

PAPER • OPEN ACCESS

The tilt-dependent potential of mean force of a pair of DNA oligomers from all-atom molecular dynamics simulations

To cite this article: Ruggero Cortini *et al* 2017 *J. Phys.: Condens. Matter* **29** 084002

View the [article online](#) for updates and enhancements.

Related content

- [Theory of phase segregation in DNA assemblies containing two different base-pair sequence types](#)
Dominic J (O') Lee, Aaron Wynveen and Alexei A Kornyshev
- [Charge renormalization of helical macromolecules](#)
D J Lee
- [Statistical mechanical model for a closed loop plectoneme with weak helix specific forces](#)
Dominic J (O') Lee

The tilt-dependent potential of mean force of a pair of DNA oligomers from all-atom molecular dynamics simulations

Ruggero Cortini^{1,2,3,4,7}, Xiaolin Cheng^{5,6} and Jeremy C Smith^{5,6}

¹ Chemistry Department, Faculty of Natural Sciences, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

² Laboratoire de Physique Théorique de la Matière Condensée, CNRS UMR 7600, Université Pierre et Marie Curie, Sorbonne Université, 4 place Jussieu, 75252 Cedex 05, Paris, France

³ Genome Architecture, Gene Regulation, Stem Cells and Cancer Programme, Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Dr. Aiguader 88, Barcelona 08003, Spain

⁴ Universitat Pompeu Fabra (UPF), Barcelona, Spain

⁵ UT/ORNL Center for Molecular Biophysics, Oak Ridge National Laboratory, Oak Ridge, TN 37830, USA

⁶ Department of Biochemistry, Cellular & Molecular Biology, The University of Tennessee, Knoxville, TN 37996, USA

E-mail: smithjc@ornl.gov

Received 24 August 2016, revised 27 October 2016

Accepted for publication 17 November 2016

Published 16 January 2017



CrossMark

Abstract

Electrostatic interactions between DNA molecules have been extensively studied experimentally and theoretically, but several aspects (e.g. its role in determining the pitch of the cholesteric DNA phase) still remain unclear. Here, we performed large-scale all-atom molecular dynamics simulations in explicit water and 150 mM sodium chloride, to reconstruct the potential of mean force (PMF) of two DNA oligomers 24 base pairs long as a function of their interaxial angle and intermolecular distance. We find that the potential of mean force is dominated by total DNA charge, and not by the helical geometry of its charged groups. The theory of homogeneously charged cylinders fits well all our simulation data, and the fit yields the optimal value of the total compensated charge on DNA to $\approx 65\%$ of its total fixed charge (arising from the phosphorous atoms), close to the value expected from Manning's theory of ion condensation. The PMF calculated from our simulations does not show a significant dependence on the handedness of the angle between the two DNA molecules, or its size is on the order of $1k_B T$. Thermal noise for molecules of the studied length seems to mask the effect of detailed helical charge patterns of DNA. The fact that in monovalent salt the effective interaction between two DNA molecules is independent on the handedness of the tilt may suggest that alternative mechanisms are required to understand the cholesteric phase of DNA.

Keywords: DNA–DNA interactions, all-atom molecular dynamics simulations, Umbrella sampling, Kornyshev-Leikin theory

(Some figures may appear in colour only in the online journal)



Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

⁷ This article belongs to the [special issue Emerging Leaders](#) which features invited work from the best early-career researchers working within the scope of *Journal of Physics: Condensed Matter*. This project is part of the *Journal of Physics* series' 50th anniversary celebrations in 2017. Ruggero Cortini was selected by the Editorial Board of *Journal of Physics: Condensed Matter* as an Emerging Leader.

1. Introduction

The interaction between double-stranded B-DNA molecules has been extensively studied experimentally and theoretically [1–3] because of its importance in biology [4] and in DNA nanotechnology [5]. Despite significant progress in our understanding of the principles of these interactions, many aspects still remain unclear and debated.

The chirality of DNA molecules is expected to give rise to an asymmetry in interactions that depends on the handedness of the interaxial angle between two DNA molecules. Such asymmetry has been experimentally seen in DNA liquid crystals. Indeed, in a wide range of conditions DNA molecules form a cholesteric phase with a giant cholesteric pitch (of the order of $\sim 1\text{--}10\ \mu\text{m}$) [6, 7]. The theoretical analysis of the cholesteric phase has proven challenging. If the two main components of the interactions in the cholesteric phase are steric effects (volume exclusion) and electrostatics, changes in the balance between the two effects will give rise to different behaviours. Whereas electrostatic interactions alone should give rise to a right-handed helical phase [8, 9], the interplay between steric and electrostatic effects can give rise to both handednesses [10–12]. Experiments show that the handedness of the cholesteric phase can be either left- or right-handed [13, 14]. The role of electrostatic interactions is still difficult to assess, given the complexity of the system. Tilt-dependent electrostatic interactions between DNA molecules are important in biology and biophysics, as they may give rise to non-trivial effects in the behaviour of supercoiled DNA plasmids [15–17].

The Kornyshev–Leikin (KL) theory of electrostatic DNA–DNA interactions explicitly takes the helical pattern of the charged groups on the DNA surface into account [18]. The theory has highlighted the importance of helical components in the electrostatic interactions, which have been observed in osmotic stress experiments [19] and x-ray DNA scattering profiles [20]. However, the importance of helix-dependent effects in the cholesteric phase and in DNA supercoiling remains to be elucidated.

The KL theory predicts that the electrostatic interaction between DNA molecules depends on the helical structure of the charge patterns in DNA. Negative charges on the phosphate groups and positive ions at the DNA surface are expected to give rise to components of the interactions that are specific to this helical structure. Depending on the amount of positive ions adsorbed on the DNA molecules, and their location, different behaviours are expected:

- (i) If the counterions are adsorbed in the grooves [21] they will form an additional helical pattern of positive charges, which follows the helices of negative charges on the phosphates. The electric field of such charge distribution is screened by the electrolyte and decays with a characteristic length of $((2\pi/H)^2 + \lambda_D^{-2})^{-1/2}$, where H is the DNA helical pitch ($\approx 34\ \text{\AA}$) and λ_D is the Debye screening length ($\approx 7\text{--}10\ \text{\AA}$ in physiological ionic conditions). Although the screening length is short, the effect of the helical structure of the charges is important, and scales with the length of the interacting DNA molecules [22].
- (ii) Important helix-specific electrostatic effects are expected when the ions are not adsorbed on the DNA surface, but are homogeneously distributed around it.
- (iii) If the counterions are specifically adsorbed by the phosphates, the effect of the interaction reflecting the

helical pattern of charges is expected to be suppressed; depending on the amount of adsorption charges may be fully quenched by counterions. In monovalent salt solutions the theory of counterion condensation by Manning predicts that about 70% of the total DNA charge is compensated by positive ions in a close layer surrounding the DNA molecules [23]. If these ions are tightly bound to the phosphate groups, it might be difficult to observe helix-specific electrostatic effects, as they might be overshadowed by thermal noise, as predicted by the KL theory [18].

It is important to stress that in the cases in which helix-specific electrostatic effects matter (cases (i) and (ii) above) the total amount of positive charge adsorbed on the DNA molecules is the most important parameter describing the behaviour of the system. If this is smaller than 70–80% of the total DNA charge, the dominant mode of interaction will be the repulsion of net charge screened by the electrolyte, equivalent to the interaction of homogeneously charged rods [18] (the so-called zero helical harmonic interaction). However, even in this case the helical components of the interaction are important, and their effects might be observed when considering DNA molecules at different azimuthal orientations, or at interaxial skew angles of different handednesses.

Ever since it was proposed, it has been questioned whether the KL theory could describe DNA–DNA interactions in ordinary solutions of monovalent salt, or whether its applicability is limited to cases in which positive ions are specifically adsorbed in the DNA grooves (such as spermine, spermidine, cobalt-hexamine, etc). In the simulations of [24, 25], no helix-specific effects were found within a model of the electrolytic solution in which the ions were not assigned specific interactions that may favour their preferential adsorption into the grooves. However, in various experimental setups it has been observed that helix-specific electrostatic effects might also be relevant in the case of monovalent salt solutions [26–28]. For a review, see [3, 29]. It is thus important to further study this problem at the level of all-atom computer simulations to understand what one might expect from such systems in solutions of simple monovalent salt.

All-atom molecular dynamics (MD) simulations have proven useful in studying DNA–DNA interactions [24, 25, 30–32]. Recent refinements of force fields and the expansion of computational power allow improved precision in the determination of the details of interactions between nucleic acids and their ionic cloud [33–35]. For a review, see [36]. So far, only one study has been reported, aimed at elucidating the effect of the tilt angle on DNA–DNA interactions, mainly in the context of understanding the importance of direct binding of DNAs through magnesium ions [37].

Here, we employ large-scale MD simulations to construct the potential of mean force (PMF) of two 24 bp DNA oligomers as a function of their interaxial angle. We investigate, in particular, the importance of the helical components of the electrostatic interaction.

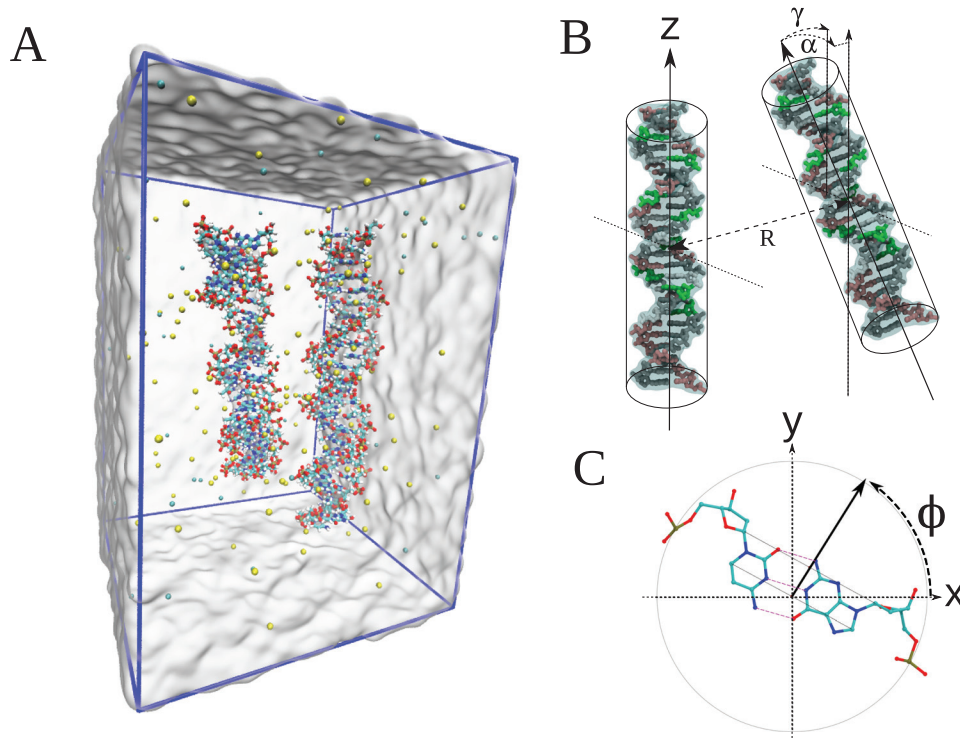


Figure 1. The geometry of the MD system. (A) A snapshot of the simulation box. Ions are shown as spheres, and the blue lines delimit the periodic simulation box. (B) Geometrical parameters. Two identical DNA oligomers are displaced by R along their centre axes, and tilted by an angle α away from their parallel orientation, which is defined positively for right-handed DNA–DNA crossovers. The angle γ is the angle away from the parallel plane. (C) is the definition of the ϕ angle for a single base pair. The centre of the minor groove is defined according to the guidelines of [39].

2. Methods

2.1. MD simulation protocol

The structure of a 24 bp-long B-DNA molecule (sequence GGTGACGAGTGAGCTACTGGGCGG) was generated using 3DNA software [38]. The DNA structure was then replicated, translated away from the centre of mass of the first molecule by a distance R , tilted by an interaxial angle α , and rotated by an angle $\Delta\Phi$ around its long axis (see figures 1(B) and (C)).

Water was added to the system by using the Solvate VMD (visual molecular dynamics) plugin [40], which adds TIP3P model water molecules to an orthorhombic cell, whose edges were calculated supposing that the distance between the edges of the periodic box and the DNA molecules must be at least 15 Å. Sodium chloride was then added to the system to reach a bulk ionic concentration of 0.15 M, which is close to the physiologically relevant value. Ions were placed, at random, in the initial configuration using the ‘autoionize’ plugin of VMD.

The CHARMM27 nucleic acid force field was used [41, 42], as implemented in the NAMD (nanoscale molecular dynamics) program [43]. NBFIX (non-bonded fix) corrections were applied to the Lennard-Jones interaction potentials between sodium atoms and the phosphate oxygens, as suggested in recent works [33, 34]: these were shown to be important in predicting the correct DNA–DNA interaction patterns in fibres of DNA under osmotic stress.

Energy minimization was carried out for 8000 steps, followed by MD simulation. A time step of 2 fs was used, constraining all bonds containing hydrogen atoms, using the SETTLE algorithm to reset their positions and velocities [44]. For the non-bonded interactions, a switching distance was set as 7 Å, the cutoff distance set to 8 Å, and the pair-list regeneration distance to 10 Å. The electrostatic interactions were evaluated using the particle mesh Ewald method [45], using a grid spacing equal to the box-edge size rounded to the closest multiple of 3. All non-bonded interactions, except electrostatic interactions, were recalculated fully every step, and electrostatic interactions were recalculated fully every two steps.

The equilibration of the system was carried out in two successive phases. In the first phase, the initial temperature of the system was set to 200 K, and was progressively heated to 350 K in the NVT ensemble, in steps of 1 K every 660 ps. During this phase, all DNA atoms were restrained around their initial positions using a harmonic potential with $k = 10$ kcal/mol/Å². The temperature of the system was fixed using Langevin dynamics, with a damping coefficient of 10 ps⁻¹. High-temperature equilibration (at 350 K) was performed for 1 ns, and then a cooling cycle applied of 1 K every 1220 ps, decreasing the temperature to the target of 298 K.

Before the second equilibration cycle, the collective variable calculation module of the NAMD was turned on. We enforced four different collective variable constraints, using the ‘colvars’ module of the NAMD, each with a characteristic harmonic spring constant k . These are: (a) R : the interaxial distance between the two DNA oligomers ($k = 10$

kcal/mol/Å²); (b) α , the tilt angle ($k = 5$ kcal/mol/deg²); (c) γ , the inward/outward tilt angle away from the xy plane ($k = 10$ kcal/mol/deg²); (d) $\Delta\Phi$: the azimuthal angle difference between the two DNA oligomers ($k = 10$ kcal/mol/deg²). See figure 1 for the definitions of these geometrical quantities.

The constraint on R is easily set by using the centre of mass of the two oligomers. The other constraints were all enforced by using the ‘spinAngle’ option in the colvars module. One defines reference coordinates for the atoms, and then the program calculates the rotation that brings back the molecule to those reference coordinates. Once a reference angle is set, a harmonic force is applied to the system. We define the reference coordinates as follows. For α , they are the position of DNA 1 when $\alpha = 0$. For $\Delta\Phi$ and for γ , they are the positions of DNA 2.

During the second equilibration cycle, simulation was carried out in the NPT ensemble. The pressure was controlled using Nosé–Hoover dynamics [46, 47], with the target pressure 1.01325 bar, a piston period of 200 fs, and a damping timescale of 100 fs. This equilibration run was 2 ns long. In this phase, and in the successive runs, harmonic constraints were applied only on one of the DNA molecules, in order to be able to sample the tilt motion of one of the DNA and construct the PMF. Furthermore, we applied harmonic restraints to the hydrogen bonds of the terminal base pairs, with a harmonic constant of 2 kcal mol⁻¹. These restraints were applied to prevent potential base-pair fraying at the DNA termini, thus improving the structural integrity of the DNA during the simulations. Using this same configuration, 8 ns of production MD run for umbrella sampling were performed. During this final phase, snapshots of the system were taken every 2000 time steps, and used for statistical analysis.

2.2. Geometrical calculations

To provide insight into the dynamics of the DNA and the ions during the simulation, we calculated several geometrical quantities for each trajectory frame.

The first important quantity is the amount of charge compensation θ , on the surface of the DNA molecules. This is defined as the fraction of positive charge which is bound *at the surface* of a DNA molecule, divided by the total DNA charge (which is $-2e/\text{bp}$). This quantity is not rigorously defined, as it depends on the distance d at which one considers a positive ion to be bound to the DNA surface. Therefore, we chose d to match the expected value of θ at large interaxial separation, which can be estimated using the Manning theory of charge condensation in the corresponding ionic conditions [23]. We determined all sodium ions within a distance d away from the surface of DNA 1 and DNA 2, their numbers denoted $n^{(1)}$ and $n^{(2)}$, and computed the fraction of charge compensation as follows:

$$\theta^{(\nu)} = \frac{n^{(\nu)}}{2N} \quad \nu = 1, 2, \quad (1)$$

where $N = 24$ is the number of base pairs in the DNA molecules.

Next, it is interesting to determine which of the ions are bound to the minor or major grooves of the DNA molecules. To do so, we first calculated the orientation of the DNA grooves at each step. This calculation enabled us to determine the *relative azimuthal orientation* of the DNA minor grooves, $\Delta\Phi$, which is defined in figure 1(C).

The first step necessary for this is to calculate the points that define the DNA centre lines. To do so, we followed the prescriptions of [39] and defined the centre line as the line connecting the mid-points of the segment that, for each base pair, connects the purine C8 and pyrimidine C6 atoms. We then computed the average vector connecting these points, which defines the vectors $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(2)}$, the centre-line directions of DNA 1 and DNA 2, respectively. Furthermore, the average spacing between the centre-line points defines the average base-pair height, l_h .

Once the centre-line points are defined, we next need to define the vector that points to the centre of the minor groove at each base-pair step. Again, in [39] we find a well-posed definition. Starting from the segment connecting the C1 atoms for each base pair, the midpoint of this is then connected to the above point that defines the DNA centre line at the same base pair (see figure 1(C)). This segment is the $\phi_i^{(\nu)}$ vector, which defines the minor groove of the ν th DNA molecule at base pair i .

In the KL theory the $\Delta\Phi$ angle is defined as the average orientation of the DNA minor grooves for *parallel* DNA molecules. Since in our simulations the DNA molecules are skewed, we must first apply a geometrical transformation to the $\phi_i^{(\nu)}$ vectors. Since we already know the vectors $\mathbf{a}_{1,2}$, we calculate the matrices $M_z^{(1)}$ and $M_z^{(2)}$, which realign such vectors to the z -axis. Then, we may calculate

$$\Delta\Phi = \frac{1}{N} \sum_{i=1}^N \arccos [(\mathbf{M}_z^{(1)} \phi_i^{(1)}) \cdot (\mathbf{M}_z^{(2)} \phi_i^{(2)})], \quad (2)$$

which is the average of the realigned orientations of the minor grooves.

Once the $\phi_i^{(\nu)}$ vectors and the $M_z^{(\nu)}$ matrices are known, we may determine which ions are bound to the DNA minor or major grooves. We define $\mathbf{b}_j^{(\nu)}$ as the position vector of the j th ion to DNA ν ($j = 1, \dots, n^{(\nu)}$). To simplify the calculation, we translate and rotate the $\phi_i^{(\nu)}$ and $\mathbf{b}_j^{(\nu)}$ vectors. First, we translated all vectors so that the first base pair was on the origin. We then applied the $M_z^{(\nu)}$ matrix to each of these vectors. We then assigned the ion to one of the DNA base pairs by simply evaluating the altitude of the ion and comparing it to the altitude of each base pair. Once this was known, we calculated the angle between the minor groove of this base pair and the azimuthal angle of the ion. We then assigned the ion to the minor or major groove depending on whether this angle was smaller than 35° or larger than 145°, respectively.

Finally, we calculated the residence time of the sodium ions at the surface of the DNA molecules. At each time frame of the production runs, ions less than 6 Å away from the DNA molecules were considered to be bound. Given the list of all

bound ions as a function of time, the residence time τ can be calculated as the number of successive time frames for which a given ion is bound. This way, we can assess the residence-time distributions for all simulation rounds.

All analyses were carried out using the MOLDYN computer cluster at ORNL.

2.3. Potential of mean force calculation

We performed two independent sets of MD simulations. In the first set, we kept the intermolecular distance R at a fixed value of 26 Å, and we performed 83 independent simulations (one for each value of the interaxial angle α) in the range $(-30^\circ, 30^\circ)$, with an increment of 0.75° . At the value of $R = 26$ Å the closest distance between the two DNA oligomers is 4–5 Å, shorter than the Debye screening length, and we expect that the electrostatic effects will be strong and that we might be able to detect significant helix-dependent effects. We performed these simulations for three independent values of $\Delta\Phi = 0, \pi/2, \pi$. Each simulation was run for a total of 10 ns, out of which the last 8 ns were used to perform statistical analyses, as detailed in section 2.1.

In the second set of simulations we kept $\Delta\Phi$ fixed at zero, and performed 29 independent simulations (one for each value of the interaxial distance R) in the range (26 Å, 40 Å), with an increment of 0.5 Å. We did this for three independent values of $\alpha = -30^\circ, 0^\circ, 30^\circ$. The simulation length was the same as above.

We then constructed the PMF for each simulation set by using the weighted histogram analysis method [48, 49], as implemented in the WHAM package [50]. Errors in the estimate of the PMF were obtained using bootstrap sampling [51].

The PMF may be compared with the results of the KL theory. For DNA oligomers, the tilt-dependent electrostatic energy was calculated within the framework of the KL theory in [9]:

$$F(R, \alpha, \Delta\Phi) \approx F_0(R, \Delta\Phi) + F_1(R, \Delta\Phi)\alpha + \frac{1}{2}F_2(R, \Delta\Phi)\alpha^2 \quad (3)$$

Formulas A.2–A.4 of [9] provide expressions for the coefficients F_0 , F_1 and F_2 . These expressions should be valid in the limit in which the interaxial angle satisfies $\alpha \ll \lambda_D/L$. In our case, this means that the maximum/minimum angle should be $\approx 10^\circ$. Note that the above expression has been calculated using $R \rightarrow \infty$ as the reference point for zero energy.

In the KL theory, three unknown parameters appear in the formulas for F_0 , F_1 and F_2 , which are denoted by θ , f_1 and f_2 . The parameter θ is the percentage of the (fixed) negative charge on the DNA (due to the charge on the phosphate groups) that is compensated by adsorbed or condensed positive ions at the DNA surfaces. The percentage of these positive ions that are adsorbed or condensed in the minor or major grooves is f_1 and f_2 , respectively.

In the case of fixed R and varying α , we should calculate the difference

$$\Delta F_\alpha = F(\alpha) - F(\alpha_0), \quad (4)$$

where α_0 is an arbitrarily chosen angle. It is then natural to choose $\alpha_0 = \alpha_{\max}$. Equation (4) simplifies to

$$\Delta F_\alpha = F_1(\alpha - \alpha_0) + \frac{1}{2}F_2(\alpha^2 - \alpha_0^2). \quad (5)$$

This expression highlights that the PMF in this case is independent of F_0 .

In the case of varying R and fixed α , we should instead fit the expression

$$\Delta F_R = F_0 + \alpha F_1 + \frac{1}{2}F_2\alpha^2, \quad (6)$$

where we have taken into account that F_0 , F_1 and F_2 tend to zero for large R , and we have chosen $R \rightarrow \infty$ as our reference point for the free energy.

To gain insight into the importance of the helical components of the electrostatic interactions, we also fitted a purely non-chiral theory: the theory of homogeneously charged cylinders (HCC). The expressions for the HCC theory may be obtained directly from the KL theory by taking into account only the zero-order term in the infinite sums. These zero-order harmonics correspond to the case of a homogeneously charged cylinder. These terms clearly depend only on the θ parameter, but not on the f_1 and f_2 parameters.

To compare the MD results to the KL and HCC theories, we fitted the expressions of the KL and HCC theories to the corresponding values of the PMF obtained from MD simulations. For the KL theory, we fitted the three fit parameters θ , f_1 and f_2 , and for the HCC theory we fitted only the θ parameter. We used an adapted version of the Levenberg–Marquardt (LM) algorithm [52] as implemented in the GNU Scientific Library [53] to perform the fit.

3. Results

Here, we illustrate the results from our all-atom simulations of a pair of 24 base-pair B-DNA oligomers. We performed two sets of independent simulations: (a) at fixed interaxial distance, for varying interaxial tilt, at three fixed values of the azimuthal angle and (b) at fixed azimuthal angle, for varying interaxial distance, at three fixed values of the interaxial tilt.

From the results of the two simulation sets we calculated the PMF as a function of the varying coordinate (α in the first case, R in the second). The data obtained were then fitted by two different theories of DNA–DNA interactions: the KL theory, and the HCC theory (see section 2.3 for details on the fitting procedure).

Ideally, we could have used this simulation protocol to calculate the full two-dimensional PMF in the $R - \alpha$ space. However, such calculations would be very computationally demanding, and are beyond the scope of the current study. Instead, we decided to focus on a small cross-sectional portion of the entire space.

3.1. Fixed tilt, variable distance

Figure 2 shows the PMF as a function of the interaxial distance R , for fixed $\Delta\Phi = 0$, and three different values of $\alpha = -30^\circ, 0^\circ, 30^\circ$.

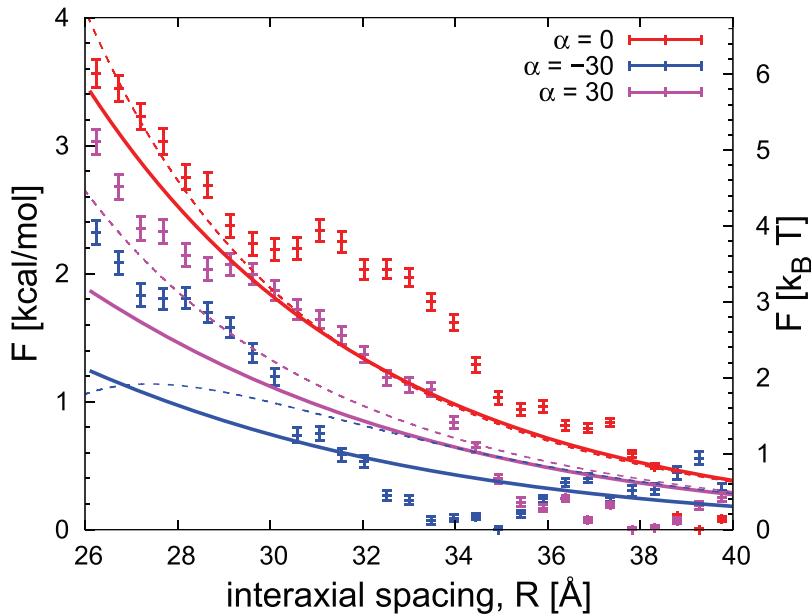


Figure 2. PMF as a function of R , calculated using the umbrella sampling method. Different curves are shown for the values of α that we studied. Continuous and dashed lines show, respectively, the fit results to the KL theory and the KL theory calculated for geometrically estimated parameters (see table 1), for each fixed value of α . The colours each correspond to a fixed value of α . Error bars correspond to estimated standard deviations using the bootstrap method (see section 2.3).

Table 1. DNA charge parameters (see section 2 for definitions) as obtained by LM least-squares fitting to the KL theory, HCC theory, and from geometrical calculations, for the simulations at fixed α and variable R .

α	KL fit				HCC fit		Geometrical		
	θ	f_1	f_2	$\bar{\chi}^2$	θ	$\bar{\chi}^2$	θ	f_1	f_2
-30°	0.602	0.000	0.222	1.833	0.645	1.516	0.554	0.356	0.168
0°	0.601	0.787	0.267	2.138	0.564	2.101	0.568	0.361	0.165
30°	0.662	1.000	0.612	1.456	0.565	1.932	0.543	0.358	0.169

In figure 2 we superpose the curves obtained by fitting the KL theory onto the estimates of the PMF as a function of R (solid lines). A great advantage of using all-atom MD simulations to analyze this system is that we have access to the positions of all the positively charged ions, so that we can estimate the values of the KL parameters θ , f_1 and f_2 with geometrical calculations from the MD, as detailed in section 2.2. The results of the fit and of the geometrical calculations are summarized in table 1. The values of the estimates for θ , the total charge compensation, are roughly consistent across the three different methods presented, and range within the physically reasonable values of 55–70%. However, the values of f_1 and f_2 , the proportion of ions in the minor and major grooves, are very different when calculated from the fit and directly from the MD geometrical calculations. The reason for this discrepancy may lie in the fact that the KL theory inadequately describes the electrostatic interactions in monovalent salt, so that the fitted values of f_1 and f_2 do not really correspond to their actual values, as calculated from the positions of the ions. However, the θ parameter is in good agreement because it does not require the finer details of the interaction profiles, whereas f_1 and f_2 do. More details for the reasons behind this inadequacy will be given in section 4.

Figure 2 shows what the KL theory gives when using the values of θ , f_1 and f_2 calculated using the geometrical calculations (dashed lines). The theoretical results and the simulation data are in good qualitative agreement, but the MD results show deviations from the smooth theoretical curves. This result will be discussed at length in section 4. We note that the fits of the HCC theory are similar in quality to the fit using the KL theory (see the $\bar{\chi}^2$ values in table 1).

3.2. Fixed distance, variable tilt

Figure 3 shows the PMF as a function of the interaxial DNA–DNA tilt angle α , at fixed interaxial distance $R = 26 \text{ \AA}$, obtained from the MD simulations (see figure 1 for the definitions of the geometrical quantities). We also show the curves obtained by performing least-squares fits to the data using the KL theory (solid lines), and the prediction of the KL theory when using the values of the parameters θ , f_1 and f_2 estimated from geometrical calculations (dashed lines).

The values of the PMF that we obtain from the MD simulations have a maximum at $\alpha = 0$. This result is readily understood in terms of the electrostatic repulsion between the two DNA oligomers. In fact, since in monovalent salt the interaction is repulsive, the two oligomers decrease their repulsion

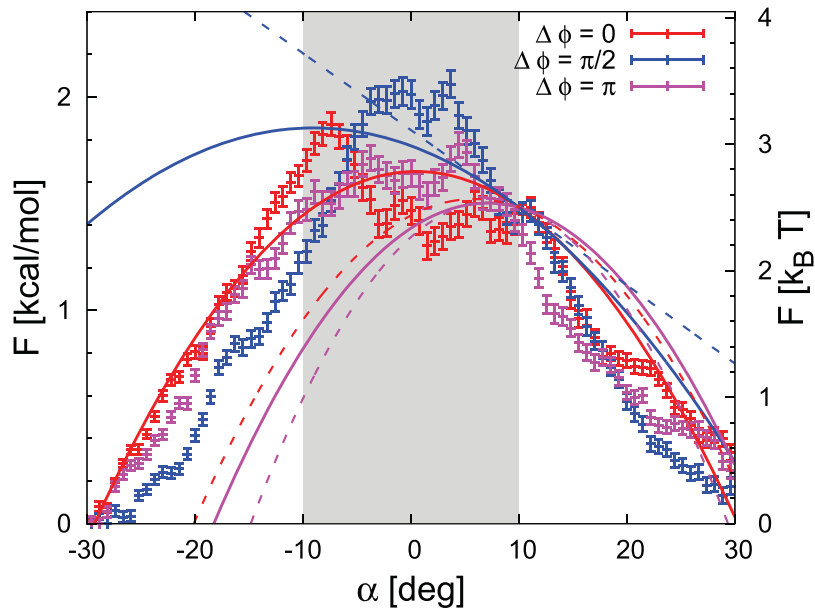


Figure 3. PMF as a function of α , calculated using the umbrella sampling method. Different curves are shown for the values of $\Delta\Phi$ that we studied. The shaded area corresponds to the values of α that satisfy the small-angle condition necessary for fitting to the KL formulas. Error bars correspond to estimated standard deviations using the bootstrap method (see section 2.3). Solid and dashed lines show, respectively, the fit to the KL theory and the prediction of the KL theory based on the parameters calculated from the positions of the ions, for each value of $\Delta\Phi$. The fit to the HCC theory lies very close to the KL fit at $\Delta\Phi = 0$, and is not shown.

Table 2. DNA charge parameters (see section 2 for definitions) as obtained by LM least-squares fitting to the KL theory, HCC theory, and geometrical calculations, for simulations at fixed $\Delta\Phi$, and variable α .

$\Delta\Phi$	KL fit				HCC fit		Geometrical		
	θ	f_1	f_2	$\bar{\chi}^2$	θ	$\bar{\chi}^2$	θ	f_1	f_2
0	0.525	0.009	0.213	0.439	0.545	0.445	0.620	0.337	0.162
$\pi/2$	0.569	0.000	0.000	3.023	0.497	0.742	0.618	0.348	0.164
π	0.665	0.140	0.000	2.771	0.527	0.527	0.616	0.340	0.160

the more they tilt, i.e. the farther apart they are. It is expected in this case that the minimum will be obtained at $\alpha = \pi/2$, that is, when the two oligomers are perpendicular to each other.

The curves for the three different $\Delta\Phi$ values studied here are similar. Table 2 shows the results of fitting the PMF curves for the three different values of $\Delta\Phi$ that we studied, along with the corresponding reduced χ^2 values. We also show the results of calculating the KL parameters θ , f_1 and f_2 from the positions of the ions, as explained in section 2. As in the case of fixed α and variable R , the θ values are consistent across the methods, but the f_1 and f_2 values are not.

The fit of the KL theory to the MD data gives good results for $\Delta\Phi = 0$. The fit to the HCC theory gives a prediction which lies very close to the KL theory at this value of $\Delta\Phi$. For $\Delta\Phi = \pi/2, \pi$ the fit to the KL theory is off from the data by as much as $\sim 2k_B T$.

To summarize, the fits of the HCC and KL theories yield similar results in both simulation sets, meaning that the simulation results do not provide evidence for an effect on the PMFs of the components of the electrostatic interactions due to the helical structure of the charges on the DNA. To

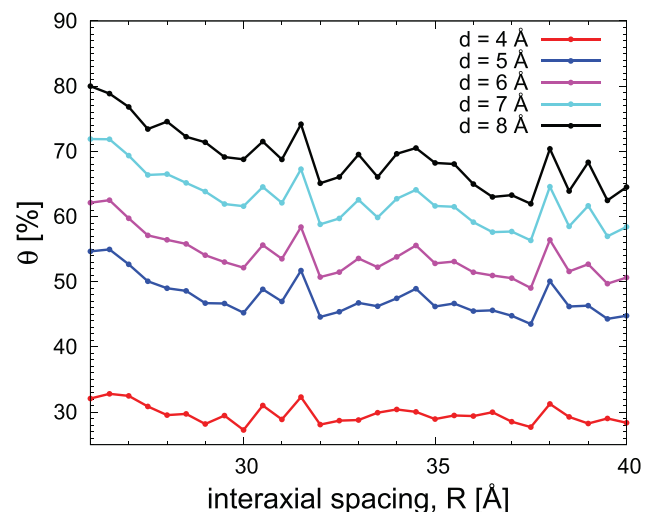


Figure 4. The value of the charge compensation θ from geometrical calculations is shown as a function of R , for different values of the distance d , chosen to discriminate between bound and non-bound ions.

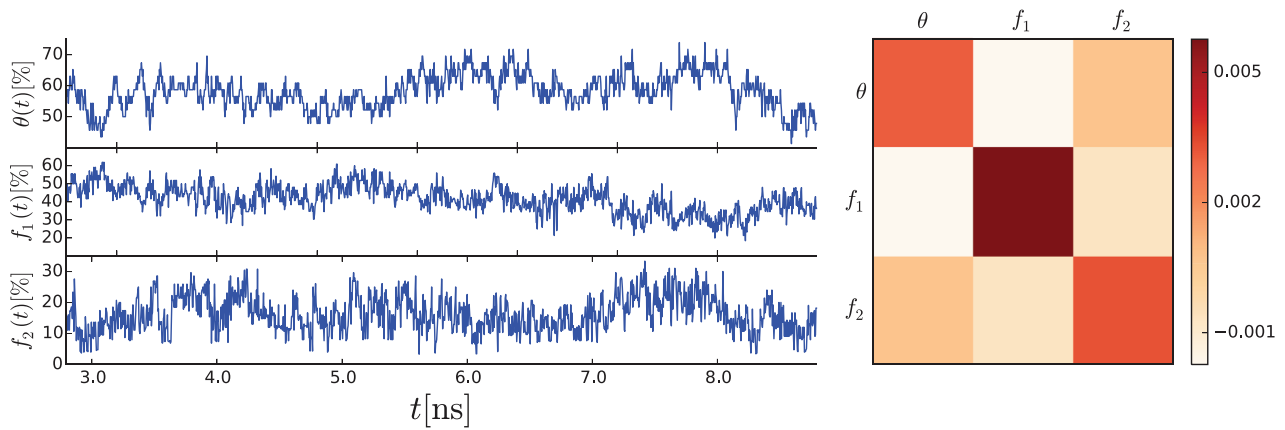


Figure 5. Variation of θ , f_1 , and f_2 as a function of time for our simulation run at $\alpha = -30^\circ$ and $R = 31.0 \text{ \AA}$. The values of the θ , f_1 and f_2 parameters were calculated as explained in section 2.2. The corresponding covariance matrix for this time series is depicted on the right.

understand this result in atomistic terms, we examined the distribution of the ions around the DNA molecules.

3.3. Ion distribution around DNA

In this section we examine the results of our simulations from the point of view of the positions of the ions around the DNA molecules. We will focus on two aspects: the geometrical positions of the ions with respect to the DNA grooves, as detailed in section 2.2, and the pair correlation functions, $g(r)$ between the ions and several components of the DNA molecule.

First, we start by examining the values of θ , f_1 and f_2 (the total charge compensation, the fraction of ions bound to the minor and major grooves, respectively), at fixed α and varying R . This case is particularly interesting because it provides a way to choose the value of d (which is the distance at which we consider an ion to be ‘bound’ to the DNA molecule: see section 2.2) to calculate the values of the KL parameters. In fact, at the largest DNA–DNA separation probed, $R = 40 \text{ \AA}$, it is expected that the value of θ should approach the Manning value of $\sim 60\text{--}70\%$, expected in our ionic conditions. We show the results of our calculations for several values of d in figure 4. With this method, we establish that $d \approx 6\text{--}7 \text{ \AA}$. Figure 4 shows that the value of θ seems to slightly increase when the interaxial separation R decreases. However, at $d = 4 \text{ \AA}$ there is almost no variation. This means that the increase in θ at small R is due to the fact that at $d > 4\text{--}5 \text{ \AA}$ the ion clouds of the two DNA molecules overlap. In such cases the same ion might be considered to be bound to both DNA molecules.

It is instructive to examine how the values of the θ , f_1 and f_2 are correlated among themselves. Figure 5 shows an example of the variation of these parameters as a function of the simulation time, and the corresponding covariance matrix. There is a non-zero positive covariance between θ and f_2 , whereas the covariance between θ and f_1 is close to zero. This suggests that the temporal variations of the total charge adsorbed on the DNA molecules is mostly due to the ions adsorbing to/desorbing from the DNA major grooves.

Next, for the chosen value of d , we examine the variation of θ , f_1 and f_2 at fixed α and varying R . Figure 6 shows the results

of our geometrical calculations to estimate these parameters. Figure 6(A) shows that there is a slight increase of θ when R decreases. This increase has a clear physical interpretation: since the interaction between the negatively charged helices is repulsive, to minimize the repulsive force there is a benefit in accumulating more positive charge on the molecules, at the expense of the demixing entropy. In panels (B) and (C) of the same figure it is possible to see that the proportion of ions in the major and minor grooves of the DNA does not change significantly with varying R .

We now examine the variation of the KL parameters calculated as before, at fixed R and varying α (figure 7). The value of θ changes slightly with the angle α , by up to around 5%, from $\sim 60\%$ to $\sim 65\%$, and is independent of $\Delta\Phi$ within the accuracy of our method. The values of f_1 and f_2 are independent of α and $\Delta\Phi$.

To better understand the increase in θ observed at $\alpha = 0$ and small values of R , we examine the pair correlation function $g(r)$ between the DNA phosphate atoms and the positive sodium ions in solution. Figure 8 shows this function for three values of R : the shortest ($R = 26 \text{ \AA}$), the longest ($R = 40 \text{ \AA}$) and an intermediate one ($R = 35 \text{ \AA}$). Overall, the behaviour of $g(r)$ is similar for the three values of R shown. A slight difference is observed for the peak at short distance, $\sim r = 3\text{--}4 \text{ \AA}$. The reason for the increase in θ with R may be due to the binding of sodium ions at close distance to the phosphorous atoms. To quantify this effect, we calculated the fraction of ions located within 5 \AA of a phosphorous atom (θ_p), as a function of the interaxial distance (figure 9). We observe an increase of the order of 2–3% at small R , which gives an estimate of the contribution to the total θ due to the binding of sodium ions close to the phosphorous atoms.

The increase in sodium ions bound to the phosphorous atoms alone is insufficient to provide an explanation for the results of our PMF data and fits. Positive ions tightly bound to the phosphate groups will neutralize the major source of the helical components to the electrostatic interaction, which is due to the fixed negative charges that follow a helical pattern. The increasing peak of the $g(r)$ function at short distance indicates that the closer the DNA molecules are, the stronger this effect is. However, an increase of only 2–3% is not enough to

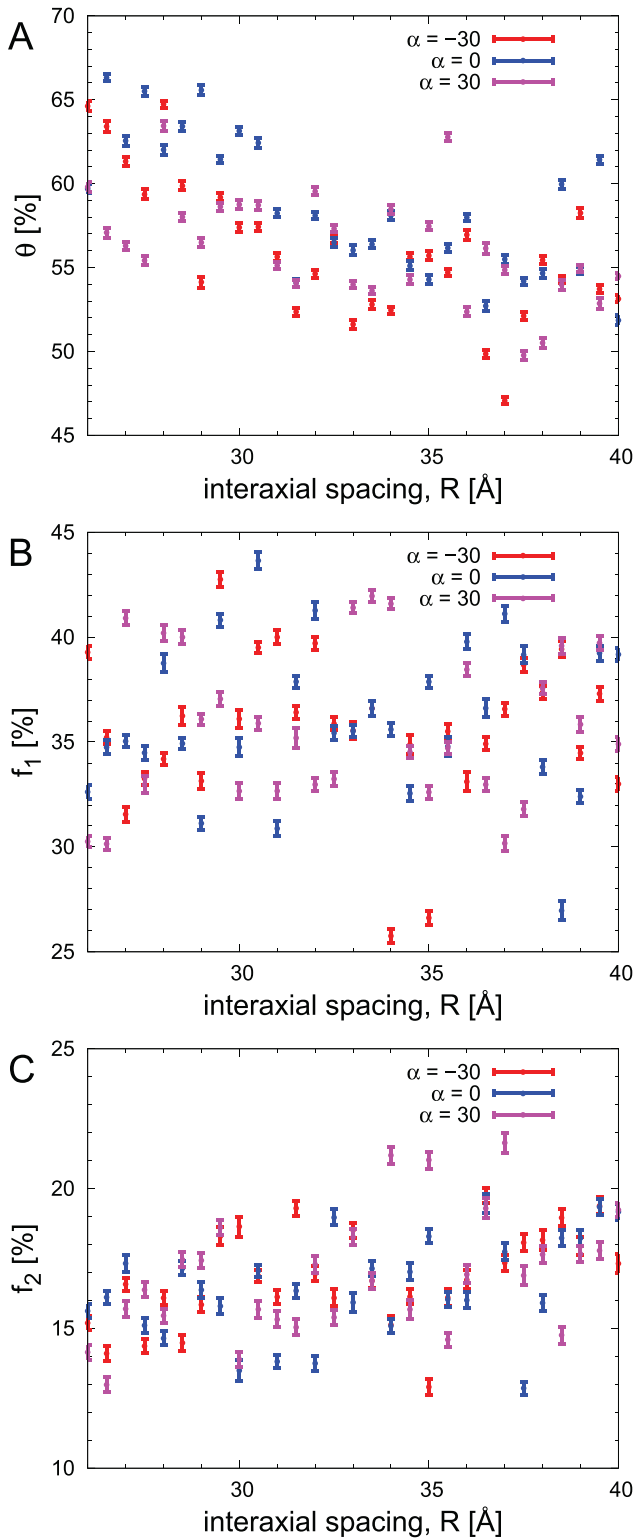


Figure 6. Variation of θ (fraction of neutralized negative charge on the DNA molecules), f_1 (its percentage in the minor groove) and f_2 (its percentage in the major groove) as a function of the interaxial separation R ((A), (B), and (C) respectively). Error bars indicate standard deviations on the mean.

explain the large difference between the prediction of the KL theory and the observed MD simulation results [18].

We calculated the timescale of sodium ions binding to the DNA molecule at a given base-pair step, and found it to

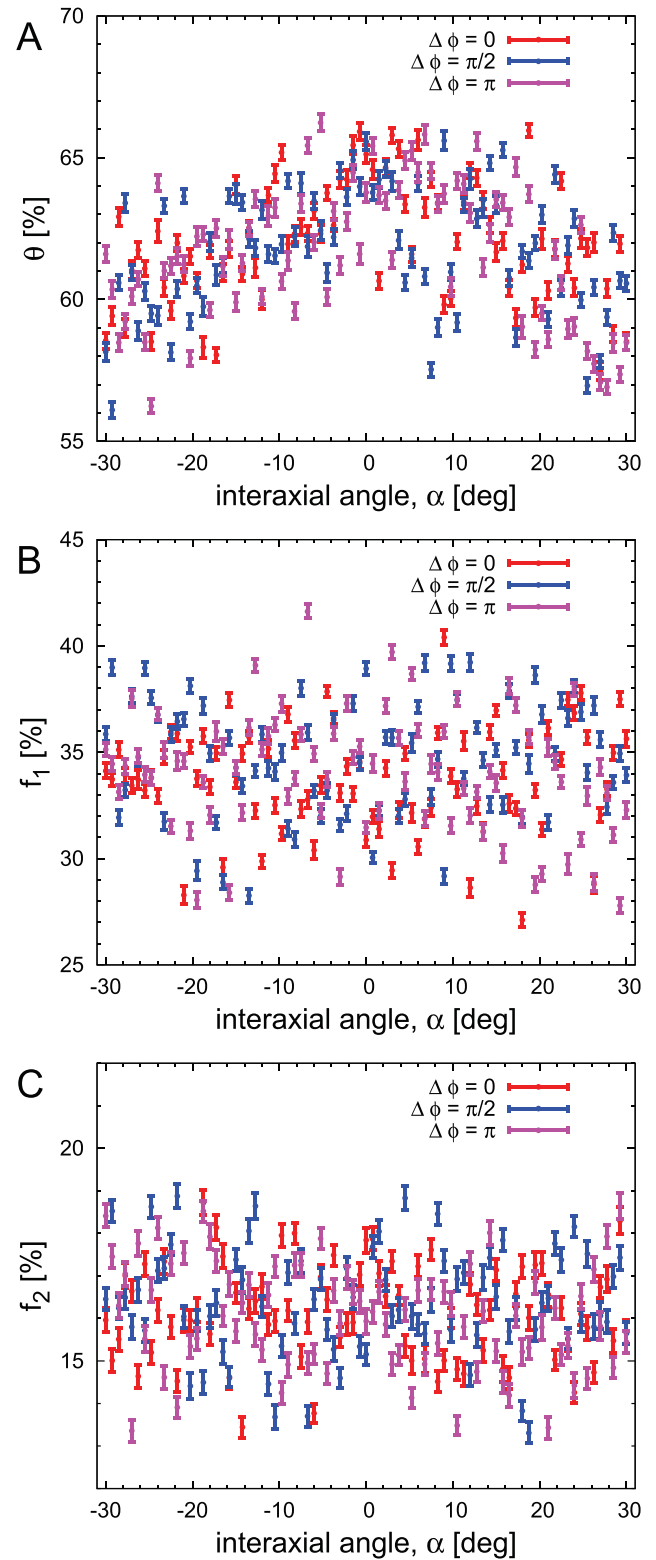


Figure 7. Variation of θ , f_1 and f_2 as a function of α , the interaxial tilt between the DNA molecules ((A)–(C), respectively). Definitions are the same as in figure 6. Error bars indicate standard deviations on the mean.

be very short. Figure 10 shows the distribution of residence times (τ) to the DNA molecules, calculated as explained in section 2.2. We estimated that the average residence time to a given base-pair step, τ , is of the order of 10 ps. Moreover, the

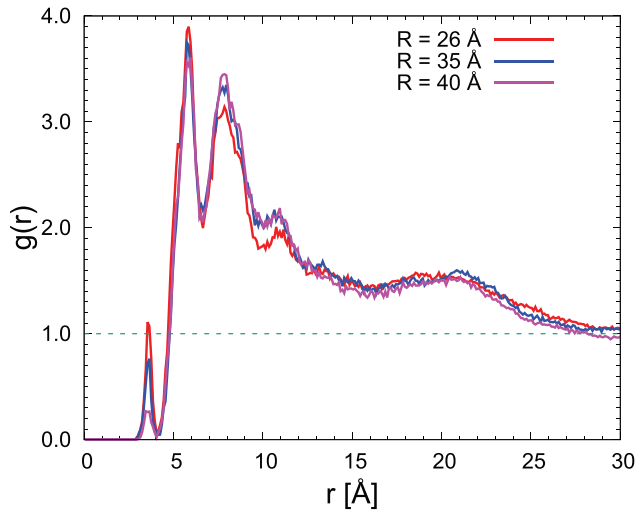


Figure 8. Pair correlation function $g(r)$ between the phosphate ions of DNA 2 and the sodium ions in solution. The two curves represent the $g(r)$ for $\alpha = 0$ and $R = 26, 35$ and 40 Å.

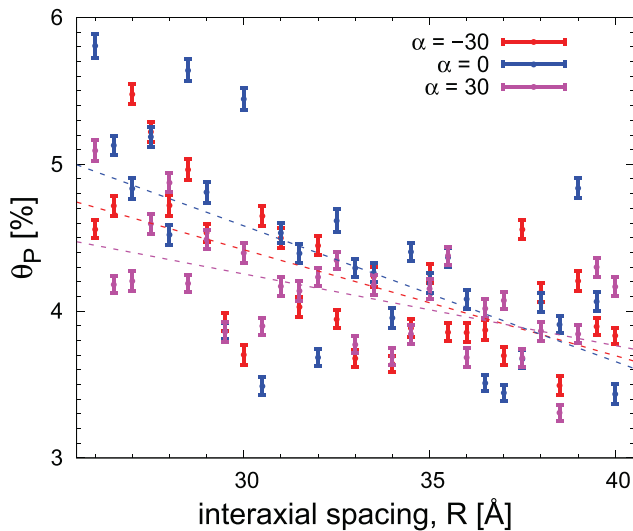


Figure 9. Contribution to the total charge compensation fraction due to sodium ions in a shell of thickness 5 Å from the phosphorous atoms of the DNA, plotted against the interaxial separation, for the three simulation sets. Dashed lines represent linear regressions to the data, and are intended as guides for the eye.

probability for larger τ values calculated from our MD simulation scales as $\sim \tau^{-4}$, so that long binding times are extremely improbable. Therefore, the ion distribution around the DNA molecules changes rapidly, due to fast ion exchange between the bulk and the DNA surface. This effect is a good indication that the (monovalent) ions in solution can reach equilibrium in the time window of our simulations.

4. Discussion

We have investigated the tilt-dependent interaction between two DNA oligomers using all-atom MD simulations. In general, we expected that the interaction between two DNA molecules in monovalent salt would be repulsive. The entropy cost of subtracting ions from the solution to minimize such

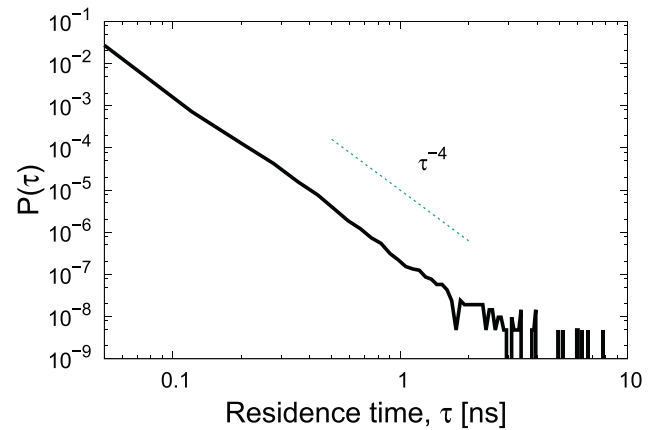


Figure 10. Distribution of residence time of sodium ions binding to the DNA molecules (thick solid line). The residence times were calculated as explained in section 2.2. The dashed line shows the τ^{-4} decay, which is roughly the one that this distribution follows.

repulsion is too high to result in a complete neutralization of the molecules. Therefore, as the molecules get closer, or more parallel to each other, this repulsion will increase. This effect is predicted by the Poisson–Boltzmann theory of electrostatic interactions in electrolytes, and is basically recapitulated by the HCC theory. The effect of the solvent is to screen the electrostatic interactions in an exponential fashion, with the characteristic length given by the Debye screening length, which in our ionic conditions is around 7 Å.

Our simulations show consistently a higher percentage of sodium ions adsorbed in the minor groove of DNA compared to the major groove, consistent with earlier MD simulations [54]. With this ion distribution, KL theory predicts strong helix-dependent effects (see dashed lines in figures 2 and 3).

The chirality of the DNA molecules suggests that the interaction between the two oligomers should be asymmetric as a function of their interaxial angle and azimuthal orientation. However, in the MD simulations the interaction is symmetric up to the accuracy of the method. These results are in agreement with primitive electrolyte model simulations of [24, 25].

Given the small number of sodium ions tightly bound to the phosphorous groups, masking of the negative charges on the phosphorous atoms due to strong binding of positive ions is not likely the reason for the absence of helix-specific electrostatic effects (see figure 8). Instead, it is possible that the basic assumption of the KL theory, that the charge distribution around the DNA molecules is fixed, is not valid in the case of DNA immersed in a solution of monovalent ions. We calculated the average residence time of sodium ions at a given DNA base-pair step, and found that it is very short (tens of picoseconds), and longer binding times are very unlikely. Therefore, the positive charge distribution around DNA must be seen as an ionic cloud which is smeared over the molecular surfaces. In this case, helix-specific electrostatic effects are still expected to be present (see Introduction). However, the distribution of ions around the DNA molecules in the simulations is affected by the geometry of the system (see figures 6 and 7), in a way which is not described within the framework of the KL theory. An additional contribution to the free energy

due to this effect should be included in the theoretical description of the system. This is perhaps one of the reasons why the KL theory seems to overestimate the effects due to the DNA chirality.

In the case of counterions that specifically adsorb in the DNA grooves, as in the systems described in [3], the predictions of the electrostatic zipper model [18] are expected to work best, and in those cases we still expect that helix-specific electrostatic effects will be relevant. In these cases, the time-scales for ion exchange at the DNA surface may be significantly longer.

The theory of homogeneously charged cylinders is in broad agreement with our simulation data. However, the dependence of the PMF on the interaxial separation shows additional complexity (see figure 2 in the range of 30–34 Å). This may be due to the fact that the electrostatic interaction is not the only interaction component, although it is surely the most significant. Hydration forces, due to structuring of the hydrogen bond network of water around DNA, are the most prominent candidates to explain the discrepancy observed (see figure 2), and the range of interaxial separation at which the difference is largest is consistent with the hydration force profile [55]. Furthermore, ion–ion and ion–water correlations are not included in the mean-field approach of the KL theory, and may also be important. It is interesting to notice that the values of θ from the fit of the PMF data to the HCC theory are consistent with the value of θ calculated from geometrical calculations, validating the results from both approaches.

Our findings shed light on the interactions between two DNA molecules separated by a close distance in monovalent salt solution. Surprisingly, we find that the mutual azimuthal orientation of the DNA oligomers has little effect on their effective interaction. This result may raise new questions about our current understanding of the complex behaviour of the cholesteric phase of DNA, which has been recently described in great detail [13, 14]. Specifically, it was thought that the electrostatic component of the interactions involved in building the cholesteric phase had to predominantly favour a right-handed cholesteric phase. In contrast, our results show that in monovalent salt the interaction of a pair of DNA oligomers does not favour any particular handedness, suggesting that alternative mechanisms are required to understand the observed experimental behaviour.

In conclusion, we have estimated the PMF of interacting DNA oligomers using all-atom MD simulations, and found that the effects due to the chiral structure of the DNA charges are masked by the thermal noise in the system. Our simulations, therefore, reveal that such effects, within the approximations built in the simulation modelling, are of the order of $1k_B T$ for the system consisting of two interacting DNAs each of 24 base pairs.

Acknowledgments

We thank A Kornyshev for useful advice and for help with editing the manuscript.

This material is based upon work supported by the US Department of Energy, Office of Science, and Office of Basic Energy Research.

This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231.

RC was supported by the United Kingdom Engineering and Physical Sciences Research Council (grant No. EP/H004319/1), the French Institut National du Cancer, Grant No. INCa 5960, by the French Agence Nationale de la Recherche, Grant No. ANR-13-BSV5-0010-03, and the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n 608959.

We acknowledge support of the Spanish Ministry of Economy and Competitiveness, 'Centro de Excelencia Severo Ochoa 2013–2017'. We acknowledge the support of the CERCA Programme/Generalitat de Catalunya.

References

- [1] Rau D and Parsegian V 1992 *Biophys. J.* **61** 246–59
- [2] Bloomfield V 1996 *Curr. Opin. Struct. Biol.* **6** 334–41
- [3] Kornyshev A, Lee D, Leikin S and Wynveen A 2007 *Rev. Mod. Phys.* **79** 943–96
- [4] Carrivain P, Courmac A, Lavelle C, Lesne A, Mozziconacci J, Paillusson F, Signon L, Victor J M and Barbi M 2012 *Soft Matter* **8** 9285–301
- [5] Chen Y J, Groves B, Muscat R A and Seelig G 2015 *Nat. Nanotechnol.* **10** 748–60
- [6] Pelta J, Livolant F and Sikorav J 1996 *J. Biol. Chem.* **271** 5656
- [7] Pelta J, Durand D, Doucet J and Livolant F 1996 *Biophys. J.* **71** 48–63
- [8] Kornyshev A and Leikin S 2000 *Phys. Rev. E* **62** 2576–96
- [9] Kornyshev A, Leikin S and Malinin S 2002 *Eur. Phys. J. E* **7** 83–93
- [10] Tombolato F and Ferrarini A 2005 *J. Chem. Phys.* **122** 054908
- [11] Stanley C, Hong H and Strey H 2005 *Biophys. J.* **89** 2552–7
- [12] Frezza E, Tombolato F and Ferrarini A 2011 *Soft Matter* **7** 9291–6
- [13] Nakata M, Zanchetta G, Chapman B, Jones C, Cross J, Pindak R, Bellini T and Clark N 2007 *Science* **318** 1276–9
- [14] Zanchetta G, Giavazzi F, Nakata M, Buscaglia M, Cerbino R, Clark N A and Bellini T 2010 *Proc. Natl Acad. Sci. USA* **107** 17497–502
- [15] Cortini R, Kornyshev A, Lee D and Leikin S 2011 *Biophys. J.* **101** 875–84
- [16] Cortini R, Lee D and Kornyshev A 2012 *J. Phys.: Condens. Matter* **24** 162203
- [17] Timsit Y, Várnai P and Wölfel S 2010 *PLoS One* **5** 430–40
- [18] Kornyshev A and Leikin S 1999 *Phys. Rev. Lett.* **82** 4138–41
- [19] Lee D, Wynveen A, Kornyshev A and Leikin S 2010 *J. Phys. Chem. B* **114** 11668–80
- [20] Kornyshev A, Lee D, Wynveen A and Leikin S 2011 *Nucl. Acids Res.* **39** 7289–99
- [21] Hsiao C, Tannenbaum E, VanDeusen H, Hershkovitz E, Perng G, Tannenbaum A R and Williams L D 2008 Complexes of nucleic acids with group I and II cations *Nucleic Acid-Metal Ion Interactions* (London: The Royal Society of Chemistry) pp 1–38
- [22] Kornyshev A and Leikin S 1997 *J. Chem. Phys.* **107** 3656
- [23] Manning G *et al* 1978 *Q. Rev. Biophys.* **11** 179–246

- [24] Allahyarov E and Löwen H 2000 *Phys. Rev. E* **62** 5542
- [25] Allahyarov E, Gompper G and Löwen H 2004 *Phys. Rev. E* **69** 041904
- [26] Baldwin G, Brooks N, Robson R, Wynveen A, Goldar A, Leikin S, Seddon J and Kornyshev A 2008 *J. Phys. Chem. B* **112** 1060–4
- [27] Danilowicz C, Lee C, Kim K, Hatch K, Coljee V, Kleckner N and Prentiss M 2009 *Proc. Natl Acad. Sci. USA* **106** 19824
- [28] Lee D, Danilowicz C, Rochester C, Kornyshev A and Prentiss M 2016 *Proc. R. Soc. A* **472** 20160186
- [29] Kornyshev A 2010 *Phys. Chem. Chem. Phys.* **12** 12352–78
- [30] Savelyev A and Papoian G A 2007 *J. Am. Chem. Soc.* **129** 6060–1
- [31] Luan B and Aksimentiev A 2008 *J. Am. Chem. Soc.* **130** 15754–5
- [32] Dai L, Mu Y, Nordenskiöld L and van der Maarel J 2008 *Phys. Rev. Lett.* **100** 118301
- [33] Yoo J and Aksimentiev A 2011 *J. Phys. Chem. Lett.* **3** 45–50
- [34] Yoo J and Aksimentiev A 2012 *J. Phys. Chem. B* **116** 12946–54
- [35] Pasi M et al 2014 *Nucl. Acids Res.* **42** 12272–83
- [36] Mocci F and Laaksonen A 2012 *Soft Matter* **8** 7289–99
- [37] Várnai P and Timsit Y 2010 *Nucl. Acids Res.* **38** 4163
- [38] Lu X J and Olson W K 2003 *Nucl. Acids Res.* **31** 5108–21
- [39] Olson W et al 2001 *J. Mol. Biol.* **313** 229–37
- [40] Humphrey W, Dalke A and Schulten K 1996 *J. Mol. Graph.* **14** 33–8
- [41] Foloppe N and MacKerell A D Jr 2000 *J. Comput. Chem.* **21** 86–104
- [42] MacKerell A D, Banavali N and Foloppe N 2000 *Biopolymers* **56** 257–65
- [43] Phillips J C, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel R D, Kale L and Schulten K 2005 *J. Comput. Chem.* **26** 1781–802
- [44] Miyamoto S and Kollman P A 1992 *J. Comput. Chem.* **13** 952–62
- [45] Darden T, York D and Pedersen L 1993 *J. Chem. Phys.* **98** 10089–92
- [46] Martyna G J, Tobias D J and Klein M L 1994 *J. Chem. Phys.* **101** 4177–89
- [47] Feller S E, Zhang Y, Pastor R W and Brooks B R 1995 *J. Chem. Phys.* **103** 4613–21
- [48] Kumar S, Rosenberg J M, Bouzida D, Swendsen R H and Kollman P A 1995 *J. Comput. Chem.* **16** 1339–50
- [49] Roux B 1995 *Comput. Phys. Commun.* **91** 275–82
- [50] Grossfield A WHAM: the weighted histogram analysis method, version 2.0.9 (<http://membrane.urmc.rochester.edu/content/wham>)
- [51] Efron B and Tibshirani R J 1994 *An Introduction to the Bootstrap* (Boca Raton, FL: CRC Press)
- [52] Marquardt D W 1963 *J. Soc. Ind. Appl. Math.* **11** 431–41
- [53] Galassi M, Gough B, Jungman G, Theiler J, Davies J, Booth M and Rossi F 2009 *GNU Scientific Library Reference Manual* 3rd edn (Bristol: Network Theory)
- [54] Feig M and Pettitt B M 1999 *Biophys. J.* **77** 1769–81
- [55] Mariani P and Saturni L 1996 *Biophys. J.* **70** 2867–74