Editors' Suggestion

Interactions between identical DNA double helices

Chun-Liang Lai,[1] Chuanying Chen,[1] Shu-Ching Ou,[1] Mara Prentiss,[2] and B. Montgomery Pettitt[1,*]
[1]*Department of Biochemistry & Molecular Biology, Sealy Center for Structural Biology and Molecular Biophysics, University of Texas Medical Branch, Galveston, Texas 77555, USA*
[2]*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*

(Received 3 October 2019; accepted 2 January 2020; published 24 March 2020)

The molecular mechanism of specific interactions between double stranded DNA molecules has been investigated for many years. Problems remain in how confinement, ions, and condensing agents change the interactions. We consider how the orientational alignment of DNAs contributes to the interactions via free energy simulations. Here we report on the effective interactions between two parallel DNA double helices in 150-mM NaCl solution using all atom models. We calculate the potential of mean force (PMF) of DNA-DNA interactions as a function of two coordinates, interhelical separation of parallel double helices and relative rotation of a DNA molecule with respect to the other about the helical axis. We generate the two-dimensional PMF to better understand the effective interactions when a DNA molecule is in juxtaposition with another. The analysis of the ion and solvent distributions around the DNA and particularly in the interface region shows that certain alignments of the DNA pair enhance the interactions. At local free energy minima in distance and alignment, water molecules and Na^+ ions form a hydrogen bonded network with the phosphates from each DNA. This network contributes an attractive energy component to the DNA-DNA interactions. Our results provide a molecular mechanism whereby local DNA-DNA interactions, depending on the helical orientation, give a potential mechanism for stabilizing pairing of much larger lengths of homologous DNA that have been seen experimentally. The study suggests an atomically detailed local picture of relevance to certain aspects of DNA condensation or aggregation.

DOI: 10.1103/PhysRevE.101.032414

I. INTRODUCTION

The interactions between pairs of double stranded DNA (dsDNA) are important for the packaging of DNA in cells and viruses [1–4], homologous recombination [5–7] and in application to certain biomaterials [8,9]. Direct force measurements [3,10–12] and theoretical studies [11,13,14] on counterion-condensed DNAs reveal that forces governing dsDNA-dsDNA interactions consist of competing components. The electrostatic repulsion has been previously modeled via the continuum solvation double-layer repulsion caused by the phosphates of the DNA backbones screened by condensed counterions [15]. At surface separations less than ~10 Å, solvation repulsions originating in part from partially ordered waters near the DNA surface also play a role [15]. Repulsion can be effectively enhanced due to the bending fluctuations of DNA dictated by its persistence length [16]. Some have postulated an attractive force due to correlated counterion fluctuations [11,17].

Treated as a charged rod in solution, the interactions between dsDNA-pairs are nonspecific. However, many DNA properties are sequence dependent in atomic detail. As a result dsDNA-dsDNA interactions could have local sequence specificity as well as geometric dependencies. Indeed, evidence from different experimental approaches demonstrated that homologous pairing is sequence dependent. Inoue *et al.* [18] revealed that in the presence of physiological concentrations of Mg^{2+} ions dsDNA molecules can efficiently and preferentially interact with the others bearing an identical sequence and length. Baldwin *et al.* [19] confirmed that dsDNA can recognize mutual sequence homology in extremely high concentrations of NaCl (>0.5 M) in the presence of crowding agents. Furthermore, Danilowicz *et al.* [20] used a parallel single molecule magnetic tweezers-based assay and specifically showed that homologous pairing can occur even in the absence of divalent cations or crowding agents. Recently Galdyshev and Kleckner [21,22] studied the sequence dependence of repeat-induced point mutations (RIPs). They discovered that a pair of homologous dsDNA can recognize each other if 3-4 base pairs are sequence matched in a periodicity of 11-12 base pairs and participating DNA molecules must be able to coalign along their lengths. The RIP experiments suggest that sequence dependent attractive interactions are governed by local attractive interactions between neighboring sequence matched regions of appropriately oriented dsDNA molecules, rather than global deformations that require sequence matching to extend along the entire length of both dsDNA molecules.

Regarding the mechanism underlying the sensing ability of homologous dsDNA, helical features and sequences of dsDNA have been considered to play an important role. One model [14,23–25] that requires sequence matching along the entire length of the dsDNA is based on mutual electrostatic complementarity. In that model, the repulsion of pairing of

*Author to whom correspondence should be addressed: mpettitt@utmb.edu

homologous dsDNAs is weakened as counterions preferentially bind in the grooves interacting with the neighboring DNA, and when the molecular features of the phosphates from two DNA double strands interdigitate or are aligned "zipper-like". In contrast to homologous sequences, nonhomologous ones cannot align well incurring a higher energy penalty to deform the duplex strands. However, this model is based on an electrostatic mean field and it cannot be reconciled by RIP data [21,22]. Another model is based on self-complementary contacts at their major-groove edges, in which short quadruplexes are formed by non-Watson-Crick hydrogen bond interactions between bases in the major grooves [26,27]. However, this model involves a large energy barrier, and provides less detail about the effects of counterions in mediating the interactions.

To further elucidate the molecular mechanism of sequence dependent dsDNA-dsDNA interactions, the present study explores the relationship among the water and ions between DNA surfaces and the resulting effective interactions. In addition, how the local alignment of DNA relative to each other contributes to the interactions is considered. We consider both distance, d, and mutual orientation, θ. We report the simulation of the effective interactions between two parallel identical dsDNAs in 150-mM NaCl solution as the potential of mean force (PMF) between dsDNA-dsDNA as a function of (d, θ) from the integral of the effective forces sampled with all-atom models.

We consider both the conventional one-dimensional PMF along the interhelical separation distance, d, as well as a two-dimensional (2D) PMF describing dsDNA molecule pairs, which additionally is sampled with relative rotation of parallel dsDNA molecules about each other, θ. For different dsDNA-dsDNA configurations we calculate the ion-distributions in the interface region and quantify the roles of Na^+ ions and water in mediating the dsDNA-dsDNA interactions. The study provides an atomically detailed model of the driving forces for DNA condensation or aggregation.

II. METHODS

A. MD simulations

Molecular dynamics (MD) simulations of a single dsDNA and two parallel dsDNA pairs were performed using NAMD 2.10 software [28] with the CHARMM 36 force field parameter set [29]. The DNA strands are rigid, have no open ends, and are effectively replicated infinitely as each end of the DNA duplex strand is covalently bonded to itself through the periodic boundaries. Particle Mesh Ewald [30] was used to calculate long-range electrostatic interactions, and van der Waals interactions were truncated at 12 Å. All bonds were constrained using the SETTLE algorithm [31] and equations of motion were integrated with a time step of 2 fs. Temperature was controlled with Langevin dynamics with a damping coefficient of 0.5/ps. The Nosé-Hoover method [32,33] with a Langevin piston was used to maintain the pressure at 1 atm with an oscillation period of 100 fs and a damping time scale of 50 fs. After 5000 steps of energy minimization, the system was heated up from 25 to 310 K with restraints on the phosphorus atoms of the DNA with a force constant of 500 kcal/(mol Å^2), except for the two base pairs at the edge of simulation box. The system was then switched to the NPT ensemble for production.

The initial structure of a 30-bp double stranded B-DNA molecule denoted as DNA1 (sequence 5'- GGGGGGGTTAT-GTCAGAACCGGCTGGGGGG) was generated and rotated so that the helical axis is the z axis. The sequence includes all possible base pair steps flanked by GC pairs. The single ds-DNA was then solvated in a TIP3P [34] water box with dimensions of 99 × 99 × 103 Å. Na^+ and Cl^- were randomly added to both neutralize the system and set the salt concentration at 0.15 M. The resulting system contained about 101 K atoms. After 200 ns of equilibration the solvent and ion atmosphere distributions were stable. Then a second duplex denoted DNA2, identical to DNA1, was generated by copying the coordinates and translating along the x axis away from the center of DNA1 by 41.5 Å including salt and solvent within 40 Å. The system with the pair of dsDNAs (DNA1 and DNA2) and sufficient salt and solvent molecules was then further equilibrated for 200 ns at which time the salt and solvent distributions were again stable.

In the calculation of the pair potential of mean force between two parallel dsDNA strands the second duplex, DNA2, was moved along the x axis and rotated about its helical axis with respect to DNA1. As a result, two collective variables, interhelical separation d and relative rotation angle θ, were considered (Fig. 1). The initial configuration of the DNA-DNA pair was extracted from the last snapshot of the solvent equilibration simulation mentioned above. We generated a series of distance sampling windows gradually reducing d from 41.5 to 22.0 Å with a step of 0.1 Å. For each window, in the *NVT* ensemble, the system was energy minimized and re-equilibrated for 100 ps. In this process, the DNA strands and the Na^+ ions within 10 Å from the DNA axis were fixed. After the windows were prepared, the configurations with d at multiples of 0.5 Å were selected for further equilibration and PMF computations. In this process, only the DNA strands remained fixed, and all ions and water were free to move. The second reaction coordinate θ was included when d was between 22.0 to 28.0 Å. This is the range where the most angular variation in the PMF would be expected. The windows of the simulation system were created by gradually rotating DNA2 about its helical axis with a step of 6°. Similar to the procedure in distance reduction, energy minimization and equilibration were performed on each window system. Then configurations 60° apart in θ were selected for further equilibration and production for PMF computations. For each simulation window, the trajectory was ∼70 ns in length, and the coordinates were saved for analysis at an interval of 2 ps.

Selected configurations of the DNA pairs with different θ are displayed in Fig. 1(a). For convenience, we denoted a configuration as (d, θ), e.g., model (22,60°) represents the configuration with $d = 22.0$ Å and $\theta = 60°$.

B. 2D-PMF calculation and error estimation

To investigate the potential of mean force (PMF) for the pairing of two parallel DNAs, we sampled the two independent reaction coordinates, d and θ, as shown in Fig. 1(b). The probability distribution of finding the pair of DNA strands,

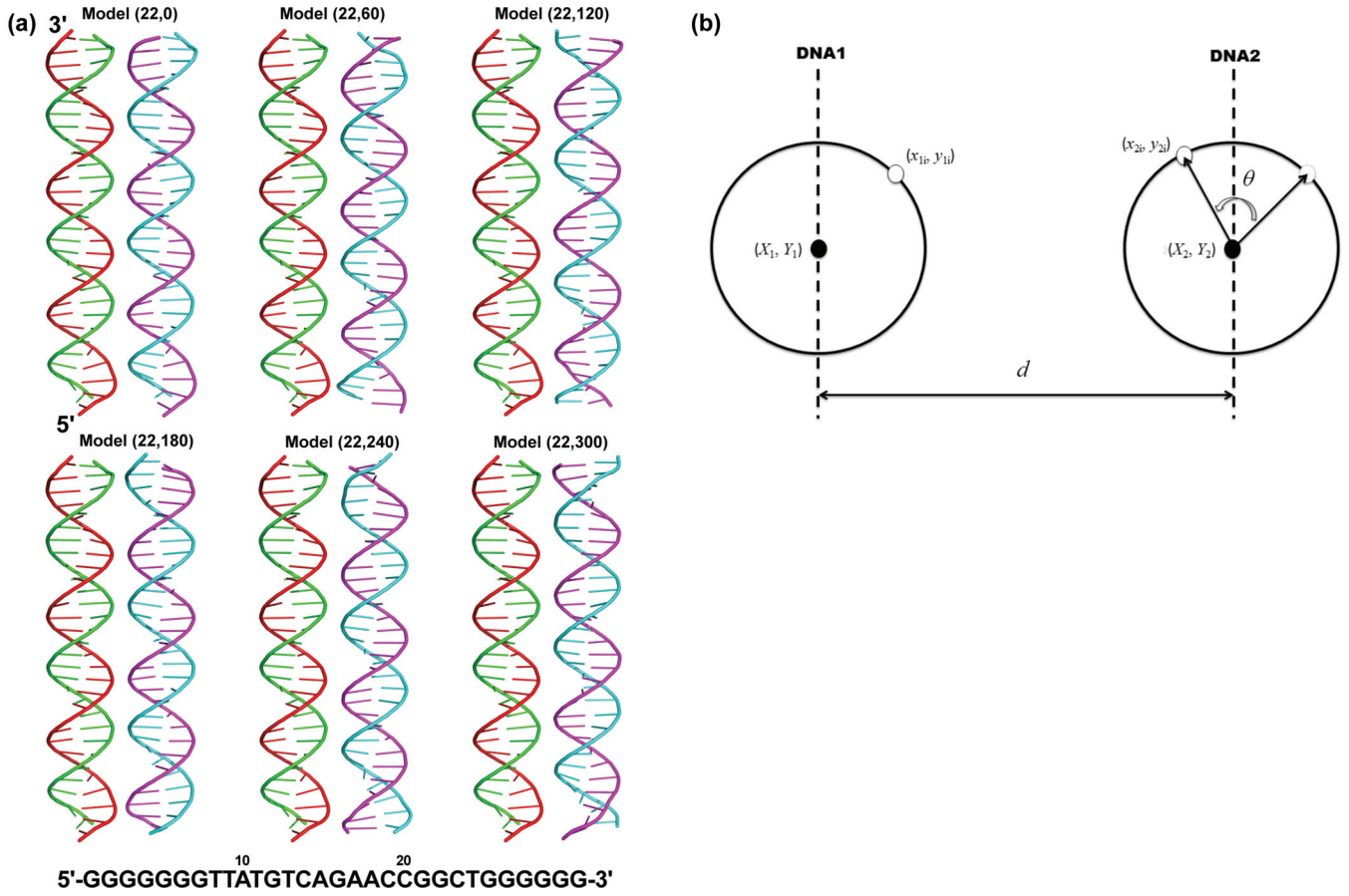

FIG. 1. (a) The configurations of the DNA pairs at $d = 22$ Å and $\theta = 0°$, 60°, 120°, 180°, 240°, and 300°. Color, red: strand A; green: strand B; magenta: strand E; cyan: strand F. (b) Coordinates of the two double helices, DNA1 and DNA2 for the PMF calculations.

$g(d, \theta)$, is related to the PMF, $W(d, \theta)$, by

$$g(d, \theta) = e^{-\beta W(d,\theta)}, \tag{1}$$

where β is the inverse of Boltzmann's constant times the absolute temperature. Details are given in the Supplemental Material [35].

III. RESULTS

A. Potential of Mean Force

We first calculated the one-dimensional (1D)-PMF as a function of the interhelical separation distance d, as shown in Fig. 2(a). The PMF for two parallel dsDNAs in the identical orientation shows the features and local minima expected from the combination of solution packing and electrostatic correlations. At distances beyond the packing correlation length we expect a smooth function given by traditional continuum electrostatics. That limit seems to be obtained beyond 30 Å in this case (Supplemental Material Fig. S1). Near a separation of 23 Å a minimum with a 2 kcal/mole barrier was found. A broader shallow feature appears around 25 Å. We will consider the mechanism that produces the local minima below.

When relative rotation angle θ was considered [Fig. 2(b)], we found that some rotations change the free energy barriers in the effective interactions between two parallel DNAs. While sequence variations may play a role we considered only a single homologous set in the current calculations. The rotation of the DNA viewed as a sequence independent backbone is related to the axial shift. If the helical pitch H is ∼34 Å, $\theta = 120°$ corresponds to a relative axial shift $\delta z = \sim 11$ Å, at which point the phosphate charge interaction energies are optimal. This is qualitatively consistent with the prediction from the "electrostatic zipper" theory [14]. That study showed that the interaction energy between two rods aligned with negatively charged helical lines of phosphates and positively charged ions is optimal at $\delta z = \sim H/4$, if 70% of absorbed ions are in the grooves. Our calculation supports the idea that the relative alignment of DNAs plays an important role in the interactions of double stranded DNAs. The features found here also reflect the effective interactions related to the orientations of the major and minor grooves. Salt and water are tightly correlated with the grooves and the phosphates, so it is necessary to investigate the relationship between the water and ions and the effective interactions.

B. Ion distribution around DNAs without rotation

For all the configurations of the DNA pairs at different d, there is considerable variation in relative free energy versus θ. (See Fig. 2 and Fig. S2B). To understand this variation we focus on the counterions, which not only screen the electrostatic repulsion between highly charged DNAs, but also

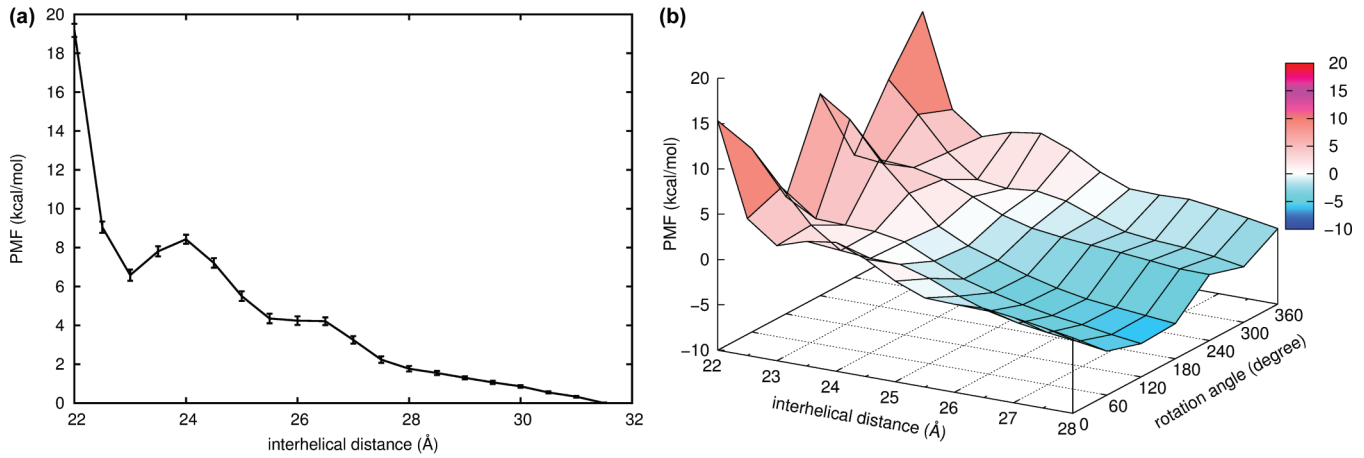

FIG. 2. Potential of mean force for two parallel dsDNAs. (a) The 1D-PMF as a function of d at $theta = 0$. The last point was set to zero. Block estimate error bars are shown at each point. (b) The 2D-PMF for two parallel dsDNAs, $W(d, \theta)$, as a function of the interhelical separation distance d and relative rotation angle θ. The single constant of integration (vertical shift) is arbitrary.

mediate other aspects of the DNA-DNA interactions and the associated packing forces.

The ion distribution around the single dsDNA [Fig. 3(a)] did not approach the continuum distribution until roughly 30 Å from the helical axis. For the double dsDNA system [Fig. 3(b)], at large interhelical distance, ∼41.5 Å, both the Na^+ and Cl^- distributions in the space between the helices remain perturbed relative to the single dsDNA system. We found significant rearrangement of the Na^+ ions in the region between the helices, particularly in the interface region of the DNAs. Importantly, two dsDNAs thus share an ion atmosphere in the interface region.

As d decreases, the Na^+ ions sample around and in between the surfaces of dsDNA. As a result, the fraction of charge around the DNA changes for $d < 60$ Å (roughly 40 Å surface to surface). The fraction of charge around the DNA is defined as the ratio of the total charge of bound positive charges to the total charge of the DNA. Many ion-DNA pair correlation functions have been proposed and used in the literature. To consider the role of ions in the range where significant correlation features appear in the PMF, we defined an ion as "bound" when it is within a boundary $dcut$ from the surface of the DNA. We wish to probe how much the ions in close proximity to the DNA, especially those associated with the grooves, are affected.

Here, as shown in Fig. 4, we estimated the fraction of counter balancing charge from the Na^+ ions, φ, around DNA1(DNA2), as the function of d for different values of $dcut$. At the surface of the strands, $dcut = 0$ Å, φ remains near ∼33–35% for all d ranging from 22 to 28 Å, which indicates that the close proximity of the DNAs to each other has essentially no influence on φ inside the grooves. φ increases with increasing $dcut$ and decreasing d. At $dcut = 6.0$ Å and $d = 26.0$ Å, φ is about ∼90%, which is beyond the 70% predicted by Manning counterion condensation theory for a single dsDNA. The large value of φ is due to the combination of Na^+ ions correlated with the DNA surface, and the Na^+ ion atmosphere shared with the neighboring DNA.

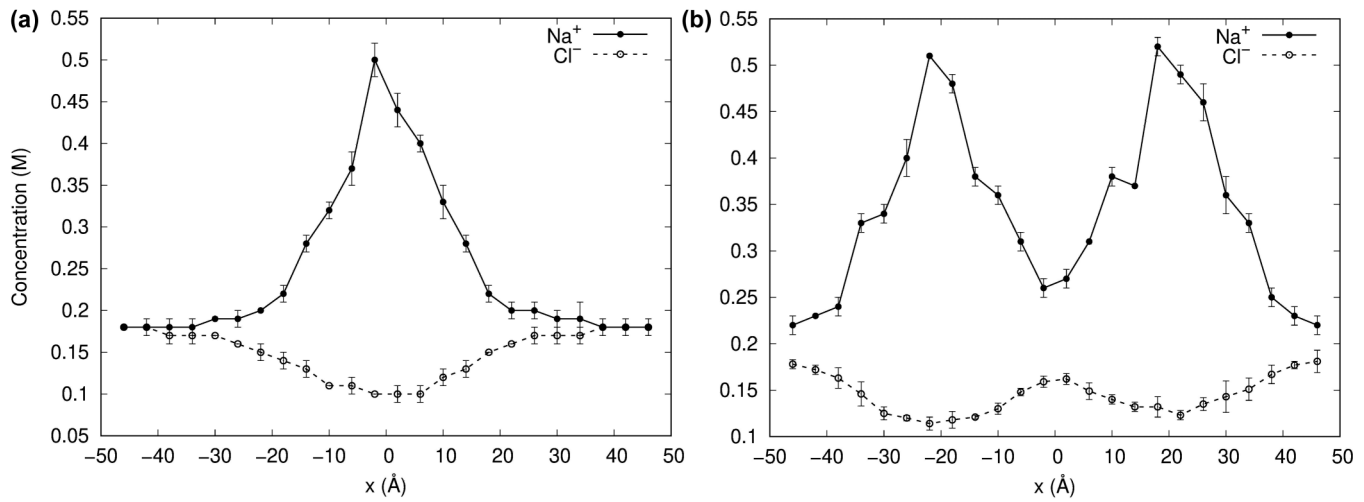

FIG. 3. Ion distributions as a function of distance averaged between 120 and 200 ns. (a) The ion distribution around a single dsDNA averaged over planes perpendicular to the helix axis. (b) The ion distribution around two dsDNAs that are positioned 41.5 Å apart.

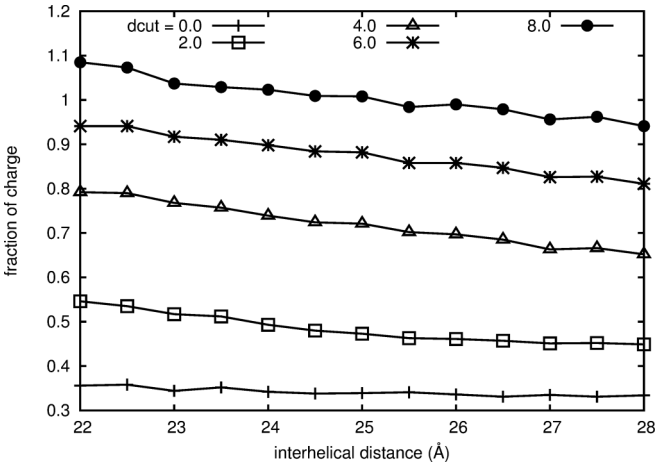

FIG. 4. The fraction of charge of the Na^+ ions around the dsDNA (DNA1) as a function of the interhelical distance at different surface boundaries, $dcut$.

In addition to the fraction of the charge of the Na^+ ions around the DNA, we also estimated the fraction of charge of the Na^+ ions in the minor groove, which remains at ∼26% for all d and $dcut$. This indicates again that the fraction of the charge of Na^+ ions inside the DNA grooves would not be expected to be strongly affected by DNA aggregation.

C. Ion distribution in the interface region without rotation

Two dsDNAs can share the ion atmosphere even at distances of over 40 Å [Fig. 3(b)]. We focus on the interface region defined as the space between the dsDNA surfaces and investigate the ion distribution by considering the position dependence of the Na^+ ions in the interface region.

We calculated the distribution of the minimum distance between the Na^+ ions and the phosphates for different d Fig. S3A and S3B). Generally, distributions of Na^+ ions around the phosphates in the interface region are similar for both DNAs. For each configuration, the distribution profile has three peaks at ∼2.4–2.5 Å, corresponding to a direct contact (P-Na), at ∼4.4–4.5 Å, corresponding to a water-mediated contact (P-W-Na) and around 6.6 Å, corresponding to a contact via two more water molecules.

Comparing the distributions among all the configurations, we found that (22,0°) and (23,0°) exhibit significant differences in the intensity of the peaks for direct contact and water-mediated. For (22,0°), the majority of Na-phosphate contacts are direct. This is reasonable because the minimum distance between surface atoms on the two dsDNAs is only 2.8 Å. The Na^+ ions can be easily trapped in this narrow space and directly absorbed by negatively charged phosphate groups. For Model (23,0°), the minimum distance between the two DNA surfaces is 3.8 Å, which provides more space. The high peak for water-mediated contact is correlated with the first minimum in potential well in the PMF [Fig. 2(a)].

We further investigated how the Na^+ ions and water mediate the interactions of the DNA pairs creating molecular features of the PMF. The DNA-DNA contacts via the Na^+ ions with the phosphates were categorized into three groups: direct ion contact (P-Na-P), hemihydrated ion contact (P-W-Na-P or P-Na-W-P), and hydrated ion contact (P-W-Na-W-P). Ion contacts mediated via three or more waters were not considered for further analysis. The contact frequencies for phosphate ion complexes formed by sodium ions with possible waters hydrogen bonded with phosphates are summarized in Table S1. Generally, we found that the occurrence of the contact P-Na-P is rare at any distance except for (22,0°), which is in a region with considerable net repulsion. The hemihydrated structures are not common at most orientations and distances. However, the fully hydrated ion structures, P-W-Na-W-P, were observed in most models, especially those in the broad second minimum between 24–25 Å, except for those at large interhelical separation ($d \geqslant 28$ Å).

The high occurrence of both P-Na-P and P-Na-W-P configurations in Model (22,0°) indicates that the DNA pairs are bridged by Na^+ ions in both a direct and water-mediated manner. Nevertheless, the bridging is not strong enough to overcome the other repulsive interactions between two DNA molecules at that separation (Fig. 2). The hydrated ion configuration (P-W-Na-W-P) is highest at (23,0°) and contributes the stability of the first minimum in the PMF.

To further illustrate Na^+ ions in mediating the dsDNA-dsDNA interactions, we identified effective sodium sites in all configurations. An effective sodium site reflects not only the density of Na^+ ions in a location but also the occupancy of this location. A site having a high density of Na^+ ions implies that either a Na^+ ion is kinetically stable at this particular site, or several mobile Na^+ ions are able to populate this site. The occupancy of a site is defined as a fraction of time during which the site is occupied by Na^+ ions. The majority of the sites have occupancy of 0.1 and 0.2 (Fig. S4), and they are either located in the major grooves and at the edges of the grooves or do not participate in the direct or water-mediated DNA-DNA interactions. Figure 5 displays the effective sodium sites colored in the occupancy ranging from 0.3 to 1.0. These effective sites are mainly located in the interface region of the dsDNAs and in the minor groove that has been observed in previous studies [36,37]. The interfacial region of Model (23,0°) has 11 effective sodium sites, almost similar to Model (22, 0°) that has 13 and much more than $0 \sim 7$ sites in the other models. At $d = 23.0$ Å, Na^+ ions are kinetically stabilized by water molecules, which are hydrogen bonded with phosphate groups of the respective DNAs and coordinated with Na^+ ions. Consequently an interaction network with hydrogen bonds and charge-dipole forces is formed via hydrated ion contact. This network might contribute to a locally attractive component of the DNA-DNA interactions, which leads to the local energy minimum.

D. Ion distribution in the interface region with rotation

The calculations of the PMF show that the DNA-DNA effective interactions can be modulated by changing the alignment of two DNAs. Particularly at $d = 22.0$ Å, the curve as a function of θ has a pronounced "W" shape [Fig. 2(b) and also Fig. S2B] with local energy minima at $\theta = 120^\circ$ and 240°. We calculated the distribution of the distance between Na^+ ions and phosphates in the interface region at different θ. The results are shown in Fig. S3C and S3D for DNA1 and DNA2, respectively.

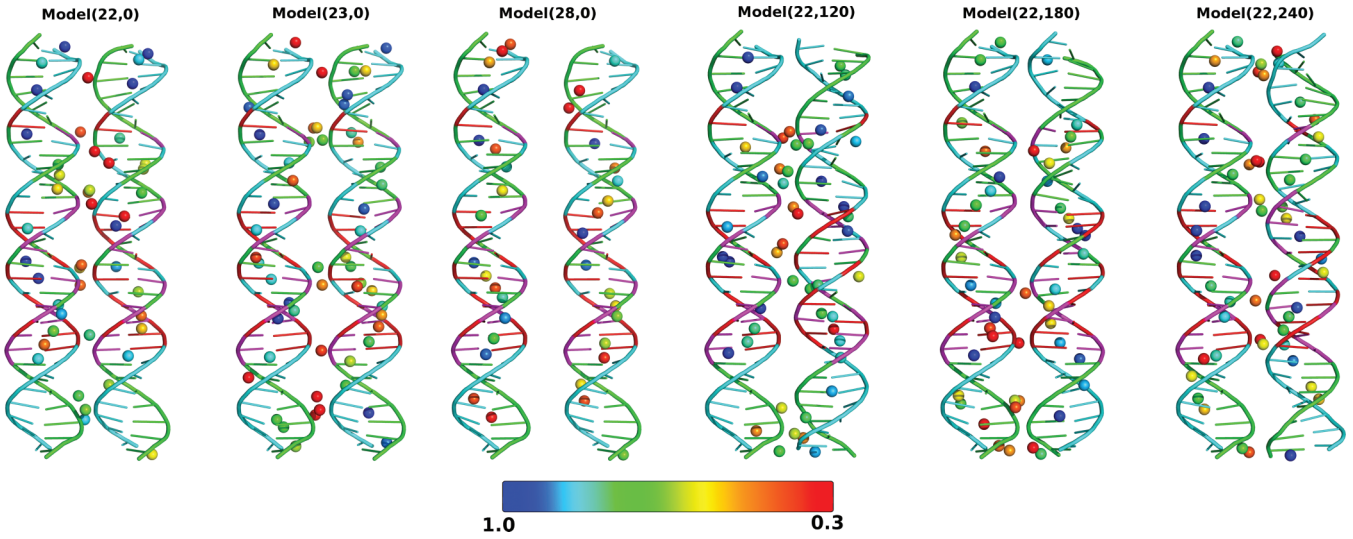

FIG. 5. Effective sodium sites in the grooves and in the interface region for different configurations. The sodium sites are colored in occupancy scaling from 1.0 to 0.3. To identify effective sodium sites in a particular region, we first divided the trajectory into several blocks and generated a sodium density map using a grid size of 0.5 Å for each block; then for the grid points whose densities are larger than a threshold (>0.01 # / point), we extracted the points that are ∼70% overlapped; we smoothed the points by density-weighted averaging the six-closest neighbors [37,38]. The sites were finally identified when the separation of the sites was larger than 2.8 Å. The DNA bases are colored: A, red; T, magenta; G, green; and C, cyan.

We observed that the distribution profiles in the interface region for DNA1 and DNA2 have some differences. For direct contact at ∼2.4–2.5 Å, the highest peak comes from Model (22,0°) for both DNAs, and then the peak shrinks in the order of (22, 240°) > (22, 180°) > (22, 120°) for DNA1, and (22, 180°) > (22, 120°) > (22, 240°) for DNA2. For water-mediated contact at ∼4.4–4.5 Å, the height of the peak decreases in the order of (22,240°) > (22,120°) > (22,0°) > (22,180°) for DNA1, and (22,240°) > (22,0°) > (22,120°) > (22,180°) for DNA2. So the contact of Na^+ ions with a dsDNA molecule favors direct contact at $\theta = 0°$ and 180°, while favoring a water-mediated one at $\theta = 120°$ and 240°.

The frequencies of the DNA-DNA contacts are listed in Table S2. Generally, the direct ion contact P-Na-P rarely happens when $\theta \neq 0$. Model (22,180°) favors hemihydrated ion contact P-W-Na-P more than hydrated ion contact P-W-Na-W-P. For both (22,120°) and (22,240°), P-W-Na-W-P is dominant. Additionally, as shown in Fig. 5, (22,120°) and (22,240°) have 12 effective sodium sites in the interface region, respectively, more than (22,180°). This finding further suggests that the hydrated ion contact network exists in (22,120°) and (22,240°), which contribute an attractive component in the DNA-DNA interactions.

These different ways of Na^+ ions to mediate the DNA-DNA interactions can be attributed to the alignment of two dsDNAs. $\theta = 0°$, 120°, 180°, and 240° correspond to $\delta z/H$, the ratio of relative axial shift δz to helical pitch H, equal to 0, 0.33, 0.50, and 0.67, respectively. When $\theta = 0°$ and 180°, $\delta z/H = 0$ and 0.50, respectively, which means the phosphate groups from the respective DNAs are head-to-head. When $\theta = 120°$ and 240°, $\delta z/H = 0.33$ and 0.67, respectively, which represent the cases when phosphate groups from one DNA partially interdigitate into the minor grooves of the other DNA. This configuration not only increases formation of the hydrated ion contact, but also provides opportunity for Na^+ ions in the minor groove to electrostatically interact with the negatively charged phosphates of the neighboring dsDNA.

We observed variations of effective sodium sites among the grooves, which are related to orientation of the DNAs and the differences caused by sequence. Na^+ ions prefer the minor groove over the major groove, and reside at A-T longer than G-C given choices in the potential model [37]. The DNA sequence here has G-C pairs at the beginning and end whereas it is more varied in the middle. So, high-occupancy sites of Na^+ ions in the minor groove appear concentrated in the middle of the sequence. Moreover, as one dsDNA rotates, the average locations of Na^+ ions in the minor groove change, which influences the interactions between two dsDNAs.

IV. DISCUSSION AND CONCLUSIONS

It has long been known that in an aqueous solution, self-aggregation of dsDNAs usually depend on counterions with charge often of 3+ or greater, e.g. $Co(NH_3)_6^{3+}$, spermidine (3+), spermine (4+) [2,4]. Divalent counterions Mn^{2+} can also induce the dsDNA condensation as well as Mg^{2+} in high concentration. For monovalent counterions, self-aggregation of dsDNAs is more rare [11,39,40]. Although screening effects from monovalent counterions are not strong enough to overcome repulsion between dsDNA and dsDNA even at large separation, it does not mean a lack of existence of local attractive contributions to the interactions of dsDNAs. Indeed, Danilowicz *et al.* [20] observed homologous pairing can occur even in monovalent solutions. Here we performed MD simulations on the pairing of two parallel identical DNA molecules in 150-mM NaCl solution. By calculating the PMF along two reaction coordinates, the interhelical separation d and the relative rotation angle θ, we explored the molecular mechanism of interactions between two dsDNAs at short surface-to-surface distances (within 8 Å). The important finding from the current

work is that potential wells were found along the 2D-PMF profiles, indicating an attractive component to the interactions in pairing of two dsDNAs. Such an attractive component is attributed to hydrated ion contact between two dsDNAs and relative helical alignment of dsDNA pairs. While the effective potential between our short, dilute strands shows the expected screened repulsion at long range, the presence of short range local minima gives a mechanism for stabilizing pairing of much larger lengths of homologous DNA, which have been seen experimentally [20].

A. Collective contributions from sodium ions and water

We found that the proximity of the two dsDNAs does not affect the fraction of charges of sodium ions φ ($\sim$33%) inside the DNA grooves, particularly in the minor groove ($\varphi = \sim$26%), but increases the fraction of charge around the DNA significantly. The ion atmosphere can be shared by two DNAs as far as a helical separation of 40 Å at this concentration.

When the DNA pairs are in close proximity ($d = 22.0$ Å for $\theta = 0^\circ$ or 180°), direct ion contact, P-Na-P, and hemi-hydrated ion contact, P-W-Na-P, dominate the interactions. However, such contacts cannot overcome the inherent electrostatic repulsion between two dsDNAs at that distance.

Nevertheless, water and Na^+ ions together can play a significant role in mediating DNA-DNA interactions. Particularly, we found that water molecules, hydrogen bonded to the phosphates from the respective DNAs, share a Na^+ ion to form five-body contact P-W-Na-W-P or an ion triplet where the central ion remains hydrated. This type of geometry dominates the DNA-DNA effective interactions in the configurations with a locally optimal separation of $d = 23.0$ Å, or rotation of $\theta = 120^\circ$ and 240° and in a broad lower minimum between 24–25 Å. It contributes an attractive component to the DNA-DNA interactions, leading to local minima in the effective potential well in the PMF.

B. Contributions from relative helical alignment

The interactions between two dsDNAs have local minima with respect to the mutual angle of orientation. In our case, $\theta = 120^\circ$ and 240° lead to an "interlocking" conformation through the minor groove, a similar conformation also found in dsDNA-dsDNA interactions in Mg^{2+} salt [40]. This "interlocking" conformation favors pairing more than the "head-to-head" conformation and satisfies aspects of the "electrostatic zipper" model [14], in which Na^+ ions in the minor groove electrostatically interact with the negatively charged phosphates of the neighboring dsDNA. The model demonstrates a helical specificity of DNA-DNA interactions. In addition, some variations of the density and occupancy of Na^+ ions inside the grooves were observed, which not only depend on the orientation of two dsDNAs but also on the sequence. Therefore, helical specificity is related to sequence specificity.

The binding of water and counterions to the grooves are sequence specific. For a single DNA molecule, hydration patterns in the major and minor grooves are different between A-T and G-C sequence [36]. A narrower hydration pattern is found in the A-T duplex minor groove. Different counterions behave differently in penetrating the grooves. Recent $\sim$1-μs MD simulations [41,42] show that both Na^+ and K^+ ions penetrate into both the minor and major groove. The occupancies can reach to 0.3 in the minor groove for both ions. Na^+ ions prefer the minor groove, and prefer A-T rather than G-C. K^+ ions prefer the major groove, and display different density patterns between A-T and G-C in the major groove [36,37]. Mg^{2+} ions are not observed in the minor groove of the DNA, and they can selectively bind in the major groove of G-C bases [43,44]. Spermine (Spm^{4+}) displays a preference for the minor groove but does not form long-lived and structurally defined complexes. Na^+ ions compete with Spm^{4+} for binding to the minor groove [45] suggesting a control mechanism for systems where polyamines modulate interduplex associations and condensation.

Sequence-dependent interactions of counterions with dsDNA can affect dsDNA-dsDNA interactions. Sequence-dependent attractions were observed experimentally in the presence of Na^+ and Mg^{2+} ions [20]. The combination of the MD simulations and experiment [46] provided evidence of sequence-dependent attraction between dsDNAs, in which in the presence of Spm^{4+} AT-rich dsDNAs associate more strongly than GC-rich ones. Spm^{4+} molecules are sterically unfavored or partially blocked by the major groove methyl groups of thymine, and as a result more spermines are condensed in the interhelical region and so can enhance the attraction.

The influence from counterions depends on their abilities to penetrate the grooves and reside in the grooves. In our case, the effective sodium sites were observed in the minor grooves of the dsDNAs. Some mutual juxtaposition geometries achieve occupancy larger than that in a single dsDNA molecule. In addition, occupancy and residence time of counterions at the grooves varies due to the proximity of the DNAs and mutual orientations. Further investigation of the linkage of those properties with d and θ will be considered in the future, as well as the type of counterions.

C. Implications for mechanisms of homologous pairing

It has been proposed that local pairing of dsDNAs in a "protein-free" environment is an initial step in homologous recombination [5,6]. That the recognition prefers identical DNA sequences to random sequences has been confirmed by experimental [18–20,47] and theoretical studies [24,25,27]. The self-complementary contact model [27] requires a large energy barrier to align two dsDNAs and form quadruplexes by non-Watson-Crick hydrogen bond interactions between bases in the major grooves. However, under the current simulation conditions we did not sample such structures.

In the mutual electrostatic complementarity model, Kornyshev and Leikin [24,25] related the recognition preference to twist angle between adjacent base pairs, $\Omega(z)$, which is related to $h[d\phi(z)/dz]$, and is sequence dependent. For random sequences, the twist angle variance is $\sim 5^\circ$. Pairing of random sequences of dsDNA requires DNA backbone deformation, which costs more energy than pairing of identical sequences of dsDNA. The model provides a global mechanism of homologous pairing which is not confirmed by the RIP data [21,22], since the RIP data indicates that sequence dependent pairing

can occur between dsDNA that are not sequence matched except in small regions separated by ∼ one helical turn. Those calculations were electrostatically mean field using torsionally rigid DNA, and so may be insensitive some local features of the DNA and the surrounding solution structure.

As observed in this study, both hydrated ion contact and relative helical alignment contribute local attractive components to interactions between dsDNAs. The former involves sequence-independent interactions, while the later satisfies "zipper-like" electrostatic complementarity [14] that is sequence related. Furthermore, we found the variations of the locations and occupancy of Na^+ ions are sequence-dependent and orientation-related, which contributes to a local mechanism in homologous pairing. So it is reasonable to hypothesize that in homologous pairing relative helical alignment enhances dsDNA-dsDNA interactions, and then the local variations of distribution of counterions play a role in the recognition preference. Further study on the sequence dependence of DNA-DNA interactions is required, which can be done by sliding a hetero sequence, such as the one used here along our helical z axis with respect to the other double strand.

ACKNOWLEDGMENTS

We thank Dr. K.-Y. Wong, Dr. C. Zhang, and Dr. G. Lynch for helpful discussions. We acknowledge partial support from the National Institutes of Health (NIH) GM066813 and the Robert A. Welch Foundation (H-0037 at the beginning and H-0013 at the end). A portion of the computational research was carried out through the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation (NSF) Grant No. ACI-1548562.

[1] V. A. Bloomfield, Curr. Opin. Struct. Biol. **6**, 334 (1996).
[2] V. A. Bloomfield, Biopolymers **44**, 269 (1997).
[3] H. H. Strey, R. Podgornik, D. C. Rau, and V. A. Parsegian, Curr. Opin. Struct. Biol. **8**, 309 (1998).
[4] V. B. Teif and K. Bohinc, Prog. Biophys. Mol. Biol. **105**, 208 (2011).
[5] B. M. Weiner and N. Kleckner, Cell **77**, 977 (1994).
[6] S. M. Burgess, N. Kleckner, and B. M. Weiner, Genes Dev. **13**, 1627 (1999).
[7] A. Barzel and M. Kupiec, Nat. Rev. Genet. **9**, 27 (2008).
[8] M. Endo, Y. Yang, and H. Sugiyama, Biomater. Sci. **1**, 347 (2013).
[9] A. Mehdinia, S. H. Kazemi, S. Z. Bathaie, A. Alizadeh, M. Shamsipur, and M. F. Mousavi, J. Pharm. Biomed. Anal. **49**, 587 (2009).
[10] D. C. Rau and V. A. Parsegian, Biophys. J. **61**, 246 (1992).
[11] R. Podgornik, D. C. Rau, and V. A. Parsegian, Biophys. J. **66**, 962 (1994).
[12] B. A. Todd, V. A. Parsegian, A. Shirahata, T. J. Thomas, and D. C. Rau, Biophys. J. **94**, 4775 (2008).
[13] B. I. Shklovskii, Phys. Rev. Lett. **82**, 3268 (1999).
[14] A. A. Kornyshev and S. Leikin, Phys. Rev. Lett. **82**, 4138 (1999).
[15] D. C. Rau, B. Lee, and V. A. Parsegian, Proc. Natl. Acad. Sci. USA **81**, 2621 (1984).
[16] H. H. Strey, V. A. Parsegian, and A. R. Podgornik, Phys. Rev. Lett. **78**, 895 (1997).
[17] I. Rouzina and V. A. Bloomfield, J. Phys. Chem. **100**, 9977 (1996).
[18] S. Inoue, S. Sugiyama, A. A. Travers, and T. Ohyama, Biochemistry **46**, 164 (2007).
[19] G. S. Baldwin, N. J. Brooks, R. E. Robson, A. Wynveen, A. Goldar, S. Leikin, J. M. Seddon, and A. A. Kornyshev, J. Phys. Chem. B **112**, 1060 (2008).
[20] C. Danilowicz, C. H. Lee, K. Kim, K. Hatch, V. W. Coljee, N. Kleckner, and M. Prentiss, Proc. Natl. Acad. Sci. USA **106**, 19824 (2009).
[21] E. Gladyshev and N. Kleckner, Nat. Commun. **5**, 3509 (2014).
[22] E. Gladyshev and N. Kleckner, Curr. Genet. **63**, 389 (2017).
[23] D. J. Lee, A. Wynveen, A. A. Kornyshev, and S. Leikin, J. Phys. Chem. B **114**, 11668 (2010).
[24] A. A. Kornyshev and S. Leikin, Phys. Rev. Lett. **86**, 3666 (2001).
[25] A. A. Kornyshev and A. Wynveen, Proc. Natl. Acad. Sci. USA **106**, 4683 (2009).
[26] S. McGavin, J. Mol. Biol. **55**, 293 (1971).
[27] A. K. Mazur, Phys. Rev. Lett. **116**, 158101 (2016).
[28] J. C. Phillips, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, J. Comput. Chem. **26**, 1781 (2005).
[29] K. Hart, N. Foloppe, C. M. Baker, E. J. Denning, L. Nilsson, and J. Alexander D. MacKerell, J. Chem. Theory Comput. **8**, 348 (2012).
[30] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. Pedersen, J. Chem. Phys. **103**, 8577 (1995).
[31] S. Miyamoto and P. A. Kollman, J. Comput. Chem. **13**, 952 (1992).
[32] S. Nosé, Mol. Phys. **52**, 511 (1984).
[33] W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).
[34] E. Neria, S. Fischer, and M. Karplus, J. Chem. Phys. **105**, 1902 (1996).
[35] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.101.032414 for details of methods of 2D-PMF calculations, figures of 1D-PMF profiles, the distance distributions of sodium ions and phosphates in the interface region and probabilities of occupancy at the effective sodium sites, and the table listing the contact frequencies of sodium ions, water molecules and the DNAs in the interface region.
[36] J. J. Howard, G. C. Lynch, and B. M. Pettitt, J. Phys. Chem. B **115**, 547 (2011).
[37] M. Feig and B. M. Pettitt, Biophys. J. **77**, 1769 (1999).
[38] C. Chen, B. W. Beck, K. Krause, and B. M. Pettitt, Proteins **62**, 982 (2006).
[39] G. L. Randall, B. M. Pettitt, G. R. Buck, and E. L. Zechiedrich, J. Phys.: Condens. Matter **18**, S173 (2006).
[40] B. Luan and A. Aksimentiev, J. Am. Chem. Soc. **130**, 15754 (2008).
[41] A. Perez, F. J. Luque, and M. Orozco, J. Am. Chem. Soc. **129**, 14739 (2007).

[42] P. D. Dans, I. Faustino, F. Battistini, K. Zakrzewska, R. Lavery, and M. Orozco, Nucleic Acids Res. **42**, 11304 (2014).
[43] T. K. Chiu and R. E. Dickerson, J. Mol. Biol. **301**, 915 (2000).
[44] W. Li, L. Nordenskiold, and Y. Mu, J. Phys. Chem. B **115**, 14713 (2011).
[45] N. Korolev, A. P. Lyubartsev, A. Laaksonen, and L. Nordenskiold, Biophys. J. **82**, 2860 (2002).
[46] J. Yoo, H. Kim, A. Aksimentiev, and T. Ha, Nat. Commun. **7**, 11045 (2016).
[47] D. J. Lee, C. Danilowicz, C. Rochester, A. Kornyshev, and M. Prentiss, Proc. Math. Phys. Eng. Sci. **472**, 20160186 (2016).