETAILS

Classical MD Simulation. Our starting MD trajectory for the creation of snapshots was the 24-ns trajectory described in reference 5, referred to as A/24 ns. The simulation has been carried out in explicit solvent using the Amber 7.0 program suite.[24] The starting structure was the NMR structure 1NAJ.[25] The system was neutralized using Na^+ ions distributed using the Xleap module of Amber and defined by the force-field parm99.[26] The same parameters were used for additional Na^+ and Cl^- ions added to 150 mM concentration. A periodically repeating truncated octahedron water box of size 61 × 61 × 61 $Å^3$ containing 5501 water molecules has been employed. Constant pressure conditions (1 atm) and constant temperature (300 K) maintained using the Berendsen weak-coupling technique[27] with a time constant of 5.0 ps were used. The simulations utilized the particle mesh Ewald summation[28] with a 9 Å cutoff for nonbonding interactions and integration step of 2 fs, which implies fixed bond lengths. Following a 150-ps system equilibration, we performed a 24.0-ns production run (consisting of 1-ns parts) with the force-field parm99 with parmbsc0 corrections.[29] The latter is recommended to avoid a too frequent occurrence of unrealistic α/γ-transitions found in the simulations with the older parm99 FF.[30] The snapshots for the DFT calculations have been saved in 4-ps steps, so that in total 6000 snapshots has been obtained.

As suggested by a reviewer, this trajectory has been extended, using the last point of the previous 24.0-ns production run as a starting point. We employed the Amber 12.0 program suite,[31] since the Amber 7.0 program was no longer available at our sources. The extension run was one sequential trajectory,

divided during the calculation into 5-ns parts, and was terminated after 390 ns. The snapshots for the DFT calculations have been saved in 40-ps steps. All other parameters have been kept the same as for the original 24.0-ns trajectory. The resulting (390 ns + 24 ns = 414 ns) trajectory is assigned as trajectory A and has been sampled from 0 to 24 ns of production run (see above, referred to as trajectory A/24 ns), from 24 to 124 ns of production run (referred to as A/124 ns), and from 24 to 264 ns of production run (referred to as A/264 ns), counted from the starting structure 1 NAJ. For A/24 ns and A/264 ns, 6000 snapshots for DFT calculations were available; for A/124 ns, 2500 snapshots were available.

Since the symmetry of the double helix was not monotonically improving with the extension of the trajectory (see below), we have run a second MD trajectory. This was obtained with same program and parameters as for the extension described above. The starting structure was 1NAJ equilibrated exactly as in the original trajectory described above, and a production run of 190.0 ns followed. The trajectory has been sampled from 0 to 120 ns of production run, which is further denoted as B/120 ns. Six thousand snapshots for DFT calculations have been saved with a time separation of 20 ps.

Definiton of the BII Conformation. In this work, BII conformation was assigned to all structures with $\zeta \leq 210°$. This choice has been motivated by our analysis of torsion angle histograms for the A/24 ns trajectory: 210° turned out to be an approximate average coordinate of the minimum between the BI and BII maxima in the histograms of ζ, see below. Our results can be, to some extent, dependent on the choice of the BII definition, but this dependence is expected to be a minor one since the flat regions around the minimum between BI and BII maxima are counted into the averages $\delta P/BI$ and $\delta P/BII$ with a very small weight. A detailed inspection of the dependence of $\delta P/BI$ and $\delta P/BII$ on BII definition is worth a future study, but it is beyond the scope of the current work.

CPMD Simulation.[32] Periodic boxes (10.051 × 10.051 × 10.051 $Å^3$) containing 27 water molecules and the DMP molecule in $-g-g$ or $-gt$ conformation (C−O−P−O torsion angle −60° or 180°, respectively) were created with the HyperChem program.[33] Within HyperChem and the Amber99[34,35] FF, a classical simulation was run for 100 ps with 1 fs integration time step and the temperature of 300 K to equilibrate the systems. The TIP3P[36] FF as a part of Amber99 was used for water. Then, the geometry was optimized and input into the CPMD software package.[37] Periodic boundary conditions and 4 au (0.09676 fs) time step were maintained for the CPMD calculations performed with the BLYP[38,39] functional and Vanderbilt ultrasoft pseudopotentials.[40] Energy cutoff of 25 Ry was used. The initial configuration was relaxed by five short CPMD runs comprising 200 steps. After each run, the system was quenched to the Born−Oppenheimer surface by reoptimizing the wave function. Longer, 97-ps production runs were then performed at the temperature of 300 K maintained with the Nosé−Hoover algorithm,[41,42] which also kept the system in the canonical (NVT) ensemble. During the long simulations the trajectory was saved at every 1000th step. The calculation took four months on a cluster using four processors.

Construction of Cluster Models for MD/DFT Calculations. MD snapshots have been extracted in the following way: (1) from the 0 to 24 ns of production run of trajectory A in 4-ps steps (A/24 ns), resulting in a total of 6000 frames, (2) from 24 to 124 ns of production run of trajectory A in 40-ps
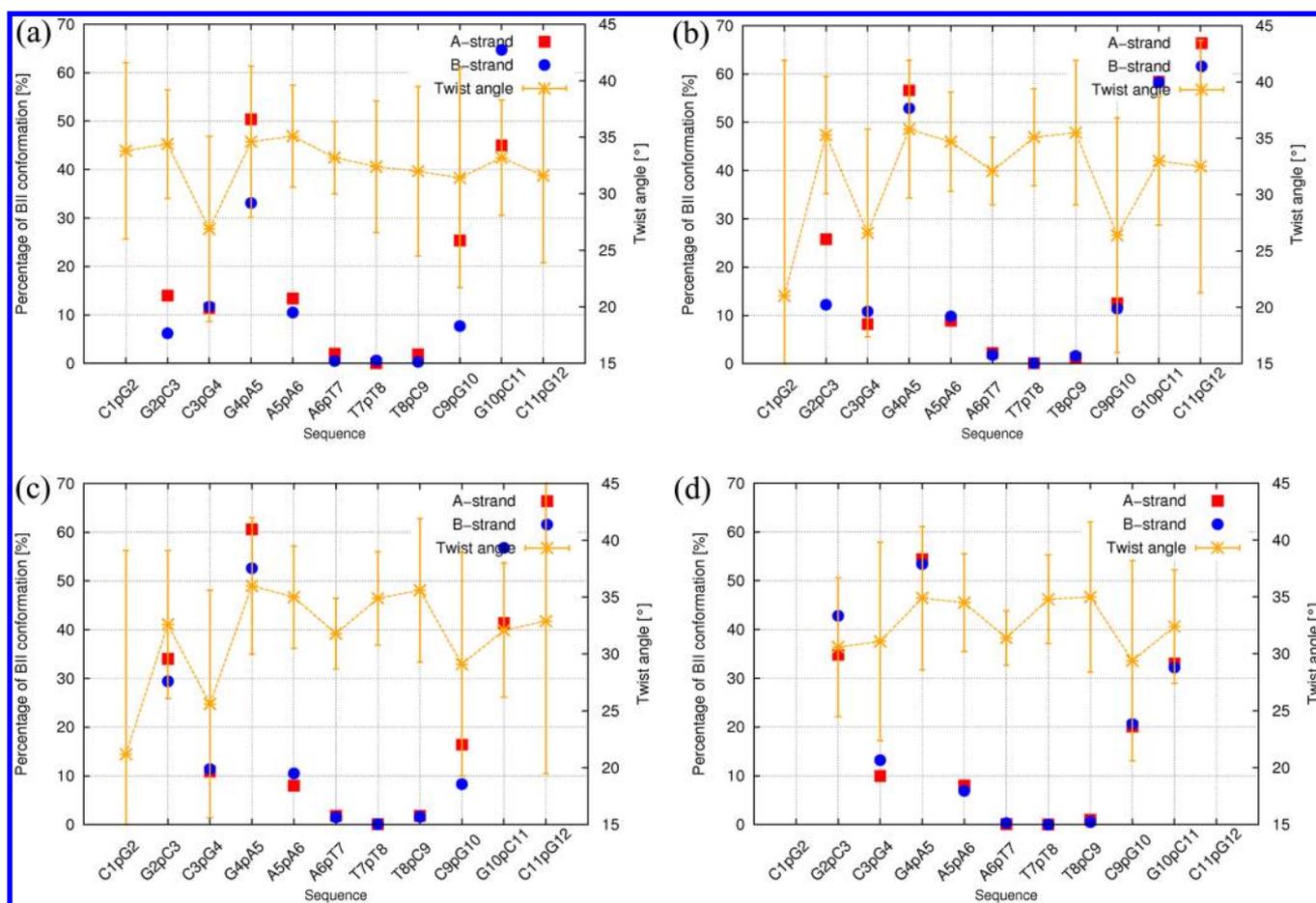
Figure 4. Development of BII populations and twist angles with time for several trajectory/sampling combinations: (a) A/24 ns, (b) A/124 ns, (c) A/264 ns, (d) 1.2 μs trajectory of Pérez et al.[23] The BII conformation was assigned to all structures with ζ ≤ 210°.

steps (A/124 ns), resulting in a total of 2500 frames, (3) from 24 to 264 ns of production run of trajectory A in 40-ps steps (A/264 ns), resulting in a total of 6000 frames, (4) from 0 to 120 ns of production run of trajectory B in 20-ps steps (B/120 ns), resulting in a total of 6000 frames.

Phosphate steps were modeled either as dimethyl phosphate (DMP) or as isopropyl ethyl phosphate (IPEP). The cluster models were cut out from the MD frames, saturating broken bonds with a H-atom bonded to a C-atom with a C−H bond length of 1.09 Å. All geometry parameters have been kept at their MD values, that is, not optimized.

Solvent embedding has been modeled either as a first solvation shell or as a first and a second solvation shell along with the closest Na⁺ ion. For the first solvation shell and the A/24 ns trajectory, the clusters included only those water molecules directly hydrogen-bonded to phosphate anionic (keto) oxygens, and no water molecules hydrogen-bonded to the ester oxygens. This was done for the sake of consistency with ref 5 and can be justified by the fact that X-ray structures reveal that ester oxygens are only responsible for ~15% of all phosphate-water contacts.[43] When extending the MD trajectory to A/124 ns and A/264 ns and when switching to B/120 ns trajectory, the first solvation shell included all water molecules directly hydrogen-bonded to both keto and ester oxygens of the phosphate. The following criteria were imposed to determine the presence of a hydrogen bond: $R(PO...H_W) \leq 2.25$ Å and $\angle(PO...H-O_W) \geq 135°$.

The second solvation shell has been defined as all water molecules, for which a hydrogen atom was at the distance of 4.65 Å or closer from the phosphorus atom. This criterion is justified below. For the Na⁺ cations, one ion per snapshot closest to the phosphorus atom has been included.

Molecular Geometries for DFT Calculations on the Grid. Out of the four relevant torsion angles in DMP, the two more distant from phosphorus were set to constant values and the other two were varied from 0° to 355° in 5° steps. A total of $72 \times 72 = 5184$ structures were constructed, which were subsequently geometry optimized at the B3LYP/6-31G(d) level. During the geometry optimization, all four relevant torsion angles have been frozen.

Molecular Geometries for CPMD/DFT Calculations. The first 1000 frames recorded during the Car−Parrinello simulation, which started from the −g−g and −gt conformation, respectively, were used for the combined CPMD/DFT calculations. DMP plus all water molecules hydrogen-bonded to any of the atoms in DMP with a H-bond distance less than or equal to 2.25 Å were extracted from the snapshots. Certain conformations of the DMP model are equivalent for symmetry reasons. The symmetry was therefore used in the construction of the CPMD map in order to obtain more data points in the bins, as described in the Supporting Information, Figure S1.

Molecular Geometries for MO Analysis. Two models of DMP with $\alpha = -60°$, $\zeta = -60°$ ($-g-g$) and with $\alpha = -60°$, $\zeta = 180°$ ($-gt$) have been geometry optimized at the B3LYP/6-31G(d) level with the α, ζ torsion angles kept frozen. The

resulting values of ε, β were $\beta = -168°$, $\varepsilon = -168°$ for $-g-g$ and $\beta = -168°$, $\varepsilon = -178°$ for $-gt$.

Chemical Shift Calculations. MD/DFT and CPMD/DFT calculations employing the model of DMP have been performed with same parameters as those described in ref 5 and in the Supporting Information. In MD/DFT calculations employing IPEP, a Gaussian-type DZVP orbital basis set[44] was used for all atoms additional to DMP. Thus, a locally dense basis set has been employed in order to describe the surroundings of phosphorus atom accurately and to keep the computational requirements limited at the same time.[45] For DFT calculations on the grid, the wave function has been calculated in Gaussian03[46] at the BP86[47,48]/IGLO-III[49] level using the "UltraFine" grid and "SCF = Tight" convergence criteria. For all other calculations, the converged Kohn–Sham molecular orbitals have been obtained in Turbomole5.6.[50−52] The molecular orbitals were transferred by Turbomole-to-Mag and Gaussian-to-Mag interface routines[53] to the property module MAG of the ReSpect code[54] and subsequently employed in NMR calculations. Chemical shift calculations in MAG employed sum-over-states density functional theory with the individual gauges for localized orbitals (IGLO)[49] choice of the gauge origin. The same choices have been employed for the molecular orbital analysis with the Boys method of localization.[55]

Calculation of Statistical Errors. The statistical errors have been determined for each MD sampling and each model/embedding combination by means of the test of statistical inefficiency described in the Supporting Information. It can be argued that the standard ways of estimating statistical errors in molecular simulations assume that the underlying process produces a normal distribution of the investigated property, whereas in our case a bimodal (BI and BII) distribution should be expected. However, the chemical shift distributions of the two conformers are rather much overlapping, despite the clear "pooling" of the structures, as is shown below. Hence, the standard procedure was followed. We note that the evaluation of the test of statistical inefficiency requires the determination of a plateau in a graph of data that is to some extent subjective. Our estimate of the introduced "error in error bar" is typically 0.02−0.03 ppm, rarely 0.04 ppm, for BI shifts, and 0.02−0.05 ppm for BII shifts.

■ RESULTS AND DISCUSSION

Molecular Dynamics Simulations of [d-(CGCGAATTCGCG)]$_2$ Starting from pdb Structure 1NAJ. From the point of view of the observables, on which we focus in this work, two properties of MD simulations are of crucial interest. First, it is the convergence of BII populations of the individual steps and the convergence of the related twist angles.[56] The BII populations should be same for corresponding steps in the two chains, and for the palindromic sequence, the twist angles should possess axial symmetry with respect to the middle step when plotted as functions of the individual steps. Second, it is important that the trajectory is converged with respect to histograms of α and ζ torsion angles, as those determine $\delta P/BI$ and $\delta P/BII$.

Figure 4(a−c) shows the BII populations and twist angles for the trajectories A/24 ns, A/124 ns, and A/264 ns. For a comparison, Figure 4(d) displays the results of the 1.2 μs trajectory of Pérez et al.[23] BII populations close to those of Pérez et al. (except for end-of-chain steps) have also been obtained in a 2400 ns simulation of DD in 150 mM KCl by

Dršata et al.[10] Clearly, the A/24 ns trajectory in Figure 4(a) is far from being converged. The A/124 ns trajectory is much closer to the 1.2 μs trajectory of Pérez et al., but the A/264 ns trajectory is again less converged and less symmetrical. To understand why this is so, we have plotted the dependence of % of BII state as a function of simulation progress for the step G4pA5. The latter step has been selected because it reveals a substantial decrease in symmetry when going from A/124 ns to A/264 ns trajectory, and at the same time, it is relatively far from both ends of the chain. The time dependence of % of BII state is shown in Figure S2 of the Supporting Information. Apparently, an approximate symmetry in strands A and B (not to be confused with trajectories A and B) is gained between ca. 25 ns and ca. 100 ns, respectively, but then, the two chains diverge again until ca. 240 ns. We have continued our MD run until 390 ns; then, it was terminated due to the observation of breaking of base pairs at one end of the duplex and distortion of one chain end into the DNA groove at 280 ns. The loss of symmetry at G4pA5 at 100 ns probably already was the first sign of the loss of symmetry in our MD simulation.

Our maximum explored trajectory A/264 ns is thus certainly not converged as concerns BI and BII populations and symmetry. As discussed above, for step G4pA5, the symmetry continually decreases after ca. 100 ns. This motivated us to explore an MD trajectory longer than A/24 ns but shorter than A/264 ns, of a length close to A/124 ns, but with higher concentration of snapshots than available for the A/124 ns trajectory. As discussed in the Computational Details section, the interval for saving snapshots for the A/124 ns was 40 ps. The length of the A/124 ns was 100 ns (from 24 to 124 ns), hence 2500 snapshots have been obtained. The resulting error bars were significantly larger than obtained for the A/264 ns trajectory with equal interval for saving snapshots (40 ps), where 6000 snapshots were available.

Accordingly, starting from 1NAJ, we have run a new MD trajectory with the 20 ps interval for saving snapshots. δP has been evaluated from the first 120 ns; thus, the trajectory is referred to as B/120 ns. (The MD run was continued and terminated at 190 ns.) As is shown in Figure 5, the B/120 ns trajectory is in terms of symmetry much closer to the 1.2 μs trajectory than is the A/264 ns and even the A/124 ns trajectory. Our exploration of base pairing for chains A vs B confirmed that the B/120 ns trajectory is much more
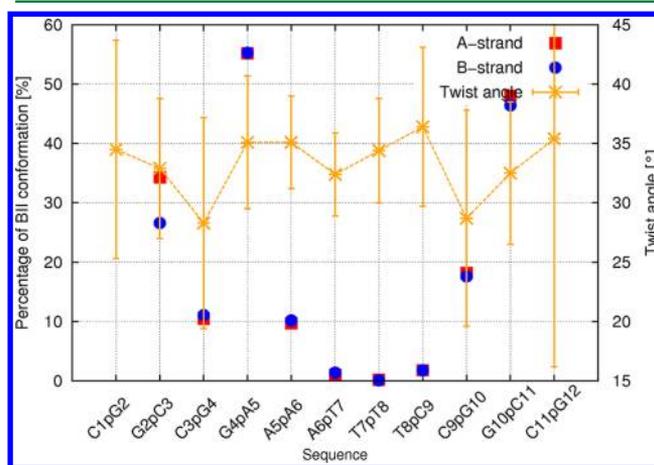
Figure 5. BII populations and twist angles for the B/120 ns trajectory. The BII conformation was assigned to all structures with $\zeta \leq 210°$.
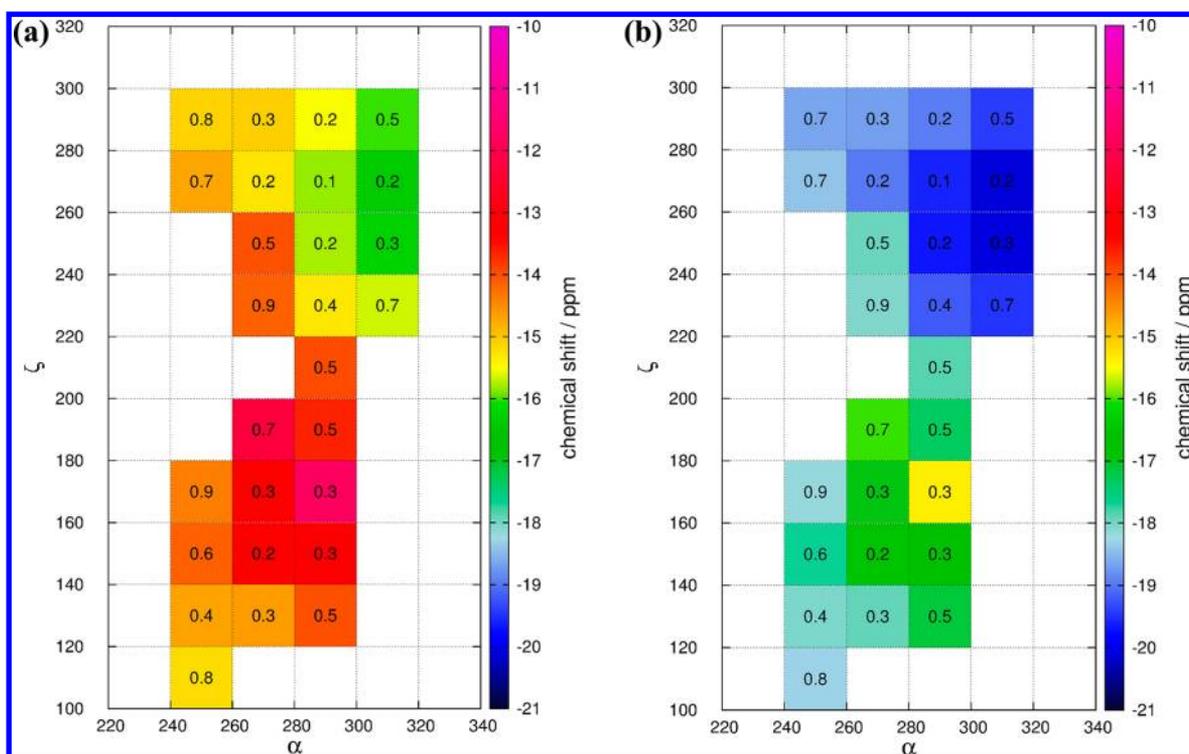
Figure 6. Dependence of δP (on color scale in ppm) on α and ζ for steps G4pA5, T7pT8, and C9pG10 computed for the models of (a) DMP and (b) IPEP, both for the A/24 ns trajectory. The numbers in the boxes give the error margins of the chemical shifts averaged over the box in question. The maps have been obtained for the sum of data obtained for the three steps in question.

symmetrical than the A/124 ns trajectory, in spite of equivalent input parameters. Thus, the numerical instability entering MD runs can very significantly influence the symmetry of the double helix.

The higher symmetry of the B/120 ns trajectory does not imply that problems of fraying at the ends of the double helix are absent. We observe corrupted behavior such as H-bond breaking and creation of non-Watson−Crick pairing, for both the A/264 ns and the B/120 ns trajectories. For the A/264 ns trajectory, until 160 ns, only short-lived unpaired states returning to Watson−Crick pairing arise. At 160 ns, pair G1−C24 splits, and at 190 ns, it undergoes a non-Watson−Crick pairing, while the pair C12−G13 splits and undergoes a non-Watson−Crick pairing only at 270 ns. For the B/120 ns trajectory, noncanonical pairing sets already on at 60 ns (pair G1−C24) and 85 ns (pair C12−G13). However, since both ends are influenced at similar time coordinates, the symmetry of the twist angles and BII populations is much higher than for the A/264 ns trajectory, cf. Figures 4 and 5.

The B/120 ns trajectory was thus selected as our most symmetrical trajectory for DFT calculations of δP. Yet, the δP/ BI and δP/BII may or may not be sensitive to BI/BII populations of individual steps and twist angles. Crucial for δP/ BI and δP/BII are correct descriptions of histograms of torsion angles, predominantly of α and ζ. The values of these angles in the two chains may but may also not depend on the BII populations, and it has not been explored yet to what precision must the histograms be converged to produce δP/BI and δP/ BII within 0.10 ppm, which is our desired precision. Consequently, in spite of the low symmetry, we have run DFT calculations also for the trajectory/sampling of A/24 ns, A/124 ns, and A/264 ns.

Focusing now on the torsion angles, in Figure S3 of the Supporting Information, we compare the histograms of torsion angles for the A/24 ns and the B/120 ns trajectories. In Figure S4 of the Supporting Information, we do the same for the B/ 120 ns and the A/264 ns trajectory. Apart from the change in BII populations with the length of the trajectory, the most apparent difference is in the β angle of the step A6pT7. Thus, the description of the β angle for A6pT7 is not converged after 120 ns. The question of how much this next-neighbor angle can influence δP/BI of A6pT7 is discussed below. Comparison of the histograms in Figures S3 and S4 with the torsion angles corresponding to the NMR structure 1NAJ[25] indicates that in the majority of cases, the NMR torsion angles fall within the full width of the half-maximum of the mainpeak of the distribution depicted in the histogram, with the following exceptions: ζ for the steps G2pC3, G4pA5, C9pG10, G10pC11, and α for the step G10pC11. In all of these cases, ζ in the histograms are shifted from the experimental ζ toward higher values. It is necessary to keep this discrepancy in mind as one possible reason for the deviation of theoretical δP from experiment, even though in the regions discussed, the δP dependence on ζ is much smaller than that on α (see below).

Another important aspect of Figures S3 and S4 is their meaning for our definition of BII conformation. The value of 210° for torsion angle ζ turns out to be an approximate average coordinate of the minimum between the BI and BII maxima. Thus, our criterion of BI vs BII conformation is the following: the BI conformation is assigned to all structures with $\zeta > 210°$, the BII conformation is assigned to all other structures.

The last aspect of our MD simulations to be discussed is the distribution of water molecules around the phosphate atoms, relevant to the definition of the second solvation sphere newly explored in this work. The graph of radial distribution function

Table 1. δP/BI for Individual Steps of DD Obtained over the MD Trajectories[a]

step	A/24 ns IPEP + 1st solvation shell	A/24 ns DMP + 1st solvation shell	A/124 ns DMP + 1st solvation shell	A/264 ns DMP + 1st solvation shell	B/120 ns DMP + 1st and 2nd solvation shell + closest Na$^+$
G2pC3	−	−15.57 ± 0.19	−16.19 ± 0.22	−16.20 ± 0.16	−15.04 ± 0.28
C3pG4	−	−15.46 ± 0.11	−15.64 ± 0.15	−15.62 ± 0.10	−14.93 ± 0.11
G4pA5	−19.61 ± 0.23	−15.70 ± 0.19	−15.77 ± 0.24	−15.66 ± 0.19	−15.11 ± 0.16
A5pA6	−	−15.64 ± 0.11	−15.98 ± 0.15	−15.99 ± 0.09	−14.99 ± 0.11
A6pT7	−	−15.75 ± 0.10	−16.33 ± 0.15	−16.39 ± 0.09	−15.64 ± 0.09
T7pT8	−19.63 ± 0.10	−15.93 ± 0.10	−16.05 ± 0.15	−16.16 ± 0.10	−15.33 ± 0.10
T8pC9	−	−15.63 ± 0.11	−16.37 ± 0.31	−16.25 ± 0.13	−15.37 ± 0.10
C9pG10	−19.09 ± 0.16	−15.46 ± 0.15	−16.09 ± 0.17	−16.06 ± 0.10	−15.03 ± 0.12
G10pC11	−	−15.68 ± 0.49	−15.91 ± 0.25	−15.80 ± 0.15	−14.85 ± 0.24

[a]Chemical shift of ^{31}P averaged over the BI parts of the trajectory and its error margin. The BI conformation was assigned to all structures with $\zeta >$ 210°.

for the distance between phosphorus and water hydrogens for trajectory A/264 ns is given in Figure S5 of the Supporting Information. The second solvation shell has been correspondingly defined as all water molecules up to the distance of 4.65 Å from phosphorus.

δP Calculations. Chemical Shift Maps for DMP vs IPEP. In our previous work,[5] we analyzed the trends in chemical shift in terms of the α and ζ torsion angles in the form of 2D maps of δP for DMP with the first solvation shell, based on the A/24 ns trajectory. A question arose as to what extent the δP maps are determined by the size of the model. In order to compare our previously calculated 2D map for the DMP model with an extended one, we have computed 2D maps using the IPEP model with the first solvation shell using the same torsion angles from the A/24 ns trajectory (Figure 6). As evidenced, the main trends are preserved, while absolute values of δP decrease by ca. 4 ppm upon model extension, in agreement with substitution of hydrogen atoms in DMP by the more bulky hydrocarbon chains in IPEP. Of importance is the fact that the error margins do not change, which indicates that the next-neighbor torsion angles β and ε, modeled more realistically in IPEP as compared to DMP, do not increase the spread of the data within the individual bins. The error margins remain inversely proportional to the population of the bins. Thus, there is no pronounced dependence on the ε and β torsion angles within the individual bins and the consideration of ε and β in the chemical shift map via their correlation to α and ζ values appears sufficient.

As indicated above, the histograms of the torsion angles from MD in several cases reveal shifts with respect to the experimental NMR structure. It should be therefore discussed to what extent the computed δP/BI and δP/BII can be influenced by these shifts. In most cases, the shift is in the torsion angle ζ, for the region of $\zeta/\alpha \sim 270°/290°$. In this region, the dependence on the ζ value in the chemical shift map is smaller than the dependence on the α value, so a minor influence on the value of δP/BI can be expected.

Sequence Dependence of δP/BI. The calculated values of δP/BI for the selected trajectories, the DMP and IPEP models, and two sizes of solvation shells are given in Table 1 and Figure 7. The presented error bars are based on the results of the tests of statistical inefficiency. These are shown for the A/24 ns, the A/124 ns, the A/264 ns, and the B/120 ns trajectories in the Figures S6, S7, S8, and S9 of the Supporting Information, respectively. The estimated statistical inefficiencies are given in Tables S1−S4 of the Supporting Information. Additionally, for the step T7pT8 and the trajectory A/264 ns, we have tested the
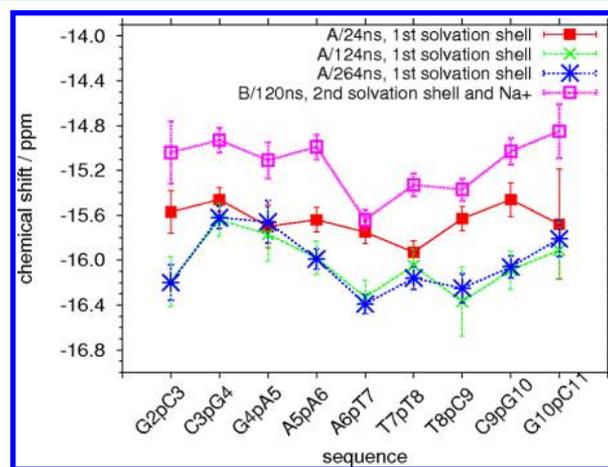
Figure 7. δP/BI as a function of sequence obtained from MD/DFT calculations for the DMP model. The BII conformation was assigned to all structures with $\zeta \leq 210°$. "2nd solvation shell and Na+" in the legend refers to employing the first solvation sphere, the second solvation sphere, and the closest Na$^+$ ion (cf. Computational Details).

effect of continually increasing the embedding from (a) the first solvation shell through (b) the first and second solvation shells to (c) the 1st + 2nd solvation shells and closest Na$^+$ ion. Our calculated values of δP/BI were (a) −16.16 ± 0.10 ppm, (b) −15.62 ± 0.25 ppm, and (c) −15.14 ± 0.10 ppm, respectively. The results of the tests of statistical inefficiency and the estimated inefficiencies are given in Figure S10 and Table S5 of the Supporting Information. Thus, both adding the second solvation shell and adding the Na+ ions has a significant influence on δP.

To check the adequacy of estimating the statistical errors, we have additionally plotted in Figure S11 of the Supporting Information the computed chemical shifts for the A/24 ns trajectory and DMP with first solvation shell in the form of histograms of δP/BI, δP/BII, and total δP. The chemical shift distributions of the two conformers are significantly overlapping, despite the clear "pooling" of the structures; thus, the standard way of estimating statistical errors is justified.

The most apparent feature of Figure 7 is the close similarity of results obtained with the A/264 ns (1st solvation shell) and A/124 ns (1st solvation shell) trajectories and their difference from the results obtained with the A/24 ns (1st solvation shell) and the B/120 ns (1st and second solvation shell + closest Na$^+$) trajectories. The trajectories A/264 ns (1st solvation shell) and A/124 ns (1st solvation shell) differ mainly in the error bars that are necessarily larger for the shorter trajectory, which is a

Table 2. δP/BII for Individual Steps of DD Obtained over the MD Trajectories[a,b]

step	A/24 ns IPEP + 1st solvation shell	A/24 ns DMP + 1st solvation shell	A/124 ns DMP + 1st solvation shell	A/264 ns DMP + 1st solvation shell	B/120 ns DMP + 1st and 2nd solvation shells and closest Na^+
G2pC3	–	-13.09 ± 0.45	-13.51 ± 0.38	-13.89 ± 0.22	-12.22 ± 0.40
C3pG4	–	-12.48 ± 0.31	-12.53 ± 0.57	-13.34 ± 0.29	-12.36 ± 0.33
G4pA5	-16.97 ± 0.23	-13.56 ± 0.19	-13.53 ± 0.20	-13.60 ± 0.15	-12.75 ± 0.14
A5pA6	–	-12.93 ± 0.26	-13.63 ± 0.49	-13.59 ± 0.30	-12.66 ± 0.33
A6pT7	–	-12.10 ± 0.71	-12.88 ± 1.34	-13.16 ± 0.76	-13.40 ± 0.91
T7pT8	$-^c$	$-^c$	$-^c$	$-^c$	$-^c$
T8pC9	–	-12.02 ± 0.84	$-^c$	-12.53 ± 0.99	-11.88 ± 0.75
C9pG10	-17.42 ± 0.27	-13.88 ± 0.25	-14.39 ± 0.40	-13.95 ± 0.22	-13.13 ± 0.24
G10pC11	–	-13.08 ± 0.55	-13.36 ± 0.22	-13.25 ± 0.18	-12.76 ± 0.26

[a]Chemical shift of ^{31}P averaged over the BII parts of the trajectory and its error margin. [b]The BII conformation was assigned to all structures with $\zeta \leq 210°$. [c]Only 1−2 structures in the trajectory corresponded to BII conformational class.

subset of the longer trajectory. However, the averages are little influenced by transition from A/124 ns to the more than twice as long A/264 ns trajectory. Thus, the A/124 ns trajectory turns out to be, to a large extent, converged with respect to δP/BI calculations.

Let us now concentrate in more detail on the similarities and differences between the various trajectory/model combinations when focusing on sequence trends in δP/BI. Trend in δP between the steps G2pC3 and C3pG4 strongly depends on the extension of the trajectory from A/24 ns to A/124 ns and A/264 ns and on increasing the solvation shell. The step G2pC3 is the penultimate steps of the duplex, as is the G10pC11 step possessing a large error bar on the δP result for the A/24 ns trajectory. It is often observed in molecular dynamics simulations that the ultimate and the penultimate steps reveal significantly larger fluctuations than the rest of the steps. Indeed, as discussed, corrupted base pairing at the ends has been observed for both trajectories A and B. Since it is usual in such cases to exclude two end steps from comparisons,[57] we adopt the same approach. With the remaining seven steps, for all trajectories and both models, δP/BI is most negative in the middle (steps A6pT7 or T7pT8) and least negative toward the ends (steps C3pG4 and C9pG10). Also for the larger IPEP model with first solvation shell, we obtained δP/BI by 0.5 ppm more negative for T7pT8 than for C9pG10 (see Table 1).

Concentrating now on the middle steps, we are in a fortunate situation since several steps have BII populations very close to zero. This is predicted jointly from our MD simulations, from the simulation of Pérez et al., and from the equation of Heddi et al. Thus, trends in δP/BI from theory and in δP from experiment can directly be compared. Experimental values of δP for A6pT7 and T7pT8 to two decimal places are equal.[58,59] This is correctly reproduced by the A/24 ns and the A/124 ns trajectories, where δP/BI for A6pT7 and T7pT8 are identical within the error bars. For the A/264 ns trajectory and the B/120 ns trajectory with the larger model, the chemical shift of T7pT8 becomes statistically less negative than that of A6pT7. This can apparently happen only due to the corresponding torsion angles α and ζ moving from their NMR values (starting structure) to their values converged within the free MD simulation, cf. Figures S3 and S4 (Supporting Information). The better agreement with experiment for the short trajectory is consistent with the indication that starting, NMR-determined (restrained MD), torsion angles seem closer to reality than their MD converged values.

Another difference between the various approaches in Figure 7 is the relative value of δP/BI for C9pG10 and C3pG4. While

the A/24 ns with the smaller solvation and the B/120 ns trajectory with the larger solvation predicts these two values to be statistically identical, the A/264 ns trajectory with the smaller model sees them as different (δP/BI for C9pG10 is more negative). Here, a comparison to experiment is not possible, since BII populations for these steps are both significant and model-dependent. Nevertheless, δP/BI is predicted to be less negative for C9pG10 than for A6pT7 jointly by all models.

Quantitatively, considering the seven innermost steps, the difference between the most and least negative δP/BI is 0.77 ppm for the A/264 ns trajectory with small solvation, and it is 0.71 ppm for the B/120 ns trajectory with large solvation. Since our calculations overestimate the experimental δP/BI−δP/BII (1.6 ppm) by a factor of ca. 1.7 (vide infra), we can divide the differences between the most and least negative δP/BI by this number to obtain an estimate of sequence-dependence of δP/BI in DD: 0.77/1.7 = 0.45 ppm and 0.71/1.7 = 0.42 ppm, respectively.

Our result, the 0.4−0.5 ppm sequence dependence of δP/BI, is sufficient to reveal that the conventionally adopted presumption of sequence independence of δP/BI and δP/BII is not justified. Thus, the population of BII states determined using this presumption and currently employed as a benchmark for MD validations most probably cannot be taken as quantitatively relevant. It might perhaps serve as a qualitative guide if, as in our particular case, the sequence dependence of δP/BI, which is not accounted for by Heddi et al., would to some extent be correlated to BII population (in our case, both are smallest in the middle and largest at the end of the duplex). Indeed, a simple model may work qualitatively if a neglected yet important property is correlated to a property, which is included.

Sequence Dependence of δP/BII. Our data for the sequence dependence of δP/BII are shown in Table 2 and Figure 8. Table 2 lists data for steps where more than 30 snapshots fell into the BII region. Figure 8 shows data for steps where the computed error bar is smaller than 0.6 ppm. In comparison to δP/BI, data for δP/BII are statistically much less conclusive due to the relatively small occurrence of BII and, hence, large error bars. Unlike for δP/BI in Figure 7, the trends for δP/BII in Figure 8 reveal different qualitative trends not only between the A/24 ns and the A/124 ns trajectories but also between the A/124 ns and the A/264 ns trajectories. Thus, the δP/BII probably cannot be considered converged, even at 264 ns, due to insufficient sampling of the BII conformations. Yet, there is one statistically significant trend observed for all
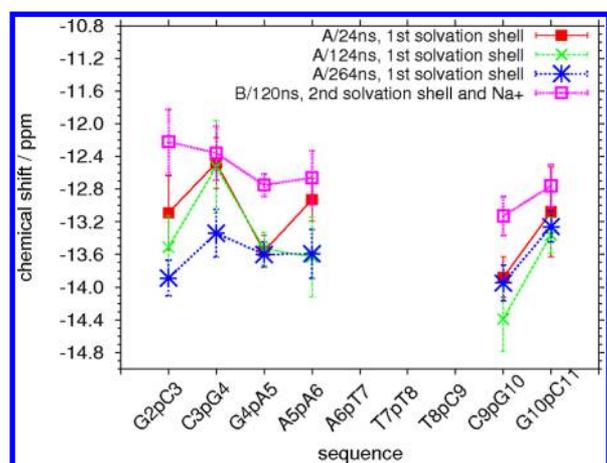
Figure 8. δP/BII as a function of sequence obtained from MD/DFT calculations for the DMP model. "2nd solvation shell and Na+" in the legend refers to employing the first solvation sphere, the second solvation sphere, and the closest Na$^+$ ion (cf. Computational Details).

trajectories and models explored: δP/BII for C3pG4 is always less negative than for C9pG10. The most conservative estimate of the difference, δP/BII(C3pG4) − δP/BII(C3pG4), is 0.10 ppm for the A/264 ns trajectory.

A last issue to be discussed at this point is the number of snapshots necessary for the δP calculation. In this study, all computed δP are averages over 6000 snapshots with the exception of the A/124 ns trajectory where only 2500 snapshots were available. While for 6000 snapshots, the minimum error bar is 0.10, for 2500 snapshots, it is 0.15. To learn more about the relation between snapshot count and error bar size, we have compared, for the B/120 ns trajectory, averages and error bars using (a) all 6000 snapshots, (b) each second snapshot, in total 3000 snapshots, and (c) each sixth snapshot, in total 1000 snapshots. The results are shown in Tables S6 and S7 (Supporting Information); the results of tests of statistical insufficiency and estimated insufficiencies are shown in Figures S9, S12, S13 and Tables S4, S8, S9 of the Supporting Information, respectively. As is evidenced in Tables S6 and S7, 1000 structures is clearly insufficient. The error bars are at least twice as large as for 6000 structures, which is a precision much too low for the present purpose. For 3000 snapshots, all but one error bar are again larger than for 6000 snapshots. In the one exceptional case of C3pG4, the decrease in the error bar is within the order of uncertainty in error bar determination, which is for BI conformations 0.02−0.03. In most cases, however, upon halving the 6000 data to 3000, error bars increase by only ca. 10−20%, so that still statistically significant conclusions about averages can be drawn, cf. Table S6 of the Supporting Information.

Sequence Dependence of Total δP and Comparison to Experiment. A successful calculation of total δP comparable to experiment requires that δP/BI and δP/BII are corrected for (1) a constant offset of the theoretical values and (2) the theoretical overestimation of the difference between δP/BI and δP/BII. In the present case, the computed values of δP/BI reveal a constant offset of ca. 11 ppm with respect to experimental values of refs 58 and 59, when we compare the steps predominantly found in the BI state. This offset originates partly from the application of a limited model[60] and partly from systematically underestimated P−O bond lengths used in the molecular dynamics as compared to the

DFT-optimized geometries.[5] The latter problem can be partially overcome by using geometries from CPMD snapshots, as discussed below. To correct for the theoretical offset, we have computed values δP/BI$_{offset}$ and δP/BII$_{offset}$ (Table S10 and S11 of the Supporting Information), determined as δP/BI$_{offset}$ = δP/BI + c, δP/BII$_{offset}$ = δP/BII + c, where the constant c has been determined so that δP/BI$_{offset}$ of the step A6pA7 computed for the trajectory and model in question matches exactly the experimental value of −4.38 ppm.[58] The step A6pA7 has been chosen as a reference for the offsetting due to the fact that this step is found almost exclusively in the BI conformation; hence, the theoretical value of δP/BI can be directly compared to the experimental chemical shift. Furthermore, for the step A6pA7, the torsion angles from the MD histograms and from NMR structures agree very well (cf. Figures S3 and S4, Supporting Information); A6pA7 thus appears as an ideal step for the computational method calibration.

The differences between δP/BI and δP/BII in Tables 1 and 2 for the A/24 ns trajectory range from 1.58 ppm (C9pG10) to 3.65 ppm (A6pT7), with an average of 2.72 ppm. The experimentally estimated average value of the difference δP/BI − δP/BII is 1.6 ppm,[1,2] indicating that present methods applied to the A/24 ns trajectory overestimate experiment by a factor of 2.72/1.6 = 1.70. This is partially caused by the limited solvent embedding, but even with the inclusion of second solvation shell and of Na$^+$, δP/BI − δP/BII is still overestimated on average by a factor of 1.5 with respect to experiment. Another possible reason might be the limitations of the BP86 density functional. Our comparative calculations done on a few snapshots with the popular hybrid B3LYP functional provided an even slightly larger value of δP/BI$_{offset}$ − δP/BII$_{offset}$; however, for δP with its large fluctuations between individual snapshots, the averaging over the whole trajectory would have to be done before a firm conclusion can be made. Scaling might be caused also by working with a limited DMP model and neglecting the sugar and base contributions to chemical shift. However, going from the DMP model to the IPEP model, the δP/BI − δP/BII difference gets even larger (cf. third and second columns of Tables 1 and 2). We believe that δP/BI − δP/BII overestimation in the calculations can be accounted for by slight bond length and bond angle responses to the torsion angle changes, which are clearly observed in CPMD simulations (see below) and which influence the values of δP/BI and δP/BII differently. To correct for the systematic offset and the scaling at the same time, we have computed "δP/BI$_{offset,scaled}$" and "δP/BII$_{offset,scaled}$" as follows. First, we determined the difference δP/BI(X) − δP/BI(A6pT7). The result is then divided by a factor of 1.7 and added to the value of δP/BI$_{offset}$(A6pT7). Similar procedure is used for δP/BII(X). The values of δP/BI$_{offset,scaled}$ and δP/BII$_{offset,scaled}$ are given in Tables S12 and S13 of the Supporting Information, with the details of calculation given in the footnotes.

For the comparison with experimental data, as most suitable appear the A/24 ns and the A/124 ns trajectories that correctly predict δP/BI for A6pT7 and T7pT8 identical within the error bars. While for the case of the A/124 ns trajectory this is partially due to the limited number of snapshots and thus larger error bars (for the related, B/120 ns trajectory, disagreement in δP/BI for A6pT7 and T7pT8 is found), the results for the A/24 ns trajectory are more statistically conclusive. We thus continue our discussion with the shortest, A/24 ns trajectory and, for comparison, also with the longest, A/264 ns trajectory. The δP/
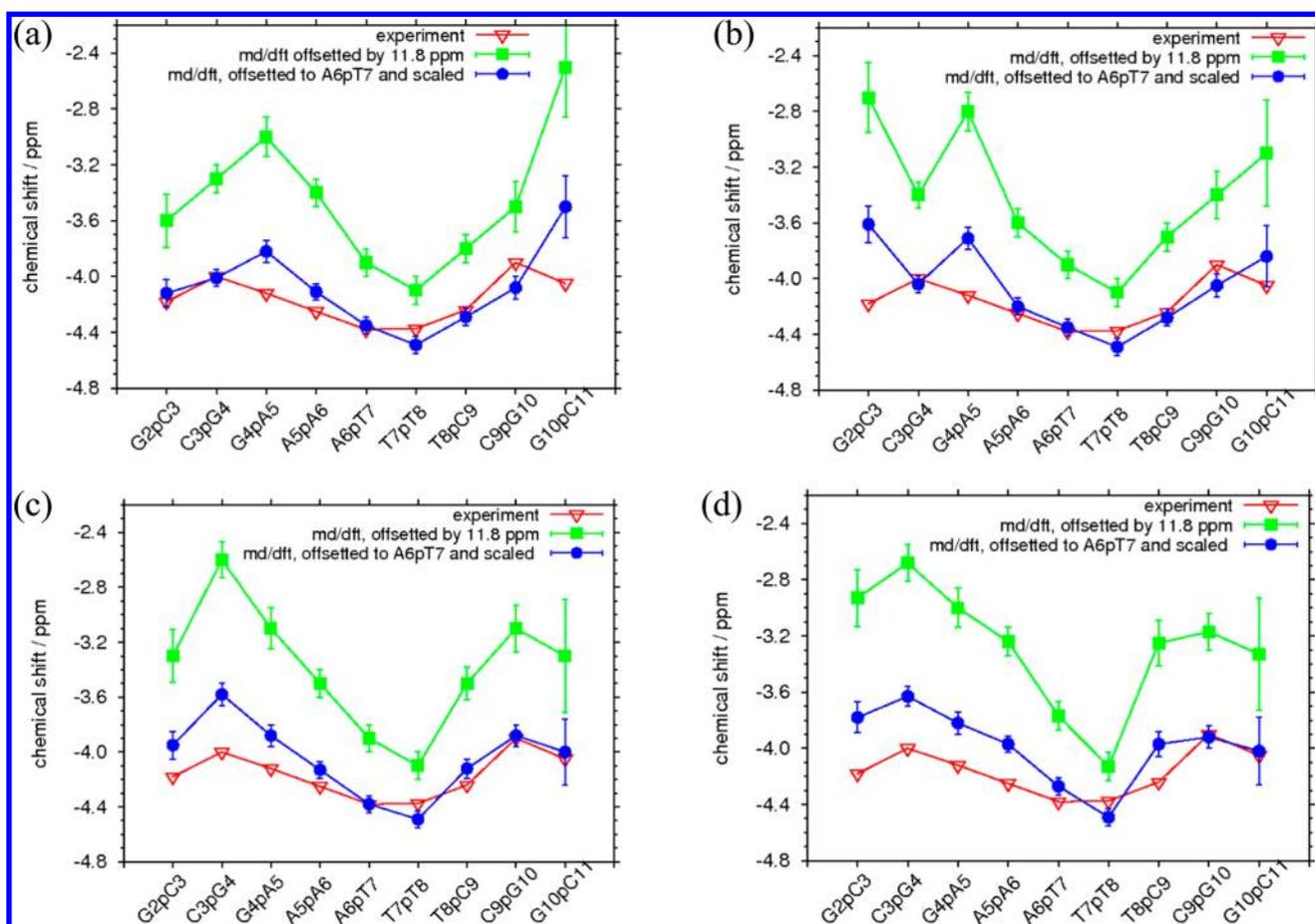
Figure 9. Experimental shifts from ref 58 (red) and theoretical shifts for the A/24 ns trajectory, computed either as weighted averages of $\delta P/BI$ and $\delta P/BII$ values from Tables 1 and 2 and shifted for graphical reasons (to better see the comparison of trends but to avoid overlap with the blue line) by 11.8 ppm up (green) or as weighted averages of $\delta P/BI_{offset,scaled}$ and $\delta P/BII_{offset,scaled}$ values from Tables S12 and S13 (blue). The weighted averages have been computed according to eq 1 except for the step T7pT8, where pure $\delta P/BI$ and $\delta P/BI_{offset,scaled}$ values are plotted, due to the absence of theoretical $\delta P/BII$ and $\delta P/BII_{offset,scaled}$ values and negligibly small percentages of BII structures. The weighting factors c_{BI}, c_{BII} have been taken from (a) our A/24 ns simulation, average BI/BII percentage over the two chains; (b) simulation of Pérez et al.,[23] average BI/BII percentage over the two chains; (c) values of Heddi et al.,[3] and (d) Schwieters et al.[22] The plotted error margins have been determined as described in the Supporting Information.

BI, $\delta P/BII$ subaverages (Tables 1 and 2) as well as the $\delta P/BI_{offset,scaled}$, $\delta P/BII_{offset,scaled}$ subaverages (Tables S12 and S13 of the Supporting Information) have been employed to calculate their weighted averages using eq 1:

$$\delta P = c_{BI}\delta P/B_I + c_{BII}\delta P/B_{II} \qquad (1)$$

Relative occurrences of BI vs BII, characterized by coefficients c_{BI} and c_{BII} have been taken from (a) our MD simulation in question, (b) from MD simulation of Pérez et al.,[23] (c) from empirical equation of Heddi et al.,[3] and (d) from restrained MD study of Schwieters et al.[22] The results are shown in Figure 9 (A/24 ns trajectory) and Figure 10 (A/264 ns trajectory).

An inspection of Figures 9 and 10 reveals that (1) after scaling (blue curves), an agreement between theory and experiment becomes much more evident; (2) the sequence trends agree with experiment better for the A/24 ns trajectory than for the A/264 ns trajectory; (3) the best agreement with experiment is found for BII populations of Heddi et al (Figure 9c); (4) a good agreement with experiment is found also for BII populations from MD simulations if steps G4pA5, G10pC11 in Figure 9a and steps G2pC3, G4pA5, G10pC11 in Figure 9b are excluded from the comparison. The latter steps are exactly

those that have within the corresponding MD simulations populations of BII larger than 30%. This suggests that the disagreement with the experimental trend can be due to overestimation of percentages of BII at the particular steps of the MD trajectory. In other words, at the steps with the high propensity for BI–BII transitions, the force field may overestimate the actual BII percentage. In Figure 9(c), on the contrary, all of the experimental trends are matched correctly indicating that the introduction of NMR restraints in the modeling of the BI vs BII populations indeed brings the description closer to reality. Quantitatively, the theoretical trends for the first three steps are, however, much too steep. This, we believe results from the quantitative inadequacy of empirical equation of Heddi et al. that neglects the sequence dependence of $\delta P/BI$.

Since, for the A/24 ns trajectory, we obtain a quite good agreement between the theory and experiment, we can use eq 1 to calculate c_{BI} and c_{BII}, when using ofsetted and scaled values of $\delta P/BI$, $\delta P/BII$ to match the experimental value of total δP exactly. Thus, by relating $\delta P(X)_{exp}$ for each step to sequence dependent $\delta P/BI_{offset,scaled}(X)$ and $\delta P/BII_{offset,scaled}(X)$ via eq 2:
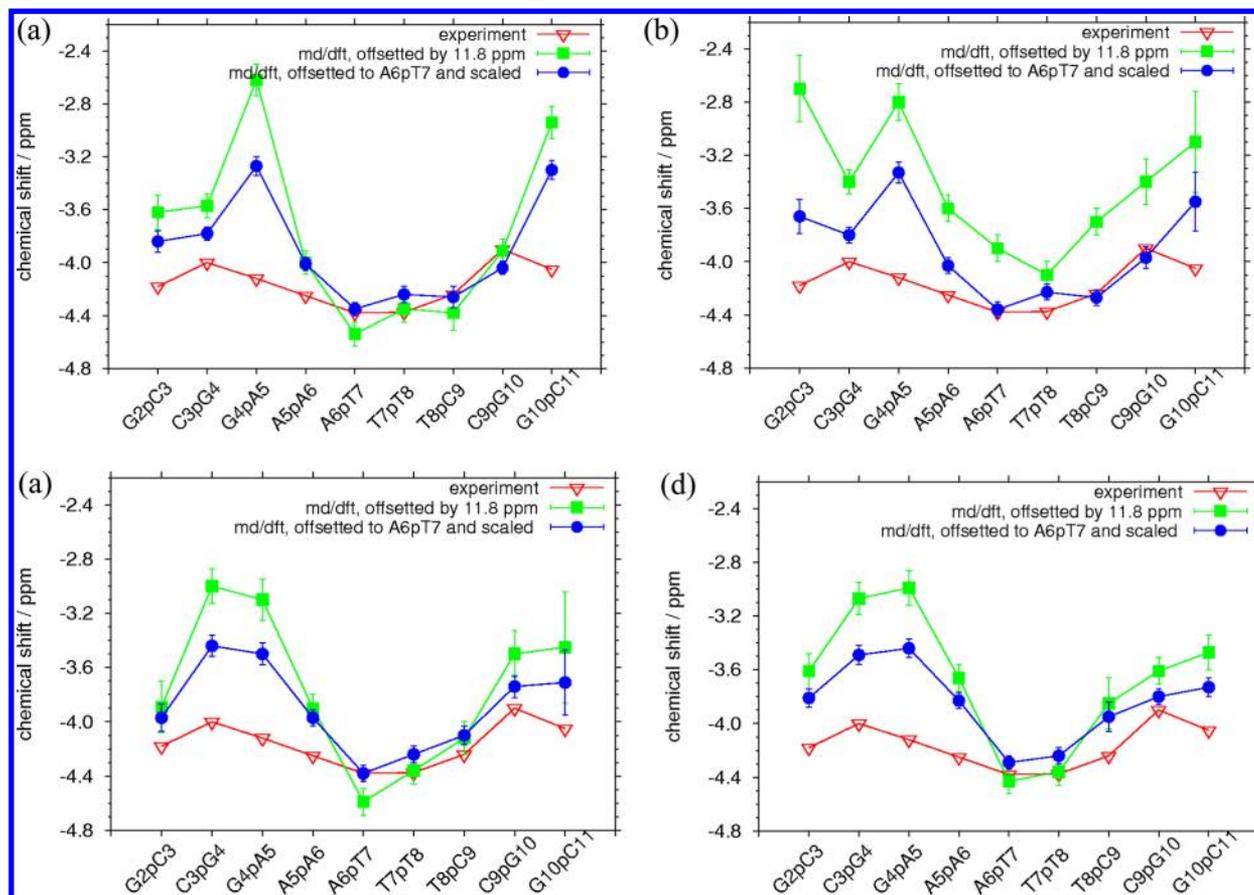
Figure 10. Experimental shifts from ref 58 (red) and theoretical shifts for the A/264 ns trajectory, computed either as weighted averages of $\delta P/BI$ and $\delta P/BII$ values from Tables 1 and 2 and shifted for graphical reasons (to better see the comparison of trends) by 11.8 ppm up (green) or as weighted averages of $\delta P/BI_{offset,scaled}$ and $\delta P/BII_{offset,scaled}$ values from Tables S12 and S13 (Supporting Information) (blue). The weighted averages have been computed according to eq 1 except for the step T7pT8, where pure $\delta P/BI$ and $\delta P/BI_{offset,scaled}$ values are plotted, due to the absence of theoretical $\delta P/BII$ and $\delta P/BII_{offset,scaled}$ values and negligibly small percentages of BII structures. The weighting factors c_{BI}, c_{BII} have been taken from (a) our A/264 ns trajectory, chain A (b) simulation of Pérez et al.,[23] average BI/BII percentage over the two chains; (c) values of Heddi et al.,[3] and (d) Schwieters et al.[22] The plotted error margins have been determined as described in the Supporting Information.

$$\delta P_{exp}(X) = c_{BI}(X)\delta P/BI_{offset,scaled}(X)$$
$$+ c_{BII}(X)\delta P/BII_{offset,scaled}(X) \qquad (2)$$

we can determine $c_{BI}(X)$ and $c_{BII}(X)$. Realizing that $c_{BI}(X) = 1 - c_{BII}(X)$, we obtain for %$B_{II}(X) = c_{BII}(X) \times 100$:

$$\%B_{II}(X) = \frac{\delta P_{exp}(X) - \delta P/BI_{offset,scaled}(X)}{\delta P/BII_{offset,scaled}(X) - \delta P/BI_{offset,scaled}(X)}$$
$$\times 100 \qquad (3)$$

The error margin of computed %B_{II} has been determined as described in the Supporting Information.

In Figure 11, the percentage of BII determined from eq 3 by using values of Tables S12 and S13 (Supporting Information) for the A/24 ns trajectory and the first solvation shell ("current NMR calculation") is plotted and compared to relative occupation of BII determined from the other sources. The best qualitative agreement with our NMR calculation is found for the results of Heddi et al. At the same time, however, our calculation suggests that for the steps G2pC3, C3pG4, and G4pA5 the percentage of BII should be smaller than suggested by Heddi et al., due to the sequence dependence of $\delta P/BI$.

The results of Schwieters et al.[22] likewise suggest significantly larger BII relative occurrence than our calculation for the first

four steps, while for the remaining five steps their error margins fall within the error margins of our calculation. For the two MD simulations, quantitative agreement is obtained with our NMR calculation for the innermost steps, while the MD description of the end steps seems clearly problematic: The simulation of Pérez et al. seems to overestimate the BII percentage of G2pC3 and G4pA5, while our MD simulation seems to overestimate the BII percentage of G4pA5 and G10pC11. Both our simulation and the simulation of Pérez et al., on the other hand, underestimate the percentage of BII for C9pG10.

δP Calculations on CPMD Snapshots. Calculations on snapshots from Car−Parrinello molecular dynamics of DMP have been done in order to see whether these can partially correct for the systematic offset of the theoretical δP as compared to the experiment. Two CPMD trajectories were employed. The first CPMD trajectory was started from a $-g-g$ structure while the second one from a $+gt$ structure. The trajectories contained all of $+g+g$, $+g-g$, $-g+g$, $-g-g$, $+gt$, $-gt$ regions, while for the *tt* conformation (both torsion angles 180° ± 20°) only two structures were encountered. The time dependence of the C−O−P−O torsion angles over the two trajectories is given in the Supporting Information (Figure S14).

The chemical shift map computed on the CPMD trajectories is given in Figure 12. Chemical shifts were averaged for bins
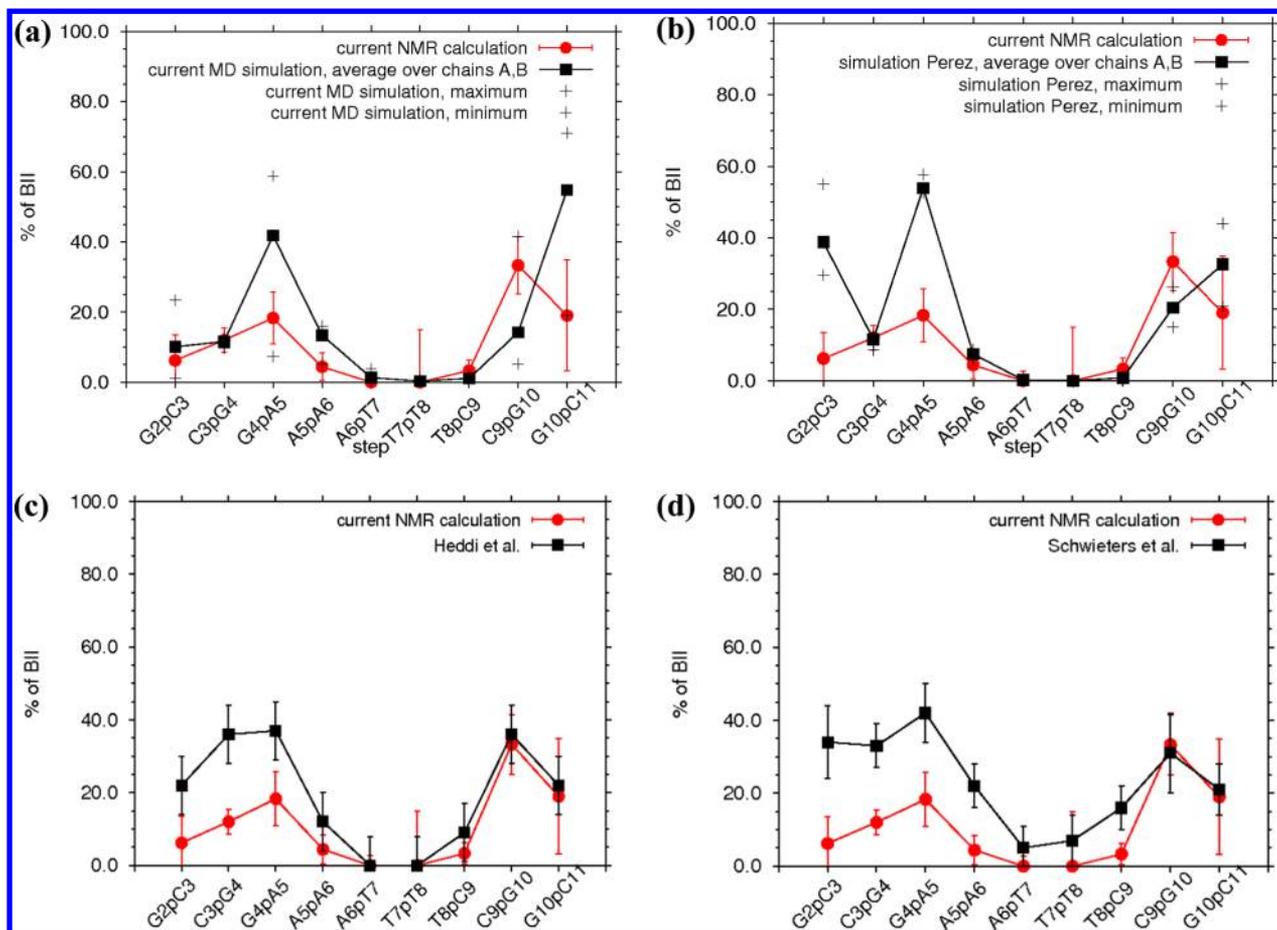
Figure 11. Percentage of time in the BII conformation. Current NMR calculation: average values determined by eq 3 using the values from Tables S12 and S13 for the A/24 ns trajectory, except for step T7pT8 where eq 3 would predict a negative percentage and the value of 0.0 is plotted. Error margins have been determined using equations S10, S11, and S12 (Supporting Information). The values from current NMR calculation are compared to percentage of BII from (a) current MD simulation, average over the two chains; (b) MD simulation of Pérez et al., taken from ref 23 (average over the two chains); (c) Heddi et al.,[3] taken the "predicted" values from Table 8 of ref 22; error margins taken from ref 3; (d) Schwieters et al., taken the $N_e = 8$ values of Table 8 of ref 22. The crosses in panels a and b indicate maximum and minimum values for the two halves of each trajectory and each strand. The black error bars in panels c and d are published error estimates. The red error bars are error estimates of our calculated data obtained as described in the Supporting Information.
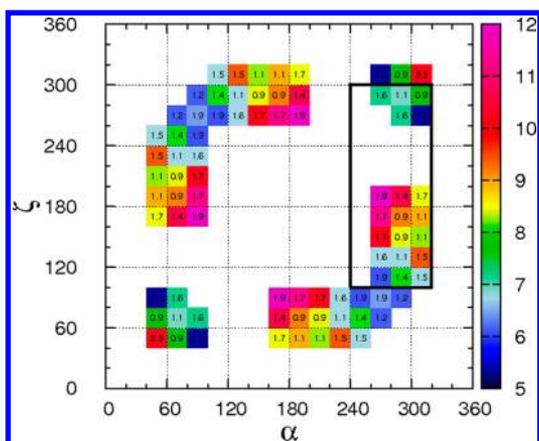
Figure 12. Chemical shift map for DMP computed on snapshots from the two CPMD trajectories described in text. The numbers in the boxes give the error margins of the chemical shifts averaged over the box in question. The highlighted rectangle corresponds to the range of values in Figure 6.

with populations of at least 25 snapshots. The error margins in Figure 12 are significantly larger than those for the chemical

shift maps from classical MD due to much smaller population of the individual bins. In terms of trends, in agreement with the classical MD maps, the least positive shifts are obtained for the *gg* conformation, while the most positive shifts are obtained when one of the angles is ∼90° (*g*) and the other is *t*. In terms of the absolute values, the whole scale is shifted to significantly more positive values. The chemical shifts for the −*g*−*g* conformation in Figure 12 are found around +7 ppm. Based on results from classical MD upon the transition from DMP to IPEP (Figure 6), the phosphorus chemical shift decreases by ca. 4 ppm. Similarly, in the case of CPMD calculations, we may expect an analogous shift of a *gg* structure of IPEP to ca. +3 ppm, which is already much closer to the experimental nucleic acid value of ca. −4 ppm[58] than in the case of classical MD snapshots, where the values of chemical shift in the middle of the *gg* region are ca. −20 ppm, cf. Figure 6b. Similar improvement of comparison between experiment and theory upon transition from classical MD to CPMD has been reported previously.[61] The reason is in the more realistic bond lengths provided by *ab initio* molecular dynamics: the CPMD distances listed in Table 3 are within ca. 0.02 Å of the DFT-optimized

Table 3. Averaged Interatomic Distances [Å] Found in the CPMD Simulation,[a] in DFT Optimization,[b] and in Classical MD Simulation[c,d]

	P = O2	P = O3	P—O4	P—O5	O4—C6	O5—C7
CPMD, −gt	1.530	1.526	1.630	1.619	1.475	1.478
CPMD, −g−g	1.524	1.524	1.627	1.627	1.475	1.475
DFT optimized, −gt	1.544	1.534	1.663	1.645	1.449	1.453
DFT optimized, −g−g	1.532	1.533	1.662	1.661	1.449	1.450
classical MD	1.477	1.477	1.599	1.595	1.424	1.423

[a]The standard errors of the mean were close to 0.001 Å in all cases. [b]Results for the DMP model with 6 H_2O molecules, the whole system optimized at the BLYP/6-31G(d) level of theory. [c]Results for a parm99 FF trajectory. [d]The numbering of atoms is given in Figure 13.

bond lengths for DMP, while the classical MD bond lengths are ca. 0.05 Å shorter than the DFT-optimized bond lengths.
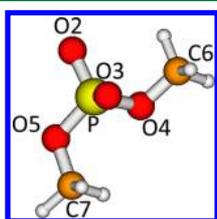
Figure 13. Atom numbering definition for Tables 3 and 4.

DFT Calculations on the Grid of α and ζ. The chemical shift map for CPMD snapshots on free dimethyl phosphate shown in Figure 12 covers a much wider torsion-angle space than the map of chemical shielding for the DD system. In order to explore the chemical shift throughout the full conformational space of 360° for α and ζ, we have performed a series of DFT calculations on a grid of α and ζ. The results are consistent with the chemical shift maps for the DD system as well as with the CMPD map for the dimethyl phosphate and are given in Figure S15 of the Supporting Information.

Analysis of MO Contributions to Phosphorus Chemical Shielding. Interpretation. Finally, let us concentrate on the interpretation of the conformational dependence of δP studied in this manuscript. Whereas the torsion angle dependence of δP in DMP has not been interpreted in terms of orbital interactions so far, stereoelectronic interpretations have been suggested for the conformational energy of DMP.[62] A stereoelectronic influence related to a specific feature of the electronic structure that results in the stabilization of one conformer over another is called the anomeric effect.[63] In the case of DMP, the −g−g (closed) geometries are more stable than the −gt (half-open) geometries,[64] which are, in turn, more stable than the tt (open) geometries.[65] The conventional view of the anomeric effect in the phosphodiester backbone is illustrated in Figure 14. The −g−g conformation is stabilized (with respect to −gt and tt conformations) through O_{ester} lone pair electron interactions with spatially proximal P—O_{ester} antibonding orbitals.[66] We may therefore expect that a similar
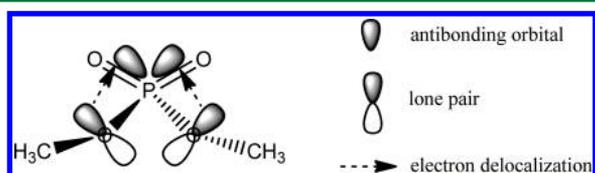
Figure 14. Conventional hypothesis of the anomeric effect responsible for the stabilization of the phosphodiester backbone in the gg geometries. Along the lines of Figure 1 in ref 67.

factor will determine the difference in ^{31}P chemical shift between the gg and the gt conformers.

Our analysis of the ^{31}P chemical shielding tensors for the −g−g and −gt conformations (see Supporting Information, Table S14) shows that while the two smaller principal components (σ_{11} and σ_{22}) vary only slightly upon the −g−g → −gt switch, the largest principal component (σ_{33}) varies by approximately 10 ppm. This is in agreement with the results of Benda et al.[7] In Table 4, the σ_{33} component is analyzed in terms of localized molecular orbital contributions, using the IGLO approach along with the Boys localization scheme. Obviously, the −g−g → −gt switch influences most the contribution of the O3−P bond and only in the second place the O4−P bond. It is thus the bond to the anionic oxygen, not to the ester oxygen, whose contribution is most influenced. Consequently, in addition to the anomeric effect described, another stereoelectronic effect involving the anionic oxygens should be expected.

As has been recently shown, the anomeric effect in the phosphodiester moiety of DNA is indeed more complicated than previously suggested.[67] Banavali and MacKerell have shown that, in addition to delocalizations involving O_{ester} lone pairs and P−O_{ester} antibonding orbitals, delocalizations involving $O_{anionic}$ lone pairs and P−$O_{anionic}$ antibonding orbitals also occur. The presence of an anomeric effect can be detected by bond length and angle changes. In agreement with the results of Banavali and MacKerell,[67] in our CPMD simulations the bond length to one of the anionic oxygens changes only slightly upon the gg→gt switch while the other is significantly lengthened, see Table 3. This indicates that indeed the changes in δP reported in this work result from a stereoelectronic effect involving the P−$O_{anionic}$ bonds.

■ CONCLUSIONS

In our previous work, we have found that δP varies within the BI and BII regions in the order of units of ppm, where the regions were defined as statistically significantly populated spaces of α and ζ in the course of free MD simulation of $[d(CGCGAATTCGCG)]_2$.[5] The main question to be answered here was whether the α, ζ dependence of δP/BI and δP/BII will remain apparent also in the sequence dependence of time averaged δP/BI and δP/BII of the canonical B-DNA duplex. This question is anything but a trivial one, since the statistically significant population of α and ζ spans a much wider region of angles than are the differences in the mean values of α and ζ, which characterize individual steps.

We have explored three molecular dynamics trajectories differing in the step size for saving snapshots and simulation length and three quantum chemical models of phosphorus surroundings. A common result obtained for all MD trajectories

Table 4. Localized Molecular Orbital (LMO) Contributions to σ_{33} in Principal Axes^a

LMO	−g−g		−gt		contrib(−gt)−contrib(−g−g)
	MO no.	contrib.	MO no.	contrib.	
LP (O4)	26	−4.7	20	−4.4	0.4
LP (O4)	28	−5.9	29	−5.8	
LP (O5)	19	−5.7	18	−6.1	−1.9
LP (O5)	29	−4.5	24	−6.0	
LP (O2)	24	−14.5	30	−13.7	0.8
LP (O3)	25	−14.5	14	−15.8	−1.3
B (O2−P)	14	6.7	15	6.3	2.6
B (O2−P)	32	21.2	26	22.0	
B (O2−P)	30	2.4	33	4.6	
B (O3−P)	15	4.4	27	−1.1	−16.1
B (O3−P)	31	21.3	28	18.9	
B (O3−P)	33	6.7	32	−1.5	
B (O4−P)	18	−180.9	16	−171.1	9.2
B (O5−P)	27	−182.6	31	−187.3	−4.5
B (O4−C6)	13	−7.5	25	−5.5	2.0
B (O5−C7)	22	−7.5	13	−8.0	−0.5

^{a}The numbering of atoms is defined in Figure 13.

and all models is that δP/BI and δP/BII are sequence dependent in the order of tenths of ppm. Disregarding two steps at each end of the double helix (which are expected to be influenced by end-effects), in all simulations and for the DMP model with both the smaller and the larger solvent/ion embedding, δP/BI is most negative in the middle (steps A6pT7 or T7pT8) and least negative toward the ends (steps C3pG4 and C9pG10). Also, for the larger IPEP model with first solvation shell, which was explored for the steps G4pA5, T7pT8, and C9pG10, we obtained δP/BI to be 0.5 ppm more negative for T7pT8 than for C9pG10.

The shortest (A/24 ns) trajectory correctly predicts the two steps A6pT7 and T7pT8 to have the smallest chemical shifts, while for the A/264 ns trajectory and the B/120 ns trajectory with the larger model, the chemical shift of T7pT8 becomes statistically less negative than that of A6pT7. This indicates that the NMR-determined values of torsion angles α and ζ that are, to a large extent, preserved during the first 24 ns of simulation are consistent with experimental values of δP, whereas extending the MD trajectory further moves to a force-field converged torsion angle values providing a worse agreement with experiment.

The difference between the most and least negative δP/BI is 0.77 ppm for the A/264 ns trajectory and DMP with first solvation shell and 0.71 ppm for the B/120 ns trajectory and DMP with first and second solvation shell and the closest Na$^+$ ion. After scaling of our computed data by a factor of 1/1.7, we obtain an estimate of sequence-dependence of δP/BI in DD of 0.4−0.5 ppm. Our data for sequence dependence of δP/BII are much less statistically conclusive due to the smaller relative occurrence of this conformation and hence larger error bars.

Our result for the 0.4−0.5 ppm sequence dependence of δP/BI is sufficient to reveal that the conventionally adopted presumption of sequence independence of δP/BI and δP/BII is not justified. Thus, the population of BII states determined using the relationship of Hartmann et al. and currently employed as a benchmark for MD validations most probably cannot be taken as quantitatively relevant. It might perhaps serve as a qualitative guide, if, as in our particular case, the sequence dependence of δP/BI that is not accounted for by Hartmann's relationship would to some extent be correlated to

BII population (in our case, both are smallest in the middle and largest at the end of the duplex). Indeed, a simple model may work qualitatively if a property that is neglected is correlated to a property that is included.

In parallel with the MD/DFT calculations of chemical shifts, we tried to find out the conditions that must be fulfilled by a MD trajectory in order to use it as a starting point for quantum chemistry calculations. The calculation of δP requires a knowledge of δP/BI, δP/BII, and the BI and BII populations. While the populations need at least 1.2 μs of trajectory to be converged, δP/BI are converged already for a A/124 ns trajectory, since upon an extension to A/264 ns, sequence trends in average values do not change significantly. Thus, the asymmetries in BII populations, different for A/264 ns than for the A/124 ns trajectory, do not seem to influence δP/BI differently. In other words, the histograms of relevant torsion angles need extension from tenths to hundreds of nanoseconds, but not to microseconds. The necessary length of δP/BII is certainly longer than 240 ns for the middle steps, but there, the population of BII seems to be negligible, so that the value of δP/BII is of little importance for the total δP. For the outer steps, an extension of trajectory beyond 240 ns seems necessary unless shortening of step and, hence, better sampling would prove sufficient to obtain better statistics. Also important is the number of snapshots necessary for δP calculations. Averaging over 6000 snapshots results in lowest error bars of 0.09−0.11 ppm, which is our desired accuracy. For 3000 structures, error bars increase by ca. 10−20%, so that still statistically conclusive outcomes are obtained. 1000 structures is insufficient, since the error bars are at least twice as large as for 6000 structures. Very recently, a parallel variable selection method has been proposed by one of us that efficiently reduces the number of clusters needed for spectroscopic properties evaluation.[68] In the future, it would be very desirable to test the method for the system and property studied in this work.

Regarding comparison with experimental δP for all inner steps, not just those being predominantly BI, the best qualitative agreement with experiment is obtained for the A/24 ns trajectory with BI and BII populations of Hartmann et al. This is consistent with the indication that starting, NMR determined, torsion angles appear to be closer to reality than

their MD-converged values, and implicates that indeed, from the qualitative point of view, the approach of Hartmann et al. is justified when sequence dependence of δP/BI is correlated to sequence dependence of BII population.

Apart from studying qualitative trends in δP/BI and δP/BII, we concentrated also on the question of getting into closer agreement with experiment regarding absolute value of δP. While part of the systematic theoretical offset certainly comes from neglecting the longe-range contributions in the model, our calculations on snapshots from CPMD overcome most of the offset of δP due to improved average bond length as compared to the MD bond length. Our analysis of the contributions to δP implies that the difference in δP of the g,g vs the g,t conformations arises due to stereoelectronic effects involving the P$-$O$_{anionic}$ bonds. The findings made in this work provide transparent structure$-\delta$P relationships for possible use as NMR restraints and suggest that NMR calculations on MD snapshots can be in the future employed for the validation of newly developed FFs.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Use of symmetry for the enhancement of chemical shift map data from Car$-$Parrinello molecular dynamics; computational details on chemical shift calculations; calculation of statistical errors; time development of percentage of BII conformation for the step G4pA5; histograms of torsion angles for inner steps of DD for the A/24 ns, the B/120 ns, and the A/264 ns trajectories; graph of radial distribution function for the distance between phosphorus and water hydrogen atom; results of the test of statistical inefficiency for the A/24 ns, A/124 ns, A/264 ns, and B/120 ns trajectories; estimated statistical inefficiencies for the A/24 ns, A/124 ns, A/264 ns, and B/120 ns trajectories; results of the tests of statistical inefficiency for the A/264 ns trajectory and step T7pT8 with increasing solvation, along with estimated statistical inefficiencies; histograms of the ^{31}P chemical shift for the BI, the BII, and the total distribution for the A/24 ns trajectory; δP/BI and δP/BII for DMP with large solvation as functions of number of snapshots along with statistical inefficiency data; δP/BI and δP/BII subject to offset and subject to offset and scaling; the determination of values and error bars in Figures 9, 10, and 11; torsion angle C$-$O$-$P$-$O vs time for CPMD on DMP; dependence of δP on α and ζ for DMP with α, ζ varied on a grid; the decomposition of ^{31}P shielding tensors in DMP (in ppm) into their diamagnetic and paramagnetic parts, respectively. This information is available free of charge via the Internet at http://pubs.acs.org

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: marketa@chemi.muni.cz.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Gorenstein, D. G. *Chem. Rev.* **1994**, *94*, 1315−1338.
(2) Gorenstein, D. G.; Meadows, R. P.; Metz, J. T.; Nikonowicz, E. P.; Post, C. B. *Advances in Biophysical Chemistry*; Bush, C. A., Ed.; JAI Press: Greenwich, 1990; pp 47−124. Although ref 1 refers to this work for the value of 1.6 ppm, we couldn't find any explicit discussion here. We deduce that it has been obtained by an extrapolation of data in Figure 28 of ref 2 to a periodical function of ε/ζ.
(3) Heddi, B.; Foloppe, N.; Bouchemal, N.; Hantz, E.; Hartmann, B. *J. Am. Chem. Soc.* **2006**, *128*, 9170−9177.
(4) Gorenstein, D. G.; Kar, D. *Biochem. Biophys. Res. Commun.* **1975**, *65*, 1073−1080.
(5) Přecechtělová, J.; Novák, P.; Munzarová, M. L.; Kaupp, M.; Sklenář, V. *J. Am. Chem. Soc.* **2010**, *132*, 17139−17148.
(6) Giessner-Prettre, C.; Pullman, B.; Prado, F. R.; Cheng, D. M.; Iuorno, V.; Ts'o, P. O. P. *Biopolymers* **1984**, *23*, 377−388.
(7) Benda, L.; Vokáčová, Z. S.; Straka, M.; Sychrovský, V. *J. Phys. Chem. B* **2012**, *116*, 3823−3833.
(8) Přecechtělová, J.; Munzarová, M. L.; Novák, P.; Sklenář, V. *J. Phys. Chem. B* **2007**, *111*, 2658−2667.
(9) Přecechtělová, J.; Padrta, P.; Munzarová, M. L.; Sklenář, V. *J. Phys. Chem. B* **2008**, *112*, 3470−3478.
(10) Dršata, T.; Pérez, A.; Oroczo, M.; Morozov, A. V.; Šponer, J.; Lankaš, F. *J. Chem. Theory Comput.* **2013**, *9*, 707−721.
(11) Heddi, B.; Foloppe, N.; Oguey, C.; Hartmann, B. *J. Mol. Biol.* **2008**, *382*, 956−970.
(12) Abi-Ghanem, J.; Heddi, B.; Foloppe, N.; Hartmann, B. *Nucleic Acids Res.* **2010**, *38*, 1119−1128.
(13) Pérez, A.; Marchán, I.; Svozil, D.; Šponer, J.; Cheatham, T. E., III; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817−3829.
(14) MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L. *J. Comput. Chem.* **2004**, *25*, 1400−1415.
(15) Holbrook, S. R.; Kim, S.-H. *J. Mol. Biol.* **1984**, *173*, 361−388.
(16) Fujimoto, B. S.; Willie, S. T.; Reid, B. R.; Schurr, J. M. *J. Magn. Reson. B* **1995**, *106*, 64−67.
(17) Borer, P. N.; Laplante, S. R.; Kumar, A.; Zanatta, N.; Martin, A.; Hakkinen, A.; Levy, G. C. *Biochemistry* **1994**, *33*, 2441−2450.
(18) Kojima, C.; Ono, A.; Kainosho, M.; James, T. L. *J. Magnet. Reson.* **1998**, *135*, 310−333.
(19) Norberg, J.; Nilsson, L. *J. Chem. Phys.* **1996**, *104*, 6052−6057.
(20) Young, M. A.; Ravishanker, G.; Beveridge, D. L. *Biophys. J.* **1997**, *73*, 2313−2336.
(21) Norberg, J.; Nilsson, L. *Biophys. J.* **2000**, *79*, 1537−1553.
(22) Schwieters, C. D.; Clore, G. M. *Biochemistry* **2007**, *46*, 1152−1166.
(23) Pérez, A.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2007**, *129*, 14739−14745.

(24) Case, D. A.; Pearlman, D. A.; Caldwell, J.W..; Cheatham III, T. E.; Wang, J.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. AMBER 7; University of California: San Francisco, 2002.

(25) Wu, Z.; Delaglio, F.; Tjandra, N.; Zhurkin, V. B.; Bax, A. J. Biomol. NMR 2003, 26, 297−315.

(26) Cheatham, T., III; Cieplak, P.; Kollman, P. A. J. Biomol. Struct. Dyn. 1999, 16, 845−862.

(27) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. J. Chem. Phys. 1984, 81, 3684−3690.

(28) Darden, T.; York, D.; Pedersen, L. J. Chem. Phys. 1993, 98, 10089−10092.

(29) Pérez, A.; Marchán, I.; Svozil, D.; Šponer, J.; Cheatham, T. E., III; Laughton, C. A.; Orozco, M. Biophys. J. 2007, 92, 3817−3829.

(30) Várnai, P.; Djuranovic, D.; Lavery, R.; Hartmann, B. Nucleic Acids Res. 2002, 30, 5398−5406.

(31) Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. AMBER 12; University of California: San Francisco, 2012.

(32) Car, R.; Parrinello, M. Phys. Rev. Lett. 1985, 55, 2471−2474.

(33) HyperChem 8.0.3; Hypercube Inc.: Gainesville, FL, 2007.

(34) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. J. Comput. Chem. 1986, 7, 230−252.

(35) Wang, J. M.; Cieplak, P.; Kollman, P. A. J. Comput. Chem. 2000, 21, 1049−1074.

(36) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. J. Chem. Phys. 1983, 79, 926−935.

(37) CPMD V3.11 Copyright IBM Corp. 1990−2006, Copyright MPI für Festkörperforschung Stuttgart 1997−2001.

(38) Becke, A. D. J. Chem. Phys. 1993, 98, 5648−5652.

(39) Lee, C.; Yang, W.; Parr, R. G. Phys. Rev. B 1988, 37, 785−789.

(40) Vanderbilt, D. Phys. Rev. B 1990, 41, 7892−7895.

(41) Hoover, W. G. Phys. Rev. A 1985, 31, 1695−1697.

(42) Frenkel, D.; Smit, B. Integration Schemes. Understanding Molecular Simulation, 2nd ed.; Computational Science Series Vol. 1; Academic Press: New York, 2002; pp 540−542.

(43) Schneider, B.; Patel, K.; Berman, H. M. Biophys. J. 1998, 75, 2422−2434.

(44) Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. Can. J. Chem. 1992, 79, 560−571.

(45) Chesnut, D. B.; Rusiloski, B. E.; Moore, K. D.; Egolf, D. A. J. Comput. Chem. 1993, 14, 1364−1375.

(46) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., 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 E.01; Gaussian, Inc.: Wallingford, CT, 2004.

(47) Becke, A. D. Phys. Rev. A 1988, 38, 3098−3100.

(48) Perdew, J. P. Phys. Rev. B 1986, 33, 8822−8824.

(49) Kutzelnigg, W.; Fleischer, U.; Schindler, M. The IGLO-Method: Ab-initio Calculation and Interpretation of NMR Chemical Shifts and Magnetic Susceptibilities. NMR-Basic Principles and Progress; Springer-Verlag: Berlin/Heidelberg, 1990; Vol. 23, pp 165−262.

(50) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. Chem. Phys. Lett. 1989, 162, 165−169.

(51) Treutler, O.; Ahlrichs, R. J. Chem. Phys. 1995, 102, 346−354.

(52) Von Arnim, M.; Ahlrichs, R. J. Comput. Chem. 1998, 19, 1746−1757.

(53) Kaupp, M.; Reviakine, R.; Malkina, O. L.; Arbuznikov, A.; Schimmelpfennig, B.; Malkin, V. G. J. Comput. Chem. 2002, 23, 794−803.

(54) Malkin, V. G.; Malkina, O. L.; Reviakine, R.; Arbuznikov, A. V.; Kaupp, M.; Schimmelpfennig, B.; Malkin, I.; Helgaker, T.; Ruud, K. MAG-ReSpect, Version 1.1; 2003.

(55) Edmiston, C.; Ruedenberg, K. Rev. Mod. Phys. 1963, 35, 457−&.

(56) Lavery, R.; Zakrzewska, K.; Beveridge, D.; Bishop, T. C.; Case, D. A.; Cheatham, T.; Dixit, S.; Jayaram, B.; Lankas, F.; Laughton, C.; Maddocks, J. H.; Michon, A.; Osman, R.; Orozco, M.; Perez, A.; Singh, T.; Spackova, N.; Sponer, J. Nucleic Acids Res. 2010, 38, 299−313.

(57) Beveridge, D. L.; Barreiro, G.; Byun, K. S.; Case, D. A.; Cheatham, T. E., III; Dixit, S. B.; Giudice, E.; Lankas, F.; Lavery, R.; Maddocks, J. H.; Osman, R.; Seibert, E.; Sklenar, H.; Stoll, G.; Thayer, K. M.; Varnai, P.; Young, M. A. Biophys. J. 2004, 87, 3799−3813.

(58) Sklenář, V.; Bax, A. J. Am. Chem. Soc. 1987, 109, 7525−7526.

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

(60) Case, D. A. NMR Parameters in Proteins and Nucleic Acids. In Calculation of NMR and EPR Parameters. Theory and Applications; Kaupp, M.; M. Bühl, V. G. Malkin, Eds.; Wiley: Weinheim, 2004; p 341.

(61) Dračínský, M.; Kamínský, J.; Bouř, P. J. Phys. Chem. B 2009, 113, 14698−14707.

(62) Cramer, C. J. J. Org. Chem. 1992, 57, 7034−7043.

(63) Wolfe, S. Acc. Chem. Res. 1972, 5, 102−111.

(64) (a) Newton, M. D. J. Am. Chem. Soc. 1973, 95, 256−258. (b) Perahia, D.; Pullman, B.; Saran, A. Biochim. Biophys. Acta 1974, 340, 299−313.

(65) Pullman, B.; Pullman, A.; Berthod, H.; Gresh, N. Theoret. Chim. Acta (Berl.) 1975, 40, 93−111.

(66) Saenger, W.; Structures and Conformational Properties of Bases, Furanose Sugars, and Phosphate Groups. Principles of Nucleic Acid Structure; Springer-Verlag: New York, 1984; p 96.

(67) Banavali, N. K.; MacKerell, A. D., Jr. J. Am. Chem. Soc. 2001, 123, 6747−6755.

(68) Kessler, J.; Dračínský, M.; Bouř, P. J. Comput. Chem. 2012, DOI: 10.1002/jcc.23143.