7018–7026 *Nucleic Acids Research, 2020, Vol. 48, No. 13* *Published online 15 June 2020*
doi: 10.1093/nar/gkaa499

Structure-guided DNA–DNA attraction mediated by divalent cations

Amit Srivastava[1,†], Raju Timsina[2,†], Seung Heo[2], Sajeewa W. Dewage[1], Serdal Kirmizialtin[1,*] and Xiangyun Qiu[2,*]

[1]Chemistry Program, Science Division, New York University Abu Dhabi, Abu Dhabi 129188, United Arab Emirates and [2]Department of Physics, George Washington University, Washington, DC 20052, USA

Received March 08, 2020; Revised May 26, 2020; Editorial Decision May 29, 2020; Accepted June 10, 2020

ABSTRACT

Probing the role of surface structure in electrostatic interactions, we report the first observation of sequence-dependent dsDNA condensation by divalent alkaline earth metal cations. Disparate behaviors were found between two repeating sequences with 100% AT content, a poly(A)-poly(T) duplex (AA-TT) and a poly(AT)-poly(TA) duplex (AT-TA). While AT-TA exhibits non-distinguishable behaviors from random-sequence genomic DNA, AA-TT condenses in all alkaline earth metal ions. We characterized these interactions experimentally and investigated the underlying principles using computer simulations. Both experiments and simulations demonstrate that AA-TT condensation is driven by non-specific ion–DNA interactions. Detailed analyses reveal sequence-enhanced major groove binding (SEGB) of point-charged alkali ions as the major difference between AA-TT and AT-TA, which originates from the continuous and close stacking of nucleobase partial charges. These SEGB cations elicit attraction via spatial juxtaposition with the phosphate backbone of neighboring helices, resulting in an azimuthal angular shift between apposing helices. Our study thus presents a distinct mechanism in which, sequence-directed surface motifs act with cations non-specifically to enact sequence-dependent behaviors. This physical insight allows a renewed understanding of the role of repeating sequences in genome organization and regulation and offers a facile approach for DNA technology to control the assembly process of nanostructures.

INTRODUCTION

Inspired by the highly charged nature of nucleic acid helices (1) and the prevalence of helix-helix interactions in biology (2), extensive efforts have been devoted to physical understanding of nucleic acid electrostatics. Double-stranded DNA (dsDNA, used interchangeably with DNA herein), the molecular form of the gene, is arguably the most studied helix, and the interaction partners of DNA range from diverse families of structural and regulatory proteins to small ligands to ubiquitous ions. Motivated by the biological significance of genome packaging—particularly in tight spaces such as viruses and sperm—and the development of DNA nanotechnology, cation-mediated DNA–DNA interactions have attracted substantial interests from experimentalists and theorists alike. The highly charged DNA backbones and well-defined structures of DNA and cations also make them an ideal model system for studying bio-molecular electrostatics.

Cations are able to interact with negatively charged DNA via non-specific electrostatic coupling and hence universally screen the like-charge repulsion between DNA. Cation valence is critically important. When cation valence is low, continuum theories of electrostatics suffice to quantitatively describe the effect of ion screening, primarily through considerations of competing Coulomb and entropy forces. Remarkably, multivalent cations (i.e. valence $\geq$ 3) can go beyond screening and induce attraction between dsDNA at sub-mM concentrations (3,4). Since continuum theories always predict DNA–DNA repulsion, the multivalent cation induced attraction (MCIA), driven by electrostatics, necessarily arises from spatial correlations between DNA charges and ions. However, the exact physical mechanisms of MCIA, i.e. how DNA charges and ions are correlated, remain unclear.

In contrast with the broad interest in the role of ions, much less studied is the role of DNA structure in DNA–

*To whom correspondence should be addressed. Tel: +971 2 628 5727; Fax: +971 2 659 0794; Email: serdal@nyu.edu
Correspondence may also be addressed to Xiangyun Qiu. Tel: +1 202 994 6537; Fax: +1 202 994 3001; Email: xqiu@gwu.edu
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.
Present address: Sajeewa W. Dewage, Department of Food Science & Technology Faculty of Livestock Fisheries & Nutrition, Wayamba University of Sri Lanka, Makandura, Gonawila 60170, Sri Lanka.

DNA interactions. Recent studies started shedding light on this issue. Triple-stranded DNA (tsDNA) was the first helix found to be condensed by divalent alkaline earth metal cations (e.g. Mg^{2+} and Sr^{2+}) (5) which are known to not condense dsDNA. The enhanced attraction between tsDNA was attributed to its more closely spaced phosphate backbones, where cations bind phosphates externally and mediate attraction via ion bridging. Note that binding in this context is non-ion-specific, and that ions are territorially localized around surface charges while retaining their hydration shells. The importance of groove geometry became evident from the resistance of dsRNA to condensation, which was attributed to the deeper RNA major grooves internalizing bound ions rather than exposing them to neighboring helices (6–8). Interestingly, the groove structures can also be tuned by DNA sequences or modifications which were observed to modulate DNA–DNA interactions in the presence of chain-shaped poly-cationic ligands, such as spermine and poly-lysine (9,10). Specifically, the methyl groups of thymine and methylated cytosine (mC) bases act as steric barriers to ligand binding in the grooves, which in turn favors external ligand binding to phosphates and mediates inter-DNA attraction via ligand/ion bridging. This leads to the significant revelation that the attraction between nucleic acids duplexes increases with the content of A-T and G-mC base pairs. Altogether, the molecular level structural features are shown to be capable of modulating DNA–ion–DNA interactions. Consequently, the molecular details of DNA and ions are both important for understanding nucleic acids electrostatics.

Here, we report the first observation of sequence-dependent dsDNA condensation by divalent alkaline earth metal cations. In addition to their lower valence, their monoatomic shape and smaller sizes are qualitatively different from the chain-shaped poly-cationic ligands with valence ≥ 3 used in previous studies of dsDNA condensation. We also note that the DNA sequence specificity on ion binding has been studied previously and reviewed in Ref. (11), whereas this study focuses on the role of the sequence specificity in DNA condensation. Rather curiously, we observed that duplexes of repeating A-T base pairs (AA-TT) are condensed by all alkaline earth metal ions tested (Mg^{2+}, Ca^{2+}, Sr^{2+} and Ba^{2+}) but not by chain-shaped divalent ligands such as putrescine. In contrast, alternating A-T and T-A base pairs (AT-TA) with the same A-T content shows nearly identical behavior to random-sequence genomic DNA—no condensation observed. We further performed all-atom molecular dynamics (MD) simulations to elucidate the molecular origins of attraction. The MD results reproduce not only DNA sequence-dependent condensation but the quantitative force-distance relationship between juxtaposed helices measured experimentally with the osmotic stress method (OSM). The main difference between AA-TT and AT-TA is shown to be additional cation binding in the major grooves of AA-TT that is exposed to the DNA interface. This represents a new mechanism of sequence-directed DNA condensation, where consecutive adenine bases enhance cation groove binding via their continuous chain of partial charges (N6 and N7), and the exposed cations are spatially correlated with phosphate groups from neighboring DNA to elicit attraction.

MATERIALS AND METHODS

Experimental method

The AA-TT homopolymeric duplex (repeating A-T base pairs) was prepared by annealing equi-molar mixture of single-stranded deoxynucleotides poly(A) and poly(T) in 1×TE (10 mM Tris 1 mM ethylenediaminetetraacetic acid (EDTA) at pH 7.5) buffer. Poly(A) and poly(T) were purchased from GE Healthcare Life Sciences in the form of lyophilized sodium salts. Both polynucleotides have nominal average lengths of ∼300 bases and are highly monodisperse (see Supplementary Figure S1a). To ensure mixing at equi-molar ratio, we carried out annealing at a series of mixing ratios and confirmed maximum hypochromicity at 1:1 poly(A):poly(T) molar ratio. A simple annealing protocol of 5-min heating at 94°C followed by room-temperature cooling led to substantial polydispersity in the AA-TT duplex (see Supplementary Figure S1b), presumably due to the homopolymeric nature of poly(A) and poly(T). We found a sequence of 5-min heating at 94°C, 12-h (overnight) agitation at 45°C (6°C below its melting temperature in 1×TE buffer) and room-temperature cooling resulted in a rather monodisperse AA-TT duplex, evident from its well-defined band shown in Supplementary Figure S1b. The 1×TE annealing buffer was chosen to give a low melting temperature (∼51°C) to avoid prolonged agitation at high temperatures. This also eliminated the possibility to form triple-stranded helices. The AT-TA duplex (alternating A-T and T-A base pairs) was purchased from GE Healthcare Life Sciences in duplex form. The as-received construct showed to be rather poly-disperse, and a similar annealing protocol was followed to obtain high monodispersity (Supplementary Figure S1c).

Genomic DNA from Salmon Testes is used as the model random sequence DNA (GNOM duplex) and was purchased from Sigma in fibrous form. The fiber was first dissolved in 1×TE buffer and then dialyzed against an excess amount of 1 M NaCl solution to remove potential contaminants. It was then re-dissolved in 1×TE buffer after ethanol precipitation. Salt chemicals (Tris, EDTA and chloride salts of alkaline earth metal elements and putrescine) were purchased from Sigma and used as received.

The centrifugal precipitation assay provides a facile method for characterizing DNA condensation. Briefly, a series of solution mixtures were made with constant DNA concentration (50 μg/ml) and varied divalent salt concentrations, and each mixture was incubated on the bench for 10 min after thorough vortexing and a pulse downspin. The mixtures were then centrifuged at 20 000 × *g* for 10 min and the supernatants were taken from the top to be measured with a UV-vis spectrometer.

The OSM measures DNA osmotic pressure as a function of DNA–DNA inter-axial distance. The method is particularly advantageous for condensed DNA arrays in ordered liquid crystalline phases, where DNA–DNA spacing can be determined within 0.1 Å precision by x-ray diffraction. In this method, the DNA array is bathed against a coexisting solution phase of osmolytes (polyethylene glycol 8000 Dalton, PEG8k, was used in this study). The osmolytes are excluded from the DNA phase and thus exert pressure on the DNA array, equivalent of applying me-

chanical pressure with a piston semi-permeable to ions and water (12). At phase equilibrium, DNA osmotic pressure equates the PEG8k osmotic pressure. Measurements at a series of PEG8k concentrations give the DNA force/pressure-spacing relationship.

Simulation method

To shed light on the sequence and concentration-dependent DNA–DNA interactions we employed MD simulation. Umbrella sampling simulations allowed computing the potential of mean force (PMF) for AA-TT (dA_{20}-dT_{20}) and AT-TA ($d(AT)_{10}$-$d(TA)_{10}$) sequences. We study these systems in the presence of Mg^{2+} ions. PMF was used to compare the osmotic pressure response of the two sequences. The concentration dependence of the assembly process was further explored by simulations performed at various ion concentrations. We studied 22, 60 and 750 mM free Mg^{2+} concentrations, which we will denote them as *low*, *mid* and *high* concentrations respectively.

Molecular simulations. Duplexes of DNA were placed in a cubic box parallel to the z-axis (Figure 2A). To mimic long DNA arrays in our experiment we set the z-axis of the simulation box to the length of the DNA strands. Periodic boundary conditions allowed extending the DNAs to infinite length. We used the AMBER99bsc0 force field parameters for DNA (13,14), TIP3P for water (15) and NBFIX (16) for ions. All simulations were carried out in constant-temperature/constant-pressure ensemble using the GROMACS 5.0.5 package (17). Details of the simulation method are summarized in Supplementary Data.

Potential of mean force calculations. The reaction coordinate was defined as the distance between the center of mass of the two DNA duplexes positioned orthogonal to the x-y plane (Figure 2A). Initial systems were prepared for the inter-DNA distances varied from 23 to 52 Å with 1 Å interval. Each system was equilibrated independently for 150 ns prior to the production runs. Details of the equilibration protocol are given in Supplementary Data. The last frame of each equilibrated trajectory was later used to initiate umbrella sampling simulations. Using harmonic restraints, with a force constant of 2000 kJ $mol^{-1}nm^{-2}$, we used 150 ns long simulations to estimate the free energy as a function of inter-DNA distance. The weighted histogram method (WHAM) (18) implemented in the GROMACS package was used to construct the free energy landscapes.

Computing the osmotic pressure of DNA arrays. We made quantitative comparisons with experimental data by directly computing osmotic pressure. Based on the assumption that inter-DNA forces in a DNA aggregate are additive (19) and DNAs are hexagonally packed, the osmotic pressure can be calculated from the inter-DNA forces as:

$$\Pi = f(d)\sqrt{(3)}/dh$$

where d is mean distance between the two nearest neighbors in the DNA array, and $f(d)$ is mean force dsDNA per unit length, estimated by fitting the PMF to a double exponential function following Ref. (20).

Ion density analysis. From equilibrium simulations, we computed cation distributions. We determined specific ion binding sites and estimated the effective charge on the DNA surface after condensation. Finally, we calculated the number of bridging ions between duplexes. Details of these analyses can be found in Supplementary Data.

RESULTS

Divalent alkaline earth metal ions such as Mg^{2+} are known to effectively screen electrostatic interactions but have not been observed to condense double stranded helices of nucleic acids. Indeed, the AT-TA duplex with 100% AT content exhibits the same behavior as the random sequence dsDNA (GNOM) with no condensation at all Mg^{2+} concentrations up to 2 M, as shown in Figure 1A. However, the AA-TT duplex starts precipitating at ∼18 mM Mg^{2+} (Figure 1A), signifying the first observation of dsDNA condensation by alkaline earth metal cations. At an initial DNA concentration of ∼50 μg/ml (∼0.25 μM AA-TT duplex or ∼0.15 mM mono-nucleotide), AA-TT condensation is virtually complete within a narrow window of ∼2 mM Mg^{2+}, leaving no detectable amount of AA-TT in the supernatant. Consistent with being driven by electrostatic interactions, Mg^{2+}-induced AA-TT condensation can be reversed by lowering [Mg^{2+}], i.e. adding low-salt buffer to the precipitate re-dissolves it. Adding Mg^{2+} condenses it again. Rather curiously, it is also possible to re-dissolve the AA-TT condensate by raising [Mg^{2+}] above 750 mM, reminiscent of re-entrant behaviors of polyamine-induced condensation of dsDNA (3). In comparison, the AT#T triplex (i.e. the AA-TT duplex plus a third poly(T) strand) studied previously by us (5) is condensed by Mg^{2+} around 7 mM and remains condensed up to 2 M Mg^{2+}. Furthermore, the AA-TT condensate gives a pronounced x-ray diffraction (XRD) peak, evidencing high degree of structural order. Its inter-axial spacing of ∼28.7 Å at 20 mM Mg^{2+} is very close to that from the liquid crystalline dsDNA phases condensed by $Cobalt^{3+}$ hexamine (27.8 Å) or $spermine^{4+}$ (28.4 Å) and is significantly smaller than the ∼29.8 Å from Mg^{2+}-condensed AT#T triplex. Therefore, these observations rule out the possibility of Mg^{2+}-induced duplex-triplex transition (i.e. AA-TT to AT#T) as the cause of AA-TT condensation, because the transition would have led to, (i) free poly(A) single strands in the supernatant (rather than complete condensation), (ii) persistent condensation up to 2 M Mg^{2+} (rather than re-dissolution above 750 mM Mg^{2+}) and (iii) similar inter-axial spacings in the AA-TT and AT#T condensates (rather than a closer spacing of AA-TT due to its smaller diameter).

Mg^{2+}-induced condensation of AA-TT duplex is qualitatively different from previous studies of DNA condensation. Stronger inter-dsDNA attractions have been reported for AT-rich duplexes compared with random or GC-rich sequences (9,10), but the multivalent cations examined therein ($spermine^{4+}$ or $hex\text{-}lysine^{6+}$) also condense random-sequence dsDNA, unlike divalent alkaline earth metal ions in this study which do not condense dsDNA in general. DNA triplex condensation by Mg^{2+} has been observed (5). Rather than DNA sequence, the ∼50% difference in linear charge density between duplex and triplex is likely

the cause. Transition metal divalent cations such as Mn^{2+} or Cu^{2+} can condense DNA duplexes (21), but their interactions with DNA are complicated by specific binding. Consequently, the AA-TT duplex is the first duplex observed to be condensed by alkaline earth metal ions (e.g. Mg^{2+}) exhibiting non-ion-specific DNA binding only. While electrostatics is presumably the driving force for Mg^{2+}-induced AA-TT condensation, specific molecular characteristics of AA-TT and/or Mg^{2+} (e.g. charge, hydration or structure) are expected to be key to this peculiar condensation behavior.

We next examined whether divalent-cation-induced AA-TT condensation is unique to Mg^{2+}. Remarkably, all alkaline earth metal cations tested (Mg^{2+}, Ca^{2+}, Sr^{2+} and Ba^{2+}) condense AA-TT between 15 and 20 mM, as shown in Figure 1B. The re-dissolution behavior at high ionic concentrations is also observed for Ca^{2+} (∼800 mM) and Sr^{2+} (∼1500 mM), whereas this cannot be ruled out for Ba^{2+} for which very high concentrations were not explored due to its relatively low solubility. It is interesting to note that Ca^{2+} condenses AA-TT again above 1200 mM, while AA-TT is soluble in Mg^{2+} above ∼750 mM. On the other hand, a chain-shaped divalent cation, putrescine^{2+}, does not condense AA-TT up to 1500 mM (Figure 1B), indicating the requirement of point-charged ions for AA-TT condensation. By way of precaution, we verified that point-charged monovalent ions (Na^{+}) cannot condense AA-TT at any concentration (Supplementary Figure S2), and adding Na^{+} to Mg^{2+}-condensed AA-TT dissolves the pellet as expected from ion competition weakening attraction. These observations indicate that AA-TT condensation is a general phenomenon for alkaline earth metal ions but not for monovalent or chain-shaped divalent cations, emphasizing the important role of cation charge density and/or shape.

To probe the nature of Mg^{2+}-mediated interactions, we measured the inter-helix force as a function of inter-axial spacing with the OSM. OSM measures inter-molecular force via an osmotic pressure equilibrium between the condensed DNA array and the surrounding phase of excluded osmolytes, which exerts pressure on the array in the form of entropy-driven depletion force. Intuitively, the osmotic force/pressure is equivalent to the mechanical force required to pressurize the array through a semi-permeable membrane, and integrating the pressure-volume relation yields the free energy change or work done on the array. Figure 1C shows the force-spacing curves of AA-TT, AT-TA, GNOM duplexes in 20 mM Mg^{2+}, slightly above the critical concentration of ∼18 mM Mg^{2+} for AA-TT condensation. AT-TA and GNOM duplexes give nearly identical force curves, displaying the characteristics of repulsive inter-helix forces at all spacings, i.e. a slow decay of repulsion with spacing extended all the way to infinity. However, AA-TT exhibits a notably different behavior, evident from its zero pressure/force (free energy minimum) at the finite spacing of ∼28.7 Å, giving rise to a convex shape. The AA-TT curve qualitatively resembles the convex-shaped force curves of Mg^{2+}-condensed AT#T triplex and Cobalt^{3+} Hexammine (Co^{3+}Hex)-condensed GNOM dsDNA of random sequence (also shown in Figure 1C), with the largest difference being an offset in DNA–DNA spacing. In addition, heat induced contraction is observed (Supplementary Figure S3), suggesting the entropy-driven nature of condensation. The physical interpretation of inter-DNA forces within the last nanometer of surface contact has been attributed to either electrostatics or hydration as the predominant contributor, while both are expected to be involved and dependent on each other. Importantly, the force-spacing relation of Mg^{2+}-condensed AA-TT is consistent with previous measurements of multivalent cation condensed DNA helices, such as Co^{3+}Hex condensed dsDNA, affirming the interstitial cations (and hydration) as the mediators of inter-DNA attraction.

The quantitative nature of the measured forces also provides a valuable test for theoretical and numerical studies. The force-spacing dependency is often referred to as the equation of state for DNA osmotic pressure (22). It is thus a thermodynamic pressure related to the PMF. We used MD simulations to compute the PMF of DNA arrays and thereby osmotic pressure. Direct comparison between experiment and molecular simulations allows assessing the accuracy of the computational method and the choice of parameters. Elucidating the atomic and energetic details can then be possible from simulations. Figure 2B compares the osmotic pressure computed from simulation and the one measured by our experiments. Good agreement between the two approaches increases our confidence in using simulations to investigate the atomic details of the sequence specificity further. At low [Mg^{2+}] the pairwise interactions between the duplexes are repulsive (see Figure 2C), marked by the absence of a minimum at PMF.

The DNA condensation is more evident in AA-TT when we simulate the duplex pair at 60 mM (mid [Mg^{2+}]). At this concentration the free energy shows a minimum of −0.13 kcal/mol per turn (Figure 2D). The minimum is located at $d \approx 29$ Å, consistent with the inter-helical distance measured in experiment (Figure 1C). The magnitude of attraction aligns well with the previously measured values of DNA–DNA attraction in the presence of multivalent cations work (23). To further benchmark the concentration effect in high [Mg^{2+}] (Figure 1), we repeated the simulations, this time at 750 mM $MgCl_2$ solution. The free energy curve is shown in Figure 2E, exhibiting a local minimum at $d \approx 29$ Å and a downhill change in energy favoring spontaneous re-dissolution -again in accord with the Figure 1A. Later, we contrasted the behavior of AA-TT with AT-TA. Figure 1C shows the dramatic difference between the two sequences. PMF results of AT-TA confirmed the repulsive nature of inter-DNA interaction at salt concentrations of low and mid [Mg^{2+}] (Figure 2F and G), again in excellent agreement with our experimental measurements. Note that the PMF is not the true free energy for the condensation process. One can estimate the free energy of the DNA condensation following Tolokh *et al*. study (24). In our experiments, the DNA duplex has a length of about 300 bp. For ordered DNA arrays the enthalpic contribution linearly grows with chain length while entropic contributions grow logarithmically. Therefore, parallel arrays can show condensation even with a small attractive interaction per base pair. Our value of attraction energy of $0.025 k_B T$/bp is in the same order of magnitude with previous experimental measurements (22,25).

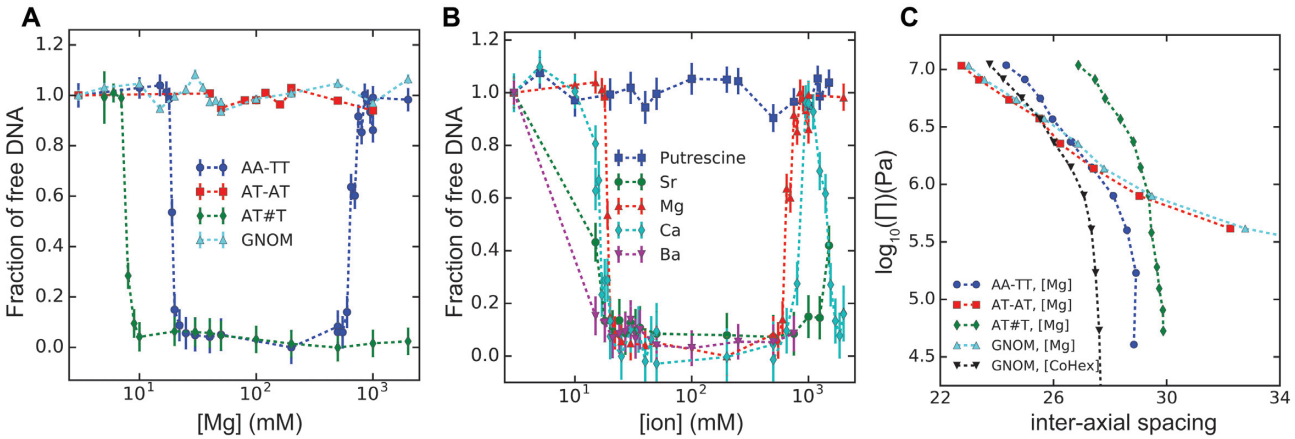

Figure 1. Precipitation assay and inter-helical forces of DNA by alkaline earth metal cations. (**A**) Mg^{2+} induced precipitation of different DNA constructs as denoted in the legend. The x-axis shows the concentration of ions in the solution and the y axis shows the amount of soluble DNA after centrifugation. (**B**) Condensation of AA-TT duplex by various divalent cations as denoted in the legend. (**C**) The force/pressure-spacing relationships between DNA helices obtained by the osmotic stress and x-ray diffraction methods. The type of DNA helix are given by the legend. The Mg^{2+} ion concentration was 20 mM and Co^{3+}Hex concentration was 1 mM. See main text for detailed discussions.

Assuming electrostatics as the driving force for Mg^{2+} induced AA-TT condensation, we looked at ion density profiles of the two sequences at mid [Mg^{2+}] (Figure 3). Details of each analysis are given in Supplementary Data. First, we projected the ion spatial distributions to the cylindrical axis of DNA denoted as $c(r)$. Figure 3A compares $c(r)$ for the two sequences at a far distance (d=36 Å). The $c(r)$ when the energy minimum (d=29 Å) is observed in Figure 3B. From the $c(r)$, a sequence dependent spatial distribution is evident. In AA-TT, Mg^{2+} ions accumulate into two regions separated by an apparent depletion zone. The first peak corresponds to major groove binding, while the second peak reflects the binding to phosphates. In contrast, AT-TA shows delocalized cation distribution; yet, the cumulative ion distributions show no marked differences between the sequences. In addition, both sequences attract more cations as the inter-DNA spacing decrease. Interestingly, some of the cations migrate from the major groove to the outer shell as the two DNAs approach each other in AA-TT. However, in AT-TA the cation atmosphere retains its structure (Figure 3A and B).

Figure 3C–E look into the cation distribution when projected onto the x axis and xy plane. A significant redistribution of ions is evident when the two helices come closer. At close distance (d=29 Å), Mg^{2+} ions accumulate at the interface, creating a cation rich region. We further divide the localized ions into two groups: (i) surface-bound ions and (ii) bridging ions (Figure 3F). The computation approach is detailed in the Supplementary Data. Here, we summarize our main findings. Our simulations suggest an increase in the surface-bound ions as the two DNA duplexes come closer (Figure 3G). The same trend is observed for low [Mg^{2+}] (Supplementary Figure S4a), yet we don't observe any sequence specificity. Unlike the surface-bound cations, the bridging ions show an excess accumulation in the case of AA-TT, resulting in a peak at the inter-DNA distance of about 29 Å (Figure 3H). In contrast, no peak is observed for AT-TA or AA-TT at low [Mg^{2+}] (Supplementary Figure S4b), suggesting a strong correlation between the existence of bridging ions and DNA condensation.

To investigate differences in the binding properties of cations that resulted in the two opposing behaviors between the two sequences, we computed radial distribution functions (RDF). RDF allowed quantifying the preferential binding of Mg^{2+} ions; a property that we found to be useful to elucidate the differences between the two sequences. We divided the DNA into three groups: minor and major groove atoms, and the backbone phosphates. Figure 4 shows our results. Phosphates condense Mg^{2+} ions non specifically. The minor grooves, while also exhibiting non-ion-specific binding, show much weaker binding. The reason for weak minor groove binding is simply volume exclusion: average width of minor groove is 4.09 ± 0.04 Å in AA-TT and it is 5.75 ± 0.01 Å in AT-TA. The effective diameter of a hexa-hydrated Mg^{2+} ion on the other hand, is about D≈8Å. The larger size of cation prohibits its direct binding to minor grooves. The major groove, on the other hand, is wide enough to accommodate cations (see details in Supplementary Figure S5). Major groove possesses high affinity to Mg^{2+} ions and it shows sequence specificity. We also looked at cation association as a function of inter-DNA spacing. Consistent with the surface bound ions (Figure 3G) the peak heights of the RDFs showed an increase as the inter-DNA spacings are reduced (Supplementary Figure S6).

The striking observation that the major groove binding shows strong sequence specificity leads us to look into pairwise interactions between Mg^{2+} ions and the atoms of major groove. We highlight the most important ones here. Simulations show that Mg^{2+} ions bind tightly to N7 and N6 atoms of adenine in AA-TT. Interestingly, the same atoms show weaker binding in the case of AT-TA (Figure 4D and E). The main reason for that lies in the unique positions of these atoms on the DNA surface as depicted in Figure 4F. In AT-TA, the two atoms are delocalized due

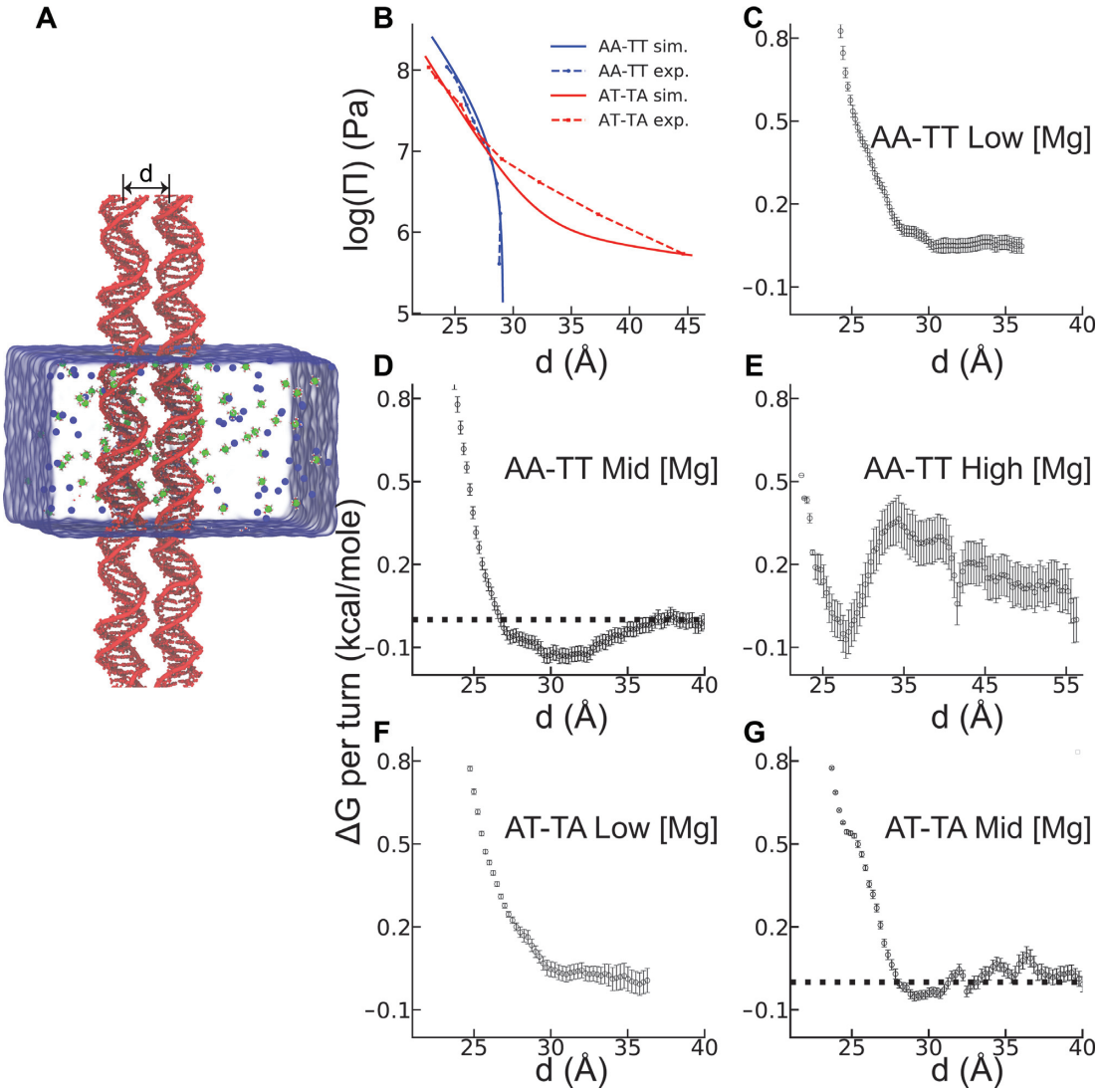

Figure 2. Simulation results for the inter-DNA interactions of the two sequences: AA-TT (dA_{20}-dT_{20}) and AT-TA ($d(AT)_{10}$-$d(TA)_{10}$). (**A**) Simulation box with the two DNAs positioned parallel with an inter-helical distance of *d*. Green and blue spheres represent hexa-hydrated $Mg^{2+}(H_2O)_6$ and Cl^- ions, respectively. Umbrella sampling simulations along the reaction coordinate *d* were used to compute the PMF between the two DNA helices. (**B**) Osmotic pressure as a function of inter-DNA spacing provides a direct comparison of the effective forces between DNA pairs with the OSM as denoted in legends (see 'Materials and Methods' section for details). Panels **C**–**G** show Gibbs free energy change as a function of the distance between DNA helices of the two sequences in different ionic conditions: (C) AA-TT at 22 mM [Mg^{2+}], (D) at 60 mM [Mg^{2+}] and (E) at 750 mM [Mg^{2+}]. (F) AT-TA at 22 mM [Mg^{2+}], and (G) AT-TA at 60 mM [Mg^{2+}]. The horizontal dashed lines are guide to eye. Statistical errors were estimated using bootstrap analysis with 100 number of bootstraps.

to alternating A and T nucleobases stacked at the groove. Alternating sequences resulted in low surface charge density at the major groove, affording weak cation associations. In the case of AA-TT, these atoms are stacked along the helix, creating high-charge density that leads to cation condensation.

As a result of extra cation condensation in AA-TT, the effective charge at the major groove increases (Supplementary Figure S7a) whereas the backbone and the minor groove still remain as negative after cation condensation. The effect of this surface charge pattern is reflected on the azimuthal shift between the two helices denoted as $\delta\phi$ here (see Supplementary Data). Interestingly, $\delta\phi$ shows sequence specificity. It localizes around 76° in AA-TT (Supplementary Figure S8) while staying at around 19° in AT-TA. The DNA orientation for both cases is shown in Supplementary Figure S9.

At last, we investigated the mechanism of re-dissolution observed in high Mg^{2+} ion concentrations (Figures 1 and 2E). We compute the surface charge on each DNA. Results are summarized in Supplementary Figure S7b. Simulations exhibit a positive net charge of +0.17e/base supporting the charge inversion mechanism. The ion-pairing between Mg^{2+} and Cl^- ions are strong at shorter distances; yet, it is not prevalent at longer separations (Supplementary Figure S10). To investigate further, we computed the net charges on each group. Results summarized in Supplementary Figure S7b demonstrate that at high salt the charge pattern diminishes, resulting in re-dissolution of the DNA pairs.

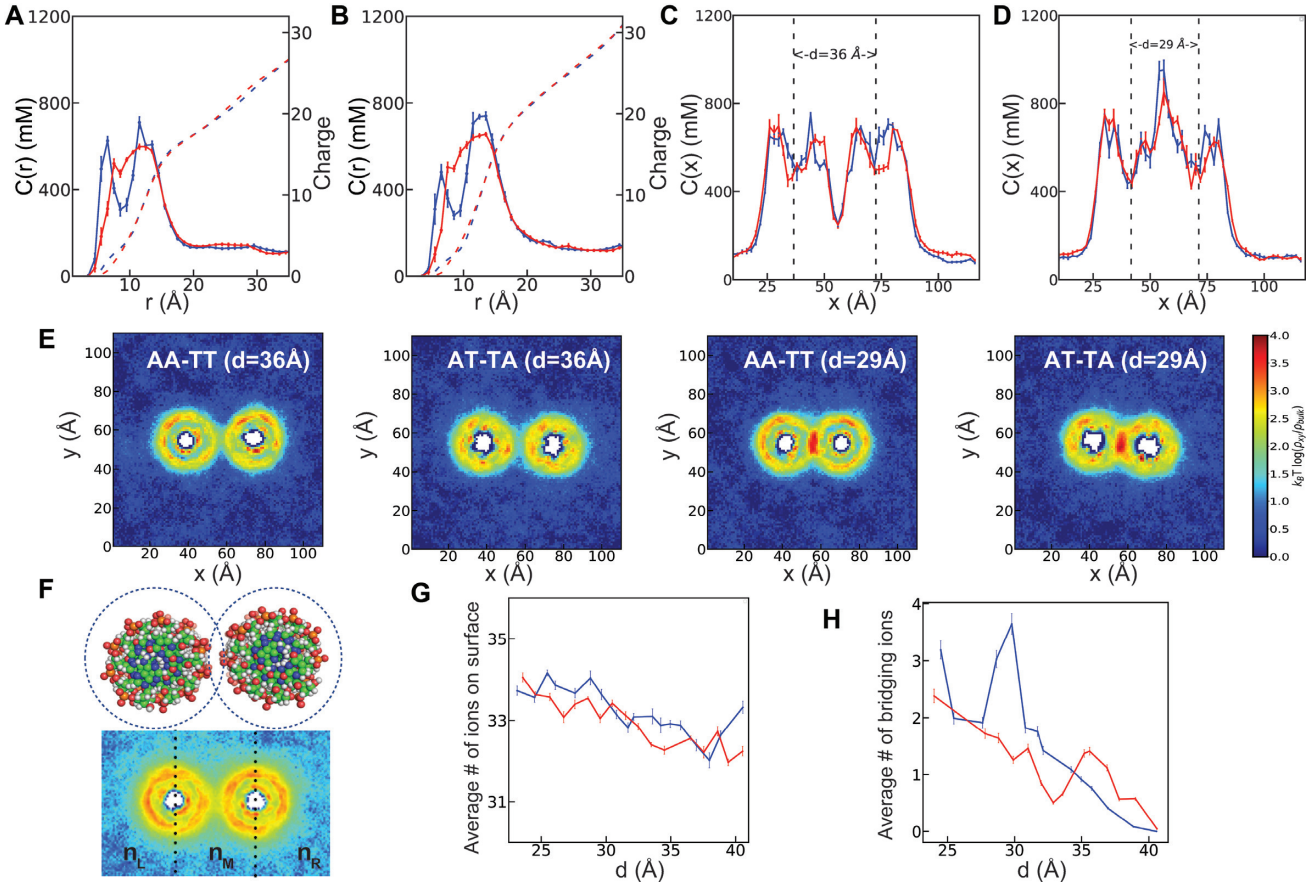

Figure 3. Mg^{2+} ion distribution around the DNA duplexes of AA-TT and AT-TA in 60 mM bulk $MgCl_2$ concentration. For panels **A**–**D** and **G** and **H**: The DNA duplex AA-TT is in blue color and AT-TA is in red color. (A) The Mg^{2+} ion distributions around the cylindrical axis of the DNA at d=36 Å (see Supplementary Data for details). (B) The same as in A, this time for d=29 Å. The Mg^{2+} ion density profiles are projected onto the Cartesian coordinate of x-axis. (C) The 1D concentration profile in inter helical distance of d=36 Å. (D) the same as C, this time for d=29 Å. The dotted lines are guide to eye showing the center of mass of each DNA. (**E**) The 2D concentration profile at inter helical distance d=36 Å and 29 Å. (**F**) Shown is the schematic description of the two binding modes. The number of the surface bound ions are computed by counting the Mg ions within 5 Å from the surface of the DNA surface (up) while the number of bridging ions, n_B is defined as the difference between the number of ions at the middle part of the simulation box separated by the dashed lines and the two sides, $n_B = n_M - (n_L + n_R)$ where n_M, n_L and n_R represent the ensemble average of the number of ions at the middle, left and right sides respectively (bottom). (G) Shows the total number of surface ions, and (H) the number of bridging ions, as a function of inter-helical separation.

DISCUSSION AND CONCLUSION

We reported sequence-dependent dsDNA condensation by divalent alkaline earth metal ions. Our data demonstrates that homopolymeric sequence of AA-TT shows condensation in the presence of divalent ions, an observation similar to tsDNA or genomic dsDNA in polyvalent conditions. Using experiments and simulations, we studied this condensation phenomenon in different salt conditions. The concentration dependence of condensation and the force-spacing dependency both provided excellent agreement between the two independent approaches. To shed light into the molecular mechanisms governing these interactions, we analyzed the computer simulations. Atomically detailed insights gained from simulations reveal cation adsorption specificity into the major grooves. This specificity is dictated by the DNA surface topology and favors angular correlation between helices.

Cation binding to major grooves results in a structural pattern of positive (the major groove) and negative (the phosphate backbone) charges along the helical axis. Attraction is mediated between two such DNA surfaces through an axial offset. This is reminiscent of the electrostatic zipper (*EZ*) model proposed for DNA condensation in poly-cations (19). The respective physical mechanisms are however qualitatively different: this study shows sequence-enhanced groove binding (SEGB) whereas the *EZ* model hypothesizes cation major groove binding *a priori* for condensation. Our simulations further show that the amount of cation binding in the major groove is small even for AA-TT, ∼+0.2e/base. A greater contribution comes from the polar groups of the major groove (e.g. C6, C4 etc.), totaling +0.53e/base. The combination gives rise to a highly charged major groove (Supplementary Figure S7a) along the side of negatively charged minor groove and phosphate backbone (−0.2, −0.8e/base, respectively). The requirement of relatively low level of cation binding for DNA condensation indicates that cation binding to the major groove can be a rather sensitive modulator. One example is that Mg^{2+} has been reported to compact λ phage DNA when pH is low-

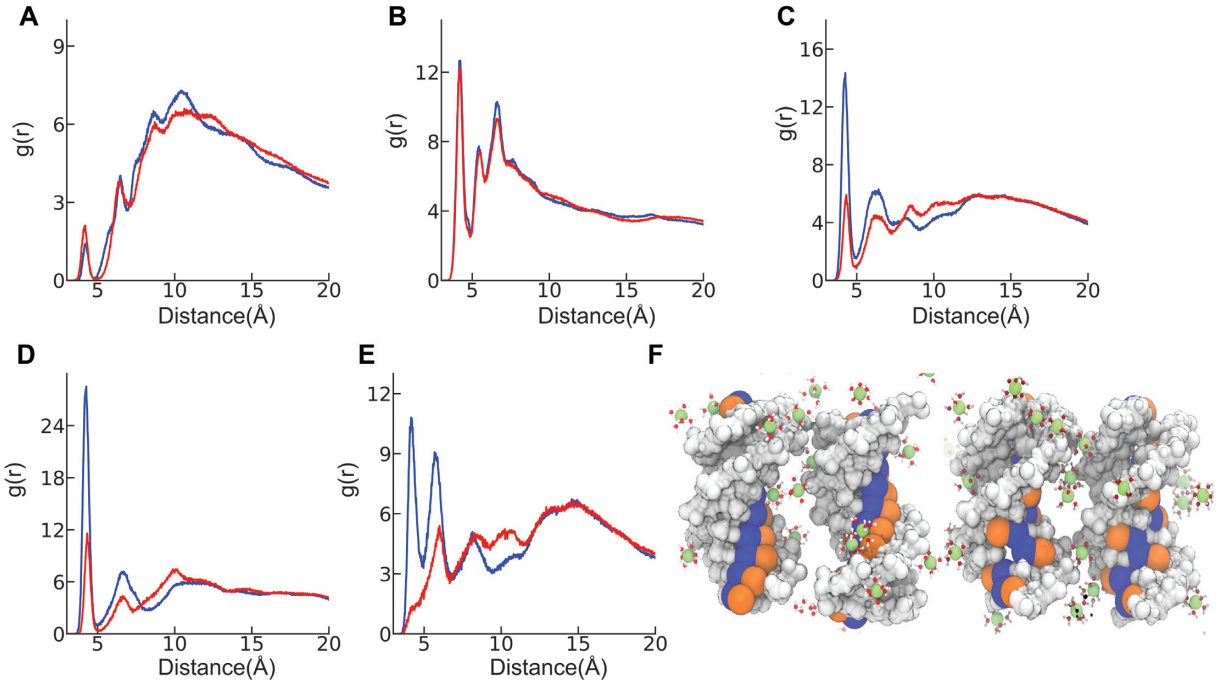

Figure 4. Cation RDF of AA-TT (blue) and AT-TA (red) in 60 mM $MgCl_2$ salt solution at inter-helical spacing of 29 Å. Panels A–E shows the RDF. (A) Minor groove (B) phosphate groups and (C) major groove. Further analysis of atoms at the major groove shows that the favorable binding is mainly due to two electronegative groups: (D) N7 atom that resides at adenine, and (E) N6 atoms at adenine. A more intricate picture of the role of these atoms in sequence specificity can be gleaned by looking at their spatial positioning on the DNA surface. (**F**) Illustrates how these two atoms are located at the surface of one turn of the DNA major groove in AA-TT (left), AT-TA (right). The two negatively charged atoms are colored orange (N7) and blue (N6). Hexa-hydrated Mg^{2+} ions are represented with ball and stick.

ered to ∼4 (26), which can be explained by H^+ association in the DNA major groove (e.g. the N7 atom of Adenine has a pKa of ∼3.5) giving the extra bit of positive charge.

Another notable consequence of the SEGB model for DNA condensation is the azimuthal shift δϕ in order to juxtapose the charge patterns of approaching DNA surfaces. The δϕ controls the distance between the major groove (positively charged) of one DNA to the phosphates of the adjacent one (negatively charged). $\delta\phi \approx 76^\circ$ observed in AA-TT gives rise to a shorter distance of (22.7 Å) for the two oppositely charged surfaces while $\delta\phi \approx 19^\circ$ in AT-TA leads to a further distance (23.7 Å), suggesting relatively weaker interactions for the same inter-DNA distance. Similar azimuthal shifts have been observed by X-ray scattering (27,28), theoretical study (29) and MD simulation (20). However, it has also been shown that azimuthal correlation is not necessary for DNA condensation by chain-shaped poly-cations which act as bridges enabled by their extended length (10). Such ion bridging cannot be achieved by alkaline earth metal ions studied here due to their small sizes, and the azimuthal shift is important for bringing the cations closer to the phosphate backbone of apposing helix.

We also observed both experimentally and through simulation that the attractive forces diminish at high salt concentrations. Various mechanisms have been proposed to explain the electrostatic interactions at high concentrations; charge reversal, and ion-pairing are among those (30–34). Unlike the attractive regime where opposing charges form an undulating positive and negative charges on the surface, in high salt condition our simulation suggest a relatively uniform charge pattern. We believe that this uniform charge pattern leads to the repulsion observed in both approaches.

Simulations (9–10,20), previous experiments (5,22–23) and our study show no evidence of DNA condensation by monovalent cations alone. The reasons behind this observation are likely that the monovalent cations have weak major groove binding (35). In addition to their weak binding, their low valence can not afford to create a positive charge pattern on the DNA surface backbone. The opposite charge pattern on the surface of the DNA leads to attraction. Hence, we surmise no condensation for monovalent cations. We would like to further point out that the mechanism by which DNA condensation occurs depends on the cation type. For chain-shaped polyvalent cations a dynamic bridging, extending between the two adjacent phosphates, serves to stabilize the complex (9,10). For point charge divalent cations, which is the focus of the current study, an alternating charge pattern along the axis leads to condensation. The charge pattern forms upon binding of the divalents to the major groove. As upon binding the major grooves became positively charged due to valence $z = 2$. Specific binding to minor groove can also create a charge pattern which in principle attracts favorably the backbone of the adjacent pair. Further studies are underway to explore the mechanism of DNA attraction in transition metal ions such as Mn^{2+} ion which is known to bind preferentially to the minor grooves in A-tracks (35).

Our data demonstrate that the condensation mechanism can be extended to alkaline earth metal ions for specifically ordered DNA sequences. Here we report the condensation in AA-TT but the observations are not limited to this se-

quence. The lessons learned from this study can lead to design sequences capable of leading to controlled assembly in divalent cations. We show that DNA duplex interactions can be engineered by sequence variability. Specific binding to grooves offers great opportunity to develop novel design strategies for DNA nanotechnology that primarily relies on hybridization. The abundance of divalent ions in vivo and the tunability of these interactions with concentration and sequence variability offers new avenues and directions to biotechnology.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Computational research was carried out on the High Performance Computing resources at New York University Abu Dhabi. S.K would like to thank to Dr Rudi Podgornik, and Dr Biman Bagchi for insightful discussions.

FUNDING

National Science Foundation [MCB-1616337 to X.Q.]; New York University Abu Dhabi faculty research grant [AD181 to S.K.]. Funding for open access charge: New York University (to S.K.).
Conflict of interest statement. None declared.

REFERENCES

1. Lipfert,J., Doniach,S., Das,R. and Herschalag,D. (2014) Understanding nucleic acid-ion interaction. *Annu. Rev. Biochem.*, **83**, 813–841.
2. Kornyshev,A.A., Lee,D.J., Leikin,S. and Wynveen,A. (2007) Structure and interaction of biological helices. *Rev. Mod. Phys.*, **79**, 943–960.
3. Pelta,J., Livolant,F. and Sikorav,J.L. (1996) DNA aggregation induced by polyamines and cobalthexamine. *J. Biol. Chem.*, **271**, 5656–5662.
4. Bloomfield,V.A. (1997) DNA condensation by multivalent cations. *Biopolymers*, **44**, 269–282.
5. Qiu,X., Parsegian,V.A. and Rau,D.C. (2010) Divalent counterion-induced condensation of triple-strand DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 21482–21486.
6. Katz,A.M., Tolokh,I.S., Pabit,S.A., Baker,N., Onufriev,A.V. and Pollack,L. (2017) Spermine condenses DNA, but not RNA duplexes. *Biophys. J.*, **112**, 22–20.
7. Tolokh,I.S., Pabit,S.A., Katz,A.M., Chen,Y., Drozdetski,A., Baker,N., Pollack,L. and Onufriev,A.V. (2014) Why double-stranded RNA resists condensation. *Nucleic Acids Res.*, **42**, 10823–10831.
8. Li,L., Pabit,S.A., Meisburger,S.P. and Pollack,L. (2014) Double-stranded RNA resists condensation. *Phys. Rev. Lett.*, **106**, 108101–108105.
9. Kang,H., Yoo,J., Sohn,B.K., Lee,S.W., Lee,H.S., Ma,W., Kee,J.-M., Aksimentiev,A. and Kim,H. (2018) Sequence-dependent DNA condensation as a driving force of DNA phase separation. *Nucleic Acids Res.*, **46**, 9401–9413.
10. Yoo,J., Kim,H., Aksimentiev,A. and Ha,T. (2016) Direct evidence for sequence-dependent attraction between double-stranded DNA controlled by methylation. *Nat. Commun.*, **7**, 11045.
11. Hud,N.V. and Engelhart,A.E. (2009) Sequence-specific DNA-Metal ion interactions. In: Hud,N.V. (ed). *Nucleic Acid-Metal Ion Interactions*, The Royal Society of Chemistry, UK, pp. 75–117.
12. Parsegian,V.A., Rand,R.P. and Rau,D.C. (1995) Macromolecules and water: probing with osmotic stress. *Methods Enzymol.*, **259**, 43–94.
13. Pérez,A., Marchán,I., Svozil,D., Sponer,J., Cheatham,T.E. III, Laughton,C.A. and Orozco,M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description α/γ conformers. *Biophys. J.*, **92**, 3817–3829.
14. Cornell,W.D., Cieplak,P., Bayly,I., Gould,I.R., Merz,K.M., Ferguson,D.M., Spellmeyer,D.C., Fox,T., Caldwell,J.W. and Kollman,P.A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, **117**, 5179–5197.
15. Jorgensen,W.L., Chandrasekhar,J., Madura,J.D., Impey,R.W. and Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926–935.
16. Yoo,J. and Aksimentiev,A. (2012) Improved parametrization of Li^+, Na^+, K^+, and Mg^{2+} ions for all-atom molecular dynamics simulations of nucleic acid systems. *J. Phys. Chem. Lett.*, **3**, 45–50.
17. Hess,B., Kutzner,C., van der Spoel,D. and Lindahl,E. (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.*, **4**, 435–447.
18. Kumar,S., Rosenberg,J.M., Bouzida,D., Swendsen,R.H. and Kollman,P.A. (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. *J. Comput. Chem.*, **13**, 1011–1021.
19. Kornyshev,A.A. and Leikin,S. (1999) Electrostatic Zipper Motif for DNA Aggregation. *Phys. Rev. Lett.*, **82**, 4138–4141.
20. Yoo,J. and Aksimentiev,A. (2016) The structure and intermolecular forces of DNA condensates. *Nucleic Acids Res.*, **44**, 2036–2046.
21. Rau,D.C. and Parsegian,V.A. (1992) Direct measurement of temperature-dependent solvation forces between DNA double helices. *Biophys. J.*, **61**, 260–271.
22. Rau,D.C. and Parsegian,V.A. (1992) Direct measurement of the intermolecular forces between counterion-condensed DNA double helices. Evidence for long range attractive hydration forces. *Biophys. J.*, **61**, 246–259.
23. Todd,B.A. and Rau,D.C. (2008) Interplay of ion binding and attraction in DNA condensed by multivalent cations. *Nucleic Acids Res.*, **36**, 501–510.
24. Tolokh,I.S., Drozdetski,A.V., Pollack,L., Baker,N.A. and Onufriev,A.V. (2016) Multi-shell model of ion-induced nucleic acid condensation. *J. Chem. Phys.*, **144**, 155101.
25. Baumann,C.G., Smith,S.B., Bloomfield,V.A. and Bustamante,C. (1997) Ionic effect on the elasticity of the single DNA molecules. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 6185–6190.
26. Wang,Y., Gao,T., Li,S., Xia,W., Zhang,W. and Yang,G. (2019) Direct demonstration of DNA compaction mediated by divalent counterions. *J. Phys. Chem. B*, **123**, 79–85.
27. Strey,H.H., Wang,J., Podgornik,R., Rupprecht,A., Yu,L., Parsegian,V.A. and Sirota,E.B. (2000) Refusing to twist: demonstration of a line hexatic phase in DNA liquid crystals. *Phys. Rev. Lett.*, **3**, 3105–3108.
28. Leforestler,A. and Livolant,F. (2009) Structure of toroidal DNA collapsed inside the phage capsid. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 9157–9162.
29. Kornyshev,A.A. and Leikin,S. (1997) Theory of interaction between helical molecules. *J. Chem. Phys.*, **107**, 3656–3674.
30. Yang,J. and Rau,D.C. (2005) Incomplete ion dissociation underlies the weakened attraction between DNA helices at high spermidine concentrations. *Biophys. J.*, **89**, 1932–1940.
31. Besteman,K., Van Ejik,K. and Lemay,S.G. (2007) Charge inversion accompanies DNA condensation by multivalent ions. *Nat. Phys.*, **3**, 641–644.
32. Wang,Y., Wang,R., Cao,B., Guo,Z. and Yang,G. (2016) Single molecular demonstration of modulating charge inversion of DNA. *Sci. Rep.*, **6**, 38628.
33. Luan,B. and Aksimentiev,A. (2010) Electric and electrophoretic inversion of the DNA charge in multivalent electrolytes. *Soft. Matter*, **6**, 243–246.
34. Grosberg,A.Y., Nguyen,T.T. and Shklovskii,B.I. (2002) Colloquium: the physics of charge inversion in chemical and biological systems. *Rev. Mod. Phys.*, **74**, 329–345.
35. Atzori,A., Liggi,S., Laaksonen,A., Porcu,M., Lyubartsev,A.P., Saba,G. and Mocci,F. (2016) Base sequence specificity of counterion binding to DNA: what can MD simulations tell us? *Candian J. Chem.*, **94**, 1181–1188.