DNA conformational flexibility study using phosphate backbone neutralization model

ARTICLE *in* SOFT MATTER · NOVEMBER 2013

Impact Factor: 4.03 · DOI: 10.1039/c3sm52345d

4 AUTHORS:

Shiyan Xiao
University of Science and Technology of Ch…
16 PUBLICATIONS **57** CITATIONS
SEE PROFILE

Hong Zhu
University of Science and Technology of Ch…
7 PUBLICATIONS **51** CITATIONS
SEE PROFILE

Lei Wang
University of Science and Technology of Ch…
10 PUBLICATIONS **71** CITATIONS
SEE PROFILE

Haojun Liang
University of Science and Technology of Ch…
152 PUBLICATIONS **2,215** CITATIONS
SEE PROFILE

Soft Matter

PAPER

View Article Online
View Journal | View Issue

DNA conformational flexibility study using phosphate backbone neutralization model†

Cite this: *Soft Matter*, 2014, **10**, 1045

Shiyan Xiao,[a] Hong Zhu,[a] Lei Wang[a] and Haojun Liang*[bc]

Received 5th September 2013
Accepted 12th November 2013

DOI: 10.1039/c3sm52345d

www.rsc.org/softmatter

Due to the critical role of DNA in the processes of the cell cycle, the structural and physicochemical properties of DNA have long been of concern. In the present work, the effect of interplay between the DNA duplex and metal ions in solution on the DNA structure and conformational flexibility is studied by comparing the structure and dynamic conformational behavior of a duplex in a normal form and its "null isomer" using molecular dynamics methods. It was found that the phosphate neutralization changes the cation atmosphere around the DNA duplex greatly, increases the major groove width, decreases the minor groove width, and reduces the global bending direction preference. We also noted that the probability of BI phosphate linkages increases significantly because of the charge reduction in the backbone phosphate groups. More importantly, we found that the electrostatic effect on the DNA conformational flexibility is dependent on the sequence; that is, the phosphate backbone neutralization induces the global dynamic bending to be less flexible for GC-rich sequences but more flexible for AT-rich sequences.

1 Introduction

Due to its functions in encoding the genetic blueprint of an entire living organism and carrying instructions for protein synthesis, DNA is one of the most important biological molecules, and the knowledge of DNA's structural and physicochemical properties is indispensable in understanding its role in the processes of the cell cycle.[1] From the perspective of DNA being considered as a polyelectrolyte, the enormous electrostatic repulsive interactions between the negatively charged phosphoryl groups of the nucleic acid backbone and the interaction between base pairs keep DNA a highly ordered and rigid molecule.[2] Under a cellular environment, counterions in solution, mostly monovalent and divalent cations, will construct a condensed ion layer surrounding the molecule, usually called an "ion atmosphere", which not only plays an important role in neutralizing the net charge of the nucleic acid and reducing the repulsion through electrostatic screening, but also affects the structural stability and functional dynamics of DNA,[3–6] such as the processes of nucleic acid packaging,[7,8] DNA–protein interactions and so on.[9]

Metal cations can modulate DNA structure and flexibility in a sequence-dependent manner,[10–13] thus spurring tremendous interest and studies aimed at the understanding of the specific DNA base–cation interactions.[14] The groove width, as well as the spatial arrangement of electronegative groups on the bases, correlates very well with sites that are identified as being particularly favorable for cation localization.[15] Based on the experimental results of magnetic relaxation dispersion,[16] X-ray diffraction methods[17] and capillary electrophoresis,[18] Denisov, Halle and other researchers found that monovalent alkali cations prefer to bind to the minor groove of AT rich sequences, especially A-tract DNA. Evidence from molecular dynamics (MD) simulations by Wilson *et al.* and Feig *et al.* support that the entrance of Na^+ into the minor groove induces groove narrowing,[19,20] while McConnell and Beveridge reported that Na^+ entry into the minor groove has little effect on the groove width.[21] In the major groove of the X-ray structure of B-form DNA, Na^+ has been identified but is poorly ordered.[22] Using MD simulations, for the Dickerson–Drew dodecamer (DDD) and duplex $(A_5G_5)\cdot(T_5C_5)$, Na^+ is observed along the guanine bases with an occupancy of 10–20% and causes a slight narrowing of the major groove,[20,21] and K^+ is in contact with the O6 and N7 of guanine and the phosphate groups of the $d(CG)_{12}$ oligomer, and binds essentially to GpC and not to CpG steps.[23] Similar phenomena have also been observed by Beveridge *et al.* in the modeling of DDD, but the correlation between ion proximity and DNA major and minor groove widths was found to be negligibly small.[24] Divalent cations, such as Mg^{2+} and Ca^{2+}, are

[a]*CAS Key Laboratory of Soft Matter Chemistry, Department of Polymer Science and Engineering, University of Science and Technology of China, Hefei, Anhui 230026, P. R. China*

[b]*CAS Key Laboratory of Soft Matter Chemistry, Collaborative Innovation Center of Chemistry for Energy Materials, Department of Polymer Science and Engineering, University of Science and Technology of China, Hefei, Anhui 230026, P. R. China. E-mail: hjliang@ustc.edu.cn; Fax: +86-551-6360-7824; Tel: +86-551-6360-7824*

[c]*Hefei National Laboratory for Physical Sciences at Microscale, University of Science and Technology of China, Hefei, Anhui, 230026, P. R. China*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c3sm52345d

much easier to distinguish from alkali cations and well ordered water molecules due to the regular octahedral geometry of their primary solvation shell in experiments.[15,25] The interaction of divalent cations with the minor groove is mediated through the hydrated water molecules, thus they are rarely found deep within the minor groove because of size and the high energetic cost required to adjust the size and the geometry of the hydration shell,[14,25,26] although there are some exceptions.[27] Ca^{2+} has been observed to locate fully hydrated in the minor groove, whereas Mg^{2+} resides on top of the minor groove of an adjacent DNA duplex[27] or bridges phosphate groups across the minor groove in the crystal of DDD.[28] A strong correlation between the position of bound divalent cations and the minor groove width is observed, which indicates that minor groove narrowing can be caused by the divalent cations on top of the groove but not those bound deep in the groove.[15,29–31] At the major groove, Ca^{2+} is found to make direct contact with the guanine at the GpG step, and Mg^{2+} is found at the GpG or GpC step but interacts with DNA through hydrogen bonding interactions between its hydrated water ligands and the guanine base.[30,32,33] In some other cases, Ca^{2+} at the major groove does not make direct contact with the DNA bases, for example a fully hydrated Ca^{2+} was observed by Kielkopf *et al.* to locate at the ApG step of CCAGTACTGG.[29] It is clear that divalent cations prefer to bind to the GpN and ApG steps,[15] and will significantly alter the local DNA structure, such as compressing the major groove.[15,30] More direct evidence for the effect of metal ions on the DNA conformational flexibility comes from the nucleic acid structural bending study of Tan *et al.*[34] Using the tightly bound ion theory to calculate the electrostatic free energy landscapes for DNA helix bending, Tan *et al.* demonstrated that the DNA bending free energy is strongly dependent on the ion concentration, valency, and size.[34] Furthermore, Li *et al.* proved that the binding of Mg^{2+} onto DNA makes DNA more rigid compared to that of Na^{+} based on the conformational entropy calculations for DNA in different ion concentration environments.[32]

Apart from the above discussions, the conformational state of the phosphate linkage is another crucial factor for the DNA structure. In addition to the canonical BI state, the BII state is a frequent feature of B-DNA,[35,36] and the reliance on the BI ↔ BII equilibrium receives special interest because of its role in the specific recognition of DNA by the cognate proteins.[37–41] Based on the ^{31}P NMR results on the JunFos oligomer, Heddi *et al.* demonstrated that BI/BII populations are affected by the type of monovalent cation, and the BII conformations would tend to be stabilized by cations in the major groove, or destabilized by cations in the minor grooves.[14] For the $C_{i-1}p_{i-1}A_{i-1}p_iG_i$ fragment of the same oligomer, it is proved that Mg^{2+} interacting with A_{i-1} and G_i increases the proportion of the p_{i-1} BII population.[42] A clear relationship between BI/BII populations and the nucleic acid sequence has been identified by Dršata *et al.*, which predicts that the CG steps have the highest BII percentage.[43]

Although extensive studies on the effect of metal ions on the B-DNA structure and dynamics have been performed, the molecular basis of such possible effects of cations on DNA remains far from well-characterized and understood.[14,42] It has been proven that the counterions can influence DNA structure by neutralizing the phosphate charge,[44] however, until recently, there has still been no consensus for which effect, the electrostatic repulsion from the backbone or the base pair stacking compression, contributes most to the DNA rigidity.[3,45–50] How much do the negatively charged phosphate groups affect the DNA base pair stacking? The main problem in studying the effect of DNA–ion interactions on the DNA properties is to separate the cause from the overall effect. As a powerful tool in studying biological processes, MD simulations have been widely applied in understanding the sequence specific localization and counterion motions of metal ions around the DNA duplex in atomic detail.[9,23,51–53] In the present work, in order to gain deeper insight into the effect of the interplay between DNA and metal ions in solution on the DNA flexibility, we followed the idea of constructing the artificial "null isomer" by reducing the charge along the DNA backbone[44,45,54] and employed MD simulations to study the effect of DNA–ion interactions on the DNA structure and conformational flexibility by comparing the behavior of the normal DNA duplex and its "null isomer".

2 Models and simulation details

2.1 Modeled systems

To quantitatively investigate the fundamental issue of electrostatic self-repulsion of the DNA backbone and the effect of DNA–cation interactions on the DNA structure and flexibility properties, Manning proposed the notion of a "null isomer" for DNA, *i.e.*, a hypothetical structure which would be adopted by the polyelectrolyte chain if the electrostatic charge on it were set to zero (without affecting the solvation). In this way, we can eliminate the electrostatic repulsive effect on the polyelectrolyte structure among its ionized groups.[49] One can achieve the "null isomer" by reducing the charges along the DNA backbone, or introducing site-specific neutralizing chemical substitutions of phosphate oxygens.[44,45,49,55] Due to the stereospecific effects of phosphate modification, the incorporation of neutral *S*-methylphosphonate groups into the major groove will effectively reduce the major groove width while leaving the minor groove width basically unaltered, and the external orientation of *R*-methylphosphonate groups produce quite the opposite effect, increasing the major groove width and decreasing the minor groove width.[44] Furthermore, the calculations of Kosikov *et al.* demonstrated that all-*R* or all-*S* methylphosphonate substitution of the phosphate oxygens produce different magnitudes and directions of bending from the uniform electrostatic neutralization of the phosphate charges. Therefore, in order to avoid the effect of stereochemistry of these phosphate substitutions, we obtained the neutralization of a given phosphate group by adding +1.00 *e* to them, including one phosphorus (P) and four oxygen (O1P, O2P, O3′, O5′) atoms, leaving the partial charges of all the other atoms of the nucleotides unchanged. Rather than evenly dividing and adding the net charge to the atoms of the phosphate group, we reduced the partial charges of these atoms by the same proportion (28.8206%) in order to reach a zero total net charge for each residue. In this model, we did not reoptimize the force field parameters or redistribute the

partial charges of the nucleotides using *ab initio* calculations, because the present work aims at investigating the effect of the interplay between DNA (especially the backbone) and the metal ions in solution on the structure and dynamic flexibility of the DNA oligomer; additionally, force field reoptimization and charge redistribution may alter the base pair stacking and the interactions between cations and the atoms located at the groove. See Fig. S1, ESI† for more details about the methodology of the neutralization, and the partial charges of the phosphate group atoms at the 5′, 3′ and central residues of DNA duplex are listed in Table S1, ESI.†

In order to study the sequence specific ion–DNA interactions, two 20-bp DNA duplexes with different sequences were employed in the present work: (1) $d(GC)_{10}\cdot d(CG)_{10}$, and (2) $d(AT)_{10}\cdot d(TA)_{10}$. For each sequence, the normally charged (charge-full), and the neutralized (charge-reduced) DNA systems have been simulated. The 5′ and 3′ ends of each DNA strand were terminated by hydroxyl groups. Each DNA system was solvated in a $\sim 80 \times 80 \times 120$ Å^3 water box composed of ~15 000 water molecules plus 150 mM (cation molar concentration) NaCl or $MgCl_2$. The system contents of all the simulation systems are listed in Table 1.

2.2 Molecular dynamics

The initial coordinates of DNA models were of the B-form and were generated using the program named *nucgen*, which comes with AMBER.[56] The modeled systems were prepared using the program *xleap* in the AMBER suite.[56] The latest parmbsc0 force field[57] was used for DNA, the Åaqvist's parameters[58] were used for Na^+, Cl^-, and Mg^{2+}, and the TIP3P model[59] was used for the water molecules.

All simulations were performed with the NAMD program,[60] and the VMD software package was employed for trajectory visualization.[61] The SHAKE/RATTLE algorithm[62] was used to constrain the bonds involving hydrogen atoms in both the solvent and the solute. Multiple time step integration was carried out using r-RESPA,[63] with a base time step of 2 fs and a secondary time step of 4 fs for long-range interactions. Periodic boundary conditions were applied and the Particle Mesh Ewald (PME) method[64] was employed to evaluate the long-range electrostatic interactions. The systems were simulated in the NPT ensemble, with a Nosé–Hoover thermostat[65] and a Langevin piston Nosé–Hoover barostat to maintain the temperature at 300 K and the pressure at 1 atm, respectively. Each system was simulated for 150 ns and the last 100 ns trajectory was used for analysis.

Table 1 List of simulation systems and their abbreviations

System	Charge	Sequence	Ions (150 mM)
GCC-Na	Charge-full	$d(GC)_{10}\cdot d(CG)_{10}$	NaCl
GCN-Na	Charge-reduced	$d(GC)_{10}\cdot d(CG)_{10}$	NaCl
GCC-Mg	Charge-full	$d(GC)_{10}\cdot d(CG)_{10}$	$MgCl_2$
GCN-Mg	Charge-reduced	$d(GC)_{10}\cdot d(CG)_{10}$	$MgCl_2$
ATC-Na	Charge-full	$d(AT)_{10}\cdot d(TA)_{10}$	NaCl
ATN-Na	Charge-reduced	$d(AT)_{10}\cdot d(TA)_{10}$	NaCl
ATC-Mg	Charge-full	$d(AT)_{10}\cdot d(TA)_{10}$	$MgCl_2$
ATN-Mg	Charge-reduced	$d(AT)_{10}\cdot d(TA)_{10}$	$MgCl_2$

3 Results and discussion

3.1 Counterion cloud around dsDNA

Root mean square deviations (RMSD). As has been widely discussed,[66–70] the force field used here can accurately describe DNA conformations with a wide variety of relative base pair positions and the interactions between base pairs without additional modification, thus is able to provide stable trajectories for DNA. The convergence of a single trajectory can be verified by the small RMSD and the high relative similarity between the first and second equivalent block of the same trajectory. In addition, we calculated the RMSD and its standard deviation for both the whole DNA duplex and its backbone (Table 2). It is shown that the RMSD standard deviation of the backbone is slightly larger (by 0.05–0.1 Å) than that of the whole DNA duplex, which indicates that the backbone is somewhat more flexible than the base pairs that are buried in the central helix, protected by the outside casing constructed by phosphate and sugar groups, and reinforced internally by strong stacking forces. As we have seen, for the charge-full form of DNA duplex, the flexibility of $d(AT)_{10}\cdot d(TA)_{10}$ is reduced in the presence of Mg^{2+} ion compared to Na^+, with a decrease of ~0.2 Å for the standard deviation of RMSD. Whereas the situation is different for $d(GC)_{10}\cdot d(CG)_{10}$; that is, the standard deviations of RMSD are almost the same for GCC-Na and GCC-Mg. This phenomenon is different from the observations of Li *et al.*,[32] who performed principal component analysis (PCA) and conformational entropy calculations for an isolated random sequence double-stranded 23-mer DNA in NaCl and $MgCl_2$ electrolytes and stated that the global motion (or conformational flexibility) of Mg-DNA is reduced compared with that of Na-DNA.[32] Our observations show that the effect of cations on the conformational flexibility is sequence dependent, probably due to the differences in binding mode for cations at poly GC segments and poly AT segments.

For the backbone of the charge-reduced form of the DNA duplex, a 0.1–0.2 Å decrease in RMSD standard deviation has been observed for $d(GC)_{10}\cdot d(CG)_{10}$ in both Na^+ and Mg^{2+} electrolytes compared to the charge-full form, but there are little differences between these two forms of $d(AT)_{10}\cdot d(TA)_{10}$. This behavior indicates that the interaction between the cations and negatively charged phosphate groups contributes more to the DNA global structural fluctuation for dGC than for dAT. These phenomena can be explained by the properties of cation distribution around DNA. As demonstrated by Li *et al.*,[32] dGC and dAT have the same total neutralization fraction values in the presence of Na^+, while the occupancy percentage of Na^+ is higher at the backbone of dGC than dAT; and in the Mg^{2+} electrolyte, more Mg^{2+} accumulates around dGC than dAT. Because of the charge reduction of phosphate groups, the lower cation occupancy percentage may help to weaken the interplay between the DNA backbone and cations in solution, therefore

Table 2 MD-averaged (with the standard deviation in parentheses) root mean square deviation (RMSD), the major groove width, minor groove width and the global bending angle of $d(GC)_{10}\cdot d(CG)_{10}$ and $d(AT)_{10}\cdot d(TA)_{10}$ in Na^+ and Mg^{2+} electrolytes. It should be noted that the standard deviations for major and minor groove widths of each isolated DNA duplex correspond to the deviation between MD simulation-averaged groove widths of designated base pairs

	MD-RMSD (DNA) (Å)	MD-RMSD (backbone) (Å)	Major groove (Å)	Minor groove (Å)	Bending angle[a] (°)
GCC-Na	3.13(0.84)	3.47(0.90)	18.74(0.24)	13.10(0.17)	16.58(0.26)
GCN-Na	2.51(0.68)	2.81(0.71)	18.76(0.21)	12.97(0.16)	14.29(0.32)
GCC-Mg	3.06(0.83)	3.44(0.90)	18.53(0.36)	13.32(0.35)	10.09(0.35)
GCN-Mg	2.90(0.73)	3.28(0.77)	18.68(0.20)	12.85(0.23)	9.79(0.32)
ATC-Na	3.37(1.01)	3.75(1.05)	20.83(0.14)	13.35(0.15)	16.31(0.32)
ATN-Na	3.09(0.97)	3.43(1.06)	20.66(0.07)	11.43(0.21)	16.44(0.39)
ATC-Mg	3.22(0.80)	3.64(0.89)	20.54(0.19)	13.51(0.34)	17.88(0.33)
ATN-Mg	3.17(0.98)	3.38(0.89)	20.71(0.11)	11.66(0.27)	19.79(0.29)

[a] The global bending angle is calculated for the central twelve base pair oligonucleotide.

reducing the structural fluctuation of nucleic acids. Obviously, this effect will be more significant for dGC compared to dAT.

The ion atmosphere. In order to investigate the overall shape of the ion atmosphere around the DNA duplex, we plot the radial distribution function (RDF) of the cation as a function of the distance from the DNA axis (Fig. 1) and the cation density map around the DNA (Fig. S2, ESI†). The reference axis system located on each base is defined according to the methodology proposed by Curuksu *et al.*,[71] with its *z*-axis perpendicular to the base plane. The DNA axis is defined as the sum of the *z*-axis unit vector of each nucleotide for the central ten base pair portion. The RDF calculations are performed for the central ten base pair portion in order to avoid the end effects of the DNA oligonucleotide.

As can be seen, all ion radial distribution profiles plotted in Fig. 1 reach 1.0 for ion–DNA distances greater than 20 Å. The plateau at large ion–DNA distances suggests that our simulations reached equilibrium. In the outer layer of the ion atmosphere, the radial profiles of cations drop rapidly as the ion–DNA distance exceeds 15 Å. More interestingly, the radial distribution profiles decay faster for Mg^{2+} than for Na^+ for all of the investigated systems, indicating the stronger binding of Mg^{2+} to DNA compared with that of Na^+. A similar phenomenon has also been observed by Yoo *et al.*[33] In the inner layer of the ion atmosphere for the charge-full form of the two nucleic acid sequences investigated, there are several small peaks but no distinct sharp peak can be recognized in the radial distribution profiles of Na^+, especially for $d(GC)_{10}\cdot d(CG)_{10}$. We attribute this phenomenon to the closely distributed electronegative atoms which act as the binding sites for cations. Additionally, the interactions between the ions and binding sites are not sufficient to withstand the thermal fluctuation of the hydrated monovalent cations, and thus no specific cation-binding sites exist in either of the grooves.[33,72] This is why the density maps of Na^+ are broadly distributed along the grooves (see Fig. S2, ESI†). In contrast, the peaks of the ion distribution curve of Mg^{2+} are sharp and the heights of these peaks are twice as high as those of Na^+, suggesting the stronger electrostatic interaction between DNA and Mg^{2+} due to its larger net charge. It is noted that there is only one peak located at an ion–DNA distance of ~12.0 Å for the ATC-Mg system, whereas two distinct peaks are observed at ~5.0 Å and ~12.0 Å for the GCC-Mg system, with their heights reduced to approximately two-thirds that of ATC-Mg. This behavior indicates that few Mg^{2+} can penetrate into the grooves of ApT and TpA steps and maintain stability, while most Mg^{2+} ions near to the DNA bridge the phosphate oxygen atoms of the opposite stand or locate on top of the grooves. The behavior of Mg^{2+} around the $d(AT)_{10}\cdot d(TA)_{10}$ is similar to that of Mg^{2+} at the A-tract region, in which Mg^{2+} is almost absent at the minor and major grooves.[32] For $d(GC)_{10}\cdot d(CG)_{10}$, Mg^{2+} can penetrate deep into the major grooves and make direct and quite stable contact with the DNA bases at the N7 and O6 of guanine, N4 of cytosine, and even bridge the O6 guanines of two consecutive GC/CG base pairs in some segments; at the backbone, they reside on top of the grooves or bridge the phosphate groups, but no Mg^{2+} ion has been observed to penetrate deep into the minor groove (see

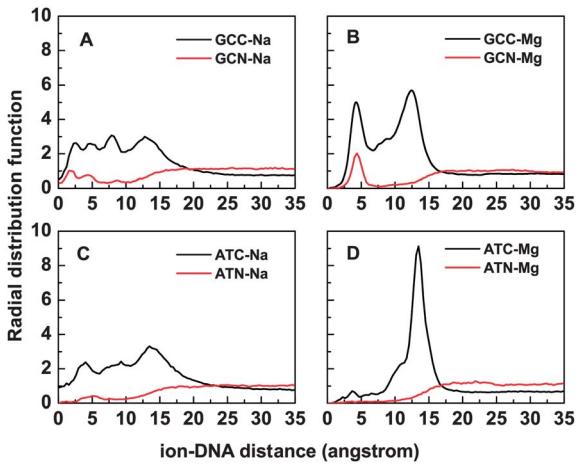

Fig. 1 Radial distribution of metal ions as a function of the distance from the DNA axis, for the charge-full form and charge-reduced form of $(GC)_{10}\cdot d(CG)_{10}$ (GCC, GCN) and $d(AT)_{10}\cdot d(TA)_{10}$ (ATC, ATN) in Na^+ (A and C) and Mg^{2+} (B and D) electrolytes, respectively. The calculations are performed for the central ten base pair portion.

Fig. S2, ESI†). Our results presented here are in accordance with the previous experimental studies in solution which state that at least one guanine is necessary for a dinucleotide step to localize a divalent cation in the major groove and that interactions with GC-rich sequences can involve inner shell coordinates, whereas interactions with AT-rich sequences do not.[15,73,74]

For the charge-reduced form of the DNA duplex, it is shown that the ion radial distributions are featureless in either the inner or the outer layer region. For the inner layer region, the RDF value is pretty small compared to that of the corresponding charge-full state; and in the region of ion–DNA distances larger than 15 Å, the ions are almost free with no obvious binding behavior. This behavior of cations is also confirmed by the ion density map in Fig. S2, ESI.† This indicates that there are few cations penetrating deep into the grooves and making stable connections with the phosphate groups on the backbone because of the reduced charges of the phosphate group. We speculate that the highly negatively charged phosphate groups on the backbone play an important role in helping the cations in solution to move close then penetrate deep into the DNA grooves. The behavior difference in the ion atmosphere between the charge-full and charge-reduced DNA can also be understood from Manning's counterion condensation theory. For charge-full duplexes, which have larger values of the Manning parameter (denoted as ζ, $\zeta = l_B/b$, where l_B is the Bjerrum length and $1/b$ represents the line density of fixed charges along the DNA strand. $\zeta = 4.2$ for normal DNA duplex) compared to their corresponding charge-reduced states, it is reasonable to observe a higher density of metal ions around the DNA. Whereas the value of $1/b$ approaches 0 in the "null isomer" state, which means a smaller value of ζ thus a smaller RDF value in the inner layer region around the DNA duplex. It is also interesting to note that a small peak exists at ~5.0 Å in the radial distribution profile for the GCN-Mg system, with some hydrated Mg^{2+} ions bound at electronegative sites in the major groove of the GC sequence (bridging the O6 atoms of guanines of consecutive GC/CG base pairs, see Fig. S2, ESI†). This phenomenon again proves the presence of guanine bases is important for Mg^{2+} binding onto the major groove, and to some extent, helps to keep Mg^{2+} in this groove.

3.2 Structural properties

Groove width. P_i and P_i' are defined as the 5′-phosphates of the complementary nucleotides that comprise bases pair i, with the prime denoting the complementary stand (*i.e.*, strand 3′ → 5′). The major and minor groove widths are calculated as follows: the distance of $P'_{i-2}\cdots P_{i+2}$ across the minor groove with three intervening base pairs is the minor groove width, and the distance of $P_{i-2}\cdots P'_{i+2}$ across the major groove with four intervening base pairs is the major groove width.[44] In the calculations, the first three and the last four base pairs are ignored due to their high mobility.

The average groove width values along with their corresponding standard deviations are displayed in Table 2. It should be noted that the standard deviation for major or minor groove widths of each isolated DNA duplex correspond to the deviation between MD simulation-averaged groove widths of designated base pairs. It means that the standard deviation values reflect differences of groove widths caused by the interaction between the metal ions and the DNA duplex. Except for the minor groove width of ATN-Na, the standard deviation of groove width is decreased in the charge-reduced form compared to the corresponding charge-full form in either $MgCl_2$ or NaCl electrolyte. This behavior indicates that the structural fluctuation of the nucleic acid duplex is reduced because of the charge reduction of backbone phosphates, and this is in accordance with the conclusion derived from the RMSD data.

It is shown that the charge reduction of backbone phosphates increases the major groove width and decreases the minor groove width, with only one exception: the major groove of $d(AT)_{10}\cdot d(TA)_{10}$ in Na^+ (see Table 2 and Fig. S4, ESI†). The observations presented are in line with the simulation of DNA duplex whose phosphates are substituted by *R*-methylphosphonates.[44,75] For the charge-full form of the DNA duplex, the minor groove width is around 13 Å, and the electrostatic repulsive interactions between the phosphate groups are strong. Previous studies show that the entrance of monovalent cations (mostly, Na^+) into the minor groove, or Na^+ and Mg^{2+} on top of the minor groove will cause groove narrowing because of the compensation of positive charges of metal ions to the high negative charges of phosphate groups.[19,20] Thus, for the minor groove, we attribute the decrease in groove widths in the charge-reduced form of the DNA duplex to the weakened repulsive interactions between the phosphates on the DNA backbone, due to the phosphate neutralization. On the contrary, for the major groove of the charge-full form of the DNA duplex, metal ions localize on top of or penetrate deep into the groove, thus help to compress the groove width by attracting the phosphate groups on the backbone. This correlation has been verified by the experimental results which state that the CpG step of DDD is found to be narrowed by the binding of Na^+ or Mg^{2+} or Ca^{2+}. This effect is greater for divalent ions than monovalent ions, that is, a more significant compression of the major groove can be observed for Mg^{2+} and Ca^{2+} ions, which have higher net charges.[21,29,30] For the charge-reduced nucleic acid duplex, the weakened attractive interactions between the phosphates and cations on the major groove help less to compress the major groove, thus we see a widened major groove width compared to the charge-full form of DNA, with only the one exception of $d(AT)_{10}\cdot d(TA)_{10}$ in Na^+.

BII state. The major BI and the minor BII backbone phosphate substates are determined by the difference of ε and ζ dihedral angles ($\varepsilon - \zeta < 0$ for BI, and $\varepsilon - \zeta > 0$ for BII). The equilibrium between BI ↔ BII backbone phosphate conformers is examined because of its potential role in high-amplitude sugar–backbone motions and protein–DNA recognition. A geometrical illustration of BI and II conformers is presented in Fig. S9, ESI.†

Overall, for the charge-full form duplexes, the BII population of $d(GC)_{10}\cdot d(CG)_{10}$ is higher (20–40%) than that of $d(AT)_{10}\cdot d(TA)_{10}$ (~10%) in either NaCl or $MgCl_2$ electrolyte (see Fig. 2 and 3). These results are in accordance with the NMR results of Nikolova *et al.*,[76] who stated that the BII conformer

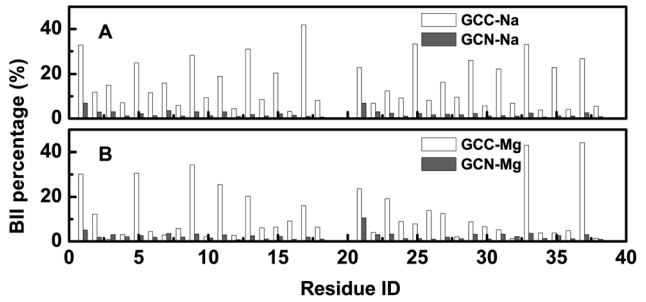

Fig. 2 Populations of the BII substate for charge-full and charge-reduced form of $d(GC)_{10}\cdot d(CG)_{10}$ in Na^+ (A) and Mg^{2+} (B) electrolytes.

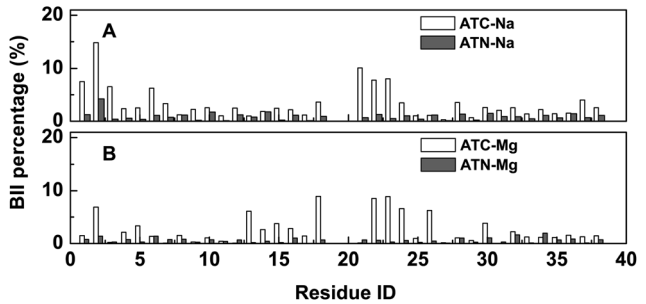

Fig. 3 Populations of the BII substate for charge-full and charge-reduced form of $d(AT)_{10}\cdot d(TA)_{10}$ in Na^+ (A) and Mg^{2+} (B) electrolytes.

occurs most frequently in CpG steps but least within A-tracts. According to the previous studies, the BII conformations would tend to be stabilized by cations in the major grooves, but destabilized by cations in the minor groove.[14] From the results and discussions above, the minor groove of ApT is a better candidate to be occupied by monovalent cations,[15,25,77] thus it is rational to see a lower BII population compared to GpC step in NaCl electrolyte. Meanwhile, GpC is known to be good at accommodating divalent cations in its major grooves but nearly no divalent cations penetrate into the minor groove, thus the presence of Mg^{2+} increases the percentage of BII conformers. In contrast, the major groove of ApT is less favorable for Mg^{2+}, implying a low percentage of BII conformers. It is also interesting to observe that there is a huge difference (~10%) for the BII populations between two successive base pairs, especially for GpCpG, which possesses a high BII population overall. Our observations confirm the conclusion derived by Guéroult *et al.*, stating that the interactions between the nucleic base pair step and hydrated Mg^{2+} are characterized by BI phosphate linkage but increase the BII population of the phosphate of its neighboring base pair step.[42]

For the charge-reduced form duplexes, it is observed that the BII population in all the systems is less than 5%, indicating that the BI phosphate linkages increase greatly if the charge of backbone phosphate groups decreases. The increase in proportion of BI conformer is attributed to the low cation occupancies in the grooves of DNA helices. Thus these behaviors are rational because fewer cations in major grooves implies smaller population of BII conformers. We also analyzed in detail the torsional angles of the phosphate backbone (see Fig. S5–S8, ESI† for the population profiles of α, β, γ, δ, ε and ζ). Detailed comparison of the population plots for all DNA sequences in Na^+ and Mg^{2+} shows that charge-full and charge-reduced duplexes present maxima at identical positions for α, β and γ, and the distributions of δ in charge-reduced forms are slightly narrower compared to their corresponding charge-full states. More apparent differences between the charge-full and charge-reduced states have been observed on ε and ζ. It is seen that the population of ε is slightly increased around 180° and −180°, but decreased slightly at the local maximum located around −90° due to the backbone phosphate neutralization. Moreover, the distribution of ζ in the charge-reduced states is slightly narrower at ~−90°, and increased evidently at ~90°. In conclusion, the phosphate neutralization does not significantly affect the backbone torsional angles, except ε and ζ, *i.e.*, the great increase in the BI phosphate linkage.

DNA bending. The global bending angle (θ) and distribution of ϕ for $d(GC)_{10}\cdot d(CG)_{10}$ and $d(AT)_{10}\cdot d(TA)_{10}$ are shown in Table 2 and Fig. 4, respectively. From Table 2, it is known that the MD-averaged global bending angles calculated from our simulations for these two nucleic acid duplexes are around 20°, which is within the bending angle range determined using experimental[78] and MD methods.[32,70] It is noted that the instinct global bending of the DNA duplex is reduced in the presence of Mg^{2+} compared to Na^+ for $d(GC)_{10}\cdot d(CG)_{10}$ (by ~5°), but enhanced for $d(AT)_{10}\cdot d(TA)_{10}$ (by ~2°). These behaviours confirm the experimental results which state that divalent ions have strong and differential effects on various sequence elements.[79,80] For the bending direction, we can see that both sequences under study tend to bend towards the central backbone in either Na^+ or Mg^{2+} electrolyte. We attribute this bending direction preference to the strong interactions between the negatively charged backbone phosphates and cations, thus helping to attract the phosphate groups and induce the bending

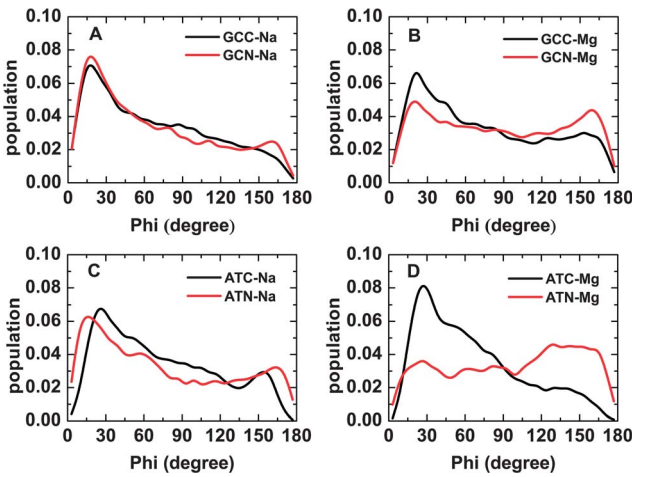

Fig. 4 The population distribution of the angle phi (ϕ) for $d(GC)_{10}\cdot d(CG)_{10}$ and $d(AT)_{10}\cdot d(TA)_{10}$. ϕ is calculated for the central sixteen base pair portion. (1) $\phi \rightarrow 0°$: DNA bending points towards the backbone of strand 1; (2) $\phi \rightarrow 180°$: DNA bending points towards the backbone of strand 2.

deformation towards the backbone. This is different from the observations that GC and AT sequences would like to bend towards the major groove reported by Curuksu *et al.* based on the restrained MD simulations for DNA duplexes.[71,81] From the iso-probability surface for directional DNA bending for $(AT)_6$ sequence,[71,81] comparable probability is observed for the phosphate direction as the groove directions. As the interactions between the DNA backbone and nucleosome contribute most to the DNA deformation upon the nucleosome, it is rational to see that the DNA duplex bends towards the backbone phosphate here.

For the charge-reduced form DNA duplex, the bending deformation is slightly reduced compared to its corresponding charge-full form for the GC-rich sequence, but slightly enhanced for the AT-rich sequence. Opposite behavior in the global bending of these two sequences has also been indicated by the experiments performed by Strauss *et al.*[80] They demonstrated that the bending of $5'$-G_3C_3 is reduced, whereas that of $5'$-A_6 tract is enhanced, if both are asymmetrically tethered by hexamethylene chains near specific phosphates. Nevertheless, the global bending is not greatly changed under the situation of backbone charge reduction, thus probably, the weak bending deformation ($\theta < 80°$ (ref. 71)) is dominated by the base pairing interactions rather than the electrostatic repulsion between phosphates on the DNA backbone. From the population distribution of angle ϕ, it is seen that the bending direction preference towards the phosphates are reduced with the reduction of backbone charges. It is believed that the phosphate neutralization significantly decreases the interactions between metal ions and the backbone phosphates, thus helps to reduce the bending direction preference. We may suspect that the negatively charged phosphates play some role in alternating the DNA duplex bending towards a particular direction.

Besides the structural properties discussed above, we also calculated the helical parameters, shown in Table S2, ESI.† It is noted that only slight changes have occurred for $d(GC)_{10}\cdot d(CG)_{10}$ in both Mg^{2+} and Na^{2+} electrolytes, indicating that the phosphate backbone neutralization slightly affects the base pair stacking in the GC sequence. In contrast to $d(GC)_{10}\cdot d(CG)_{10}$, more significant changes have been observed in $d(AT)_{10}\cdot d(TA)_{10}$, such as, the twist increases by $\sim$2.0°, whereas the twist increases are less than 0.5° in GC sequence systems. Overall, the charge reduction of backbone phosphates little affects the base pair stacking. This phenomenon is attributed to the fact that base pairs are buried in the central helix and reinforced internally by strong stacking forces. DNA twisting is known to be strongly correlated to other degrees of freedom, such as slide, roll and rise. It was proved by Olson *et al.*[82,83] and others[84] from crystal structures that an increase of twist is therefore generally accompanied by a decrease in roll, and an increase in slide and rise for a normal-form DNA duplex. From our simulation, it is shown (Table S2, ESI†) that the increase of twist for a charge-reduced duplex compared to its charge-full states is accompanied by a decrease of roll in AT sequences, but an increase of roll in GC sequences, and tiny changes for slide and rise. The difference correlation rules between charge-full and charge-reduced states are probably due to decreased geometric restrictions caused by the sugar–phosphate backbone to preserve its geometry because of the backbone phosphate neutralization.

3.3 Conformational flexibility of dsDNA in equilibrium

PCA analysis. PCA analysis was performed for all simulated systems in Cartesian coordinate space using their MD trajectories, and the magnitudes of the first 10 eigenvalues are displayed in Fig. 5, with the eigenvalues of the following 55 eigenvectors (from 11 to 65) shown in the insert. The global flexibility of the DNA duplex is dominated by the untwisting and bending movements which are represented in the first few essential deformation modes or eigenvectors of DNA (eigenvector index from 1 to 10), and the quantities of the corresponding eigenvalues reflect the motion amplitude of a certain degree of freedom.[32,85] For the charge-full form DNA, it is seen that the eigenvalues of the first few essential motions are larger for both $d(GC)_{10}\cdot d(CG)_{10}$ and $d(AT)_{10}\cdot d(TA)_{10}$ in Na^+ than in Mg^{2+}, indicating that the harmonic force constant of global motion is larger for duplexes in Mg^{2+} than in Na^+, thus the dynamic conformational flexibility of DNA is reduced due to the stronger interactions between Mg^{2+} and nucleic acids. This conclusion is in accordance with the result calculated using the CHARMM parameters.[86] It is also noted that the eigenvalues of global deformation motions of the GC sequence are slightly smaller than those of the AT sequence, which demonstrates that the AT sequence is more flexible than the GC sequence. This phenomenon proves that the DNA flexibility is sequence dependent, and also validates the experimental results[87] using the method of computational simulations.

From Fig. 5, it is interesting to note that the effect of charge reduction of DNA backbone phosphates on the global essential motions is also sequence dependent. The eigenvalues of the global deformation modes are slightly smaller for

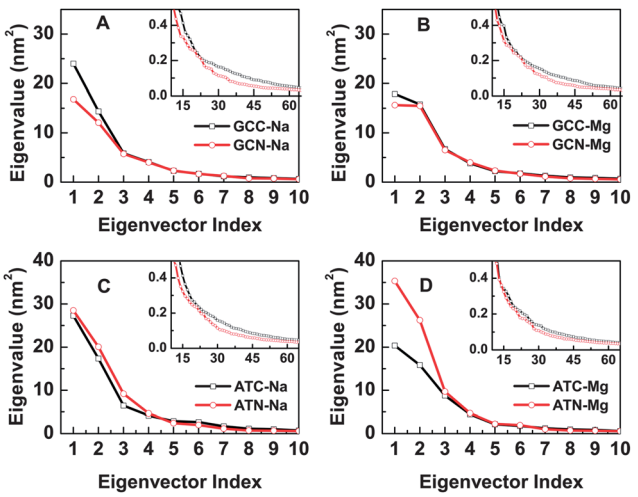

Fig. 5 Eigenvalues for the first 10 eigenvectors from principle component analysis (PCA). A and B are $d(GC)_{10}\cdot d(CG)_{10}$ in Na^+ and Mg^{2+} electrolytes, respectively; C and D correspond to $d(AT)_{10}\cdot d(TA)_{10}$. The inserts are eigenvalues of the following 55 eigenvectors (from 10 to 65) of corresponding isolated DNA duplexes.

$d(GC)_{10}\cdot d(CG)_{10}$ in the charge-reduced form compared to its charge-full form, whereas the situation is reversed for $d(AT)_{10}\cdot d(TA)_{10}$. These results demonstrate that the phosphate backbone neutralization induces the global dynamic bending to be less flexible for the GC sequence but more flexible for the AT sequence. Evidently, our simulations show that the electrostatic effect on the dynamic conformational flexibility is sequence dependent.

To date, the effect of electrostatic repulsion on the DNA dynamic bending remains controversial. The previous theoretical estimates of this contribution range from increased flexibility,[88,89] through a negligible effect[90,91] on the flexibility, to reduced flexibility.[92] Recently, Savelyev *et al.* tried to investigate to what extent electrostatics and nonelectrostatics (*i.e.*, the compressive base pair stacking) contribute to DNA's rigidity by measuring the persistence length of full charged and hypothetical uncharged coarse-grained (CG) nucleic acid duplexes using explicit ions but implicit water solvation models.[45] They gave the results that the phosphate charge reduction softens the DNA duplex, but they also pointed out that sequence specific modeling and further experiments are critically needed to address the electrostatic and nonelectrostatic contribution portions to DNA's rigidity because they used a sequence averaged DNA CG model. Okonogi *et al.* stated that phosphate backbone neutralization increases DNA flexibility through experiments for a random sequence but no specific sequence oligomer.[93] In their experiments, the phosphate backbone neutralization is achieved by substituting phosphates at six asymmetric specific sites with neutral methylphosphates, which may result in a stereochemical effect on the ion atmosphere around the nucleic acids, according to the statements of Olson and Manning *et al.*[44] Using transient electric birefringence analysis, Hagerman *et al.* concluded that the hemi-charged meroduplex had a persistence length approaching that of duplex DNA, and suggested that the stacking interaction predominates over interphosphate repulsions in determining persistence length.[94,95] In combination with previous studies, our present work heightens the importance of further experimental and theoretical investigation into how the electrostatic repulsive interaction on the DNA backbone and the sequence dependency affect the DNA flexibility.

Conformational entropy. The conformational entropies of the whole DNA duplex, the backbone combined with the sugar rings, and bases in all the simulated systems are calculated for the central sixteen nucleotide duplex portion at a time step of 5 ns, shown in Table 4. For charge-full form, DNA in NaCl electrolyte has a higher entropy value compared to that in $MgCl_2$ electrolyte. More specifically, the entropy of the GC sequence (*i.e.*, base pairs, sugar rings and phosphate groups in total) in NaCl electrolyte is approximately 15 kcal mol^{-1} larger than that in $MgCl_2$ electrolyte at room temperature, while this value is approximately 30 kcal mol^{-1} for the AT sequence. Clearly, the binding of Mg^{2+} on DNA makes it more rigid compared to that of Na^+, consistent with the conclusion demonstrated by the PCA calculations. Additionally, the entropy difference of DNA duplexes between the NaCl system and the $MgCl_2$ system obviously confirms the conclusions derived in the *ion atmosphere* section, which elucidated that divalent Mg^{2+} has a preference for GC-rich sequences, whereas AT-rich sequences are less attractive for Mg^{2+} compared to Na^+. Since the reduced electrostatic screening will increase the repulsive interactions on the DNA backbone, greater change in the electrostatic effect can be achieved by the phosphate neutralization for $d(AT)_{10}\cdot d(TA)_{10}$ due to its lower occupation compared to $d(GC)_{10}\cdot d(CG)_{10}$, thus a greater entropy decrease has been observed in the AT sequence.

It is shown from Table 4 that the entropy of the DNA duplex in the charge-reduced form is ~50 kcal mol^{-1} smaller than that in charge-full form, indicating that the phosphate backbone neutralization decreases the DNA flexibility. This result is different from the conclusion suggested by the global motion calculated by the PCA method. Nevertheless, we do not think these two conclusions are incompatible with each other. This is because the entropy values are calculated based on the consideration of all the essential motion modes rather than the first few eigenvalues representing the global dynamics bending and untwisting. In fact, the eigenvalues of the local essential deformation modes (eigenvectors from 10 to 55) possess slightly smaller values for charge-reduced DNA duplex compared to its charge-full form (see Fig. 5). It is suspected that the charge reduced backbone phosphate attracts fewer cations in solution and brings fewer fluctuations that come from the interactions between the cations and DNA duplex, thus increases the rigidity of local nucleic acid portion. For the base pairs, the backbone charge reduction slightly decreases their entropies by ~10 kcal mol^{-1} for these two types of studied DNA sequences in both Mg^{2+} and Na^+ electrolytes. This behavior indicates that the decrease in conformational flexibility of base pairs is caused by the phosphate neutralization, implying that the interactions between the base pairs are strengthened. The enhancement of stacked base pair rigidity can also be confirmed by the phenomenon that the GC and AT sequences in the charge-reduced form are systematically more rigid than those in the charge-full form, as shown in Table 3. These results are in accordance with the experimental results, which state that the

Table 3 Elastic force constants for deformations along local helical parameters of $d(GC)_{10}\cdot d(CG)_{10}$ and $d(AT)_{10}\cdot d(TA)_{10}$ in Na^+ and Mg^{2+} electrolytes. The central eighteen base pair portion is chosen for helical parameter calculations. Couplings between different parameters are neglected

	Shift[a]	Slide	Rise	Tilt[b]	Roll	Twist
GCC-Na	2.071	3.181	11.212	0.057	0.022	0.031
GCN-Na	3.336	3.030	11.999	0.060	0.028	0.067
GCC-Mg	2.204	2.892	10.962	0.057	0.024	0.034
GCN-Mg	3.343	2.960	11.939	0.060	0.028	0.065
ATC-Na	1.888	3.014	9.632	0.044	0.018	0.050
ATN-Na	2.967	3.810	10.942	0.051	0.021	0.074
ATC-Mg	1.928	3.547	10.392	0.049	0.018	0.060
ATN-Mg	2.941	3.928	11.022	0.051	0.020	0.074

[a] Shift, slide and rise are in kcal mol^{-1} $Å^{-2}$. [b] Tilt, roll and twist are in kcal mol^{-1} deg^{-2}.

melting temperature for a partially charged duplex would increase compared to the fully charged DNA.[93] We attribute this phenomenon to the weakened repulsion coming from the less charged backbone phosphates helping to stabilize the duplex; specifically, less twisting, rising and other forces to destabilize the π–π interactions between the stacked base pairs.

Anharmonicity correction to entropy. In the present work, the entropies are calculated using PCA analysis. In order to assess the anharmonicity correction to the quasi-harmonic approximation, the method developed by Baron, Hünenberger, McCammon and their coworkers[96,97] was employed to calculate not only the anharmonicities (non-Gaussian behaviors) in the probability distributions along individual essential motion modes but also the anharmonicity correction for the entropy of the DNA duplex. More details of the methodology can be found in the ESI.†

The actual probability distributions ($p'_m(\mathbf{a}_{\mathbf{q},m})$) of the transformed essential mode (m) coordinates ($\mathbf{a}_{\mathbf{q},m}$) (eqn (S5), ESI†), derived from the MD trajectories (the charge-full and charge-reduced forms of GC sequence in Na^+ are used as examples), are shown in Fig. S11 and S12, ESI.† It is demonstrated that the actual probability distributions $p'_m(\mathbf{a}_{\mathbf{q},m})$ are very close to the Gaussian function, even for the essential mode $m = 1$ in either the charge-full or charge-reduced states. The correlation coefficient R (shown in Fig. S13, ESI†) also indicates that the distributions associated with low-frequency modes are close to ideal Gaussians. These results mean that the global motion of the DNA duplex is quite close to the harmonic mode. The anharmonicity correction (ΔS_{qh}^{ah}) for the entropy calculated using PCA is displayed in Table 4. It is shown that ΔS_{qh}^{ah} is approximately −0.1 kcal mol^{-1} for all the systems investigated. This indicates that the harmonic approximation is appropriate for the evaluation of the DNA conformational entropies.

Table 4 Conformational entropy (S_{qh}) of the whole DNA, the backbone and sugar part (labeled as BB-sugar), and bases for $d(GC)_{10}\cdot d(CG)_{10}$ and $d(AT)_{10}\cdot d(TA)_{10}$. S_{qh} is the entropy calculated by the PCA method, and ΔS_{qh}^{ah} is the value of the anharmonic correction. The method developed by Baron, Hünenberger, McCammon and their coworkers[96,97] was employed to assess the anharmonicity correction to the quasi-harmonic (qh) analysis. The time-dependent problem of S_{qh} is alleviated by use of the method developed by Harris *et al.*,[98] which uses partial entropy estimates obtained for different time windows and allows us to estimate the entropy at infinite simulation time. More details can be seen in the ESI.† All values are in kcal mol^{-1}

	Bases		BB-sugar		DNA	
	S_{qh}	ΔS_{qh}^{ah}	S_{qh}	ΔS_{qh}^{ah}	S_{qh}	ΔS_{qh}^{ah}
GCC-Na	499.93	−0.10	777.93	−0.10	1166.92	−0.11
GCN-Na	482.38	−0.10	715.92	−0.10	1093.81	−0.10
GCC-Mg	493.36	−0.12	718.37	−0.13	1153.93	−0.13
GCN-Mg	483.80	−0.09	715.92	−0.09	1098.36	−0.09
ATC-Na	572.56	−0.09	761.05	−0.09	1218.63	−0.10
ATN-Na	552.09	−0.11	717.92	−0.11	1162.21	−0.12
ATC-Mg	558.49	−0.11	739.92	−0.11	1185.24	−0.12
ATN-Mg	550.81	−0.13	718.05	−0.13	1160.89	−0.13

For MD simulation, a natural question is how long the trajectory is able to achieve convergence for investigated properties. For DNA systems, Orozco *et al.* and Lankaš *et al.* proved that a good convergence can be achieved in 10 ns trajectories for (∼20)-mer GC/AT mixed DNA sequences,[99–102] and the structural properties are in good agreement with experimental results. Beveridge *et al.* demonstrated that the MD structural values for DNA parameters appeared to stabilized after 3–5 ns of MD, *i.e.*, ≈10 times the relaxation time (≈0.5 ns).[24] Moreover, the calculations performed by Li *et al.*[32] showed that conformational properties of isolated 23-mer DNA are comparable to the periodic-DNA in Na^+ and that in Mg^{2+} for 100 ns trajectories. In the current work, the DNA strand was chosen as a 20-mer, and the base pairs at the ends of the duplex are ignored in order to avoid the end effects of the DNA. For the systems investigated in the current work, similarity indicies[100] were calculated in order to verify the convergence of the analyzed trajectory portion, and the relative similarity index between the two equivalent blocks of this trajectory portion is larger than 0.95, which indicates that a stable structure of DNA duplex is well maintained in the MD simulation. Furthermore, we have performed PCA calculations for different portions of each trajectory, such as from 40 ns, 60 ns, 80 ns to the end (∼150 ns), and the calculated eigenvalues are in good agreement with each other. In the article, only the results of the trajectory portion from 40 ns to the end are shown in Fig. 5.

4 Conclusions

In the present work, we have performed MD simulations on the charge-full and uniform charge-reduced DNA duplex, and studied the effect of interplay between DNA and metal ions in solution on the DNA structure and conformational flexibility. At the structural level, the phosphate neutralization greatly changes the cation atmosphere around the DNA duplex, increases the major groove width and decreases the minor groove width. Although the DNA global bending angle is not greatly changed along with the backbone charge reduction, the bending direction preference is reduced. Furthermore, it is also noted that the proportion of BI phosphate linkages increase greatly if the net charge of the backbone phosphate groups is decreased.

More importantly, the electrostatic effect on the dynamic conformational flexibility is dependent on the DNA sequence. The phosphate backbone neutralization induces the global dynamics bending to be less flexible for GC sequences but more flexible for AT sequences compared to the corresponding charge-full states. Considering that there is no consensus on to what extent the electrostatic repulsion contributes to the DNA rigidity, our results prove that the sequence specific effect is critical in addressing the question of how electrostatic interaction on the DNA backbone affects the DNA flexibility.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Grant nos. 20934004 and 91127046) and

the National Basic Research Program of China (Grant nos. 2012CB821500 and 2010CB934500). Part of the computations were performed at the Supercomputing Center of University of Science and Technology of China. The authors thank Dr Yuan Liu (University of Washington in Seattle) for insightful discussions.

References

1 B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Molecular Biology of The Cell*, Garland Science, New York, London, 5th edn, 2007.
2 S. Neidle, *Principle of nucleic acid structure*, Elsevier, New York, 2008.
3 V. A. Bloomfield, *Biopolymers*, 1997, **44**, 269–282.
4 K. A. Sharp and B. Honig, *Curr. Opin. Struct. Biol.*, 1995, **5**, 323–328.
5 G. S. Manning, *J. Chem. Phys.*, 1969, **51**, 924.
6 Y. Bai, M. Greenfeld, K. J. Travers, V. B. Chu, J. Lipfert, S. Doniach and D. Herschlag, *J. Am. Chem. Soc.*, 2007, **129**, 14981–14988.
7 D. C. Rau, B. Lee and V. A. Parsegian, *Proc. Natl. Acad. Sci. U. S. A.*, 1984, **81**, 2621–2625.
8 C. M. Knobler and W. M. Gelbart, *Annu. Rev. Phys. Chem.*, 2009, **60**, 367–383.
9 A. D. MacKerell Jr and L. Nilsson, *Curr. Opin. Struct. Biol.*, 2008, **18**, 194–199.
10 N. V. Hud and J. Feigon, *J. Am. Chem. Soc.*, 1997, **119**, 5756–5757.
11 N. V. Hud, V. Sklenář and J. Feigon, *J. Mol. Biol.*, 1999, **286**, 651–660.
12 X. Shui, L. McFail-Isom, G. G. Hu and L. D. Williams, *Biochemistry*, 1998, **37**, 8341–8355.
13 V. Tereshko, G. Minasov and M. Egli, *J. Am. Chem. Soc.*, 1999, **121**, 3590–3595.
14 B. Heddi, N. Foloppe, E. Hantz and B. Hartmann, *J. Mol. Biol.*, 2007, **368**, 1403–1411.
15 N. V. Hud and M. Polak, *Curr. Opin. Struct. Biol.*, 2001, **11**, 293–301.
16 V. P. Denisov and B. Halle, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 629–633.
17 H. R. Drew and R. E. Dickerson, *J. Mol. Biol.*, 1981, **151**, 535–556.
18 N. C. Stellwagen, S. Magnusdottir, C. Gelfi and P. G. Righetti, *J. Mol. Biol.*, 2001, **305**, 1025–1033.
19 D. Hamelberg, L. McFail-Isom, L. D. Williams and W. D. Wilson, *J. Am. Chem. Soc.*, 2000, **122**, 10513–10520.
20 M. Feig and B. M. Pettitt, *Biophys. J.*, 1999, **77**, 1769–1781.
21 K. J. McConnell and D. Beveridge, *J. Mol. Biol.*, 2000, **304**, 803–820.
22 M. Soler-López, L. Malinina, J. Liu, T. Huynh-Dinh and J. A. Subirana, *J. Biol. Chem.*, 1999, **274**, 23683–23686.
23 P. Auffinger and E. Westhof, *J. Mol. Biol.*, 2000, **300**, 1113–1131.
24 S. Y. Ponomarev, K. M. Thayer and D. L. Beveridge, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 14771–14775.
25 J. A. Subirana and M. Soler-López, *Annu. Rev. Biophys. Biomol. Struct.*, 2003, **32**, 27–45.
26 H. Millonig, J. Pous, C. Gouyette, J. A. Subirana and J. L. Campos, *J. Inorg. Biochem.*, 2009, **103**, 876–880.
27 G. Minasov, V. Tereshko and M. Egli, *J. Mol. Biol.*, 1999, **291**, 83–99.
28 C. C. Sines, L. McFail-Isom, S. B. Howerton, D. VanDerveer and L. D. Williams, *J. Am. Chem. Soc.*, 2000, **122**, 11048–11056.
29 C. L. Kielkopf, S. Ding, P. Kuhn and D. C. Rees, *J. Mol. Biol.*, 2000, **296**, 787–801.
30 T. K. Chiu and R. E. Dickerson, *J. Mol. Biol.*, 2000, **301**, 915–945.
31 J. Liu and J. A. Subirana, *J. Biol. Chem.*, 1999, **274**, 24749–24752.
32 W. Li, L. Nordenskiöld and Y. Mu, *J. Phys. Chem. B*, 2011, **115**, 14713–14720.
33 J. Yoo and A. Aksimentiev, *J. Phys. Chem. B*, 2012, **116**, 12946–12954.
34 Z.-J. Tan and S.-J. Chen, *Biophys. J.*, 2008, **94**, 3137–3149.
35 D. Djuranovic and B. Hartmann, *J. Biomol. Struct. Dyn.*, 2003, **20**, 771–788.
36 B. Heddi, N. Foloppe, N. Bouchemal, E. Hantz and B. Hartmann, *J. Am. Chem. Soc.*, 2006, **128**, 9170–9177.
37 B. Wellenzohn, W. Flader, R. H. Winger, A. Hallbrucker, E. Mayer and K. R. Liedl, *Biochemistry*, 2002, **41**, 4088–4095.
38 C. Tisné, M. Delepierre and B. Hartmann, *J. Mol. Biol.*, 1999, **293**, 139–150.
39 B. Hartmann, M. R. Sullivan and L. F. Harris, *Biopolymers*, 2003, **68**, 250–264.
40 D. G. Gorenstein, *Chem. Rev.*, 1994, **94**, 1315–1338.
41 D. Djuranovic and B. Hartmann, *Biophys. J.*, 2005, **89**, 2542–2551.
42 M. Guéroult, O. Boittin, O. Mauffret, C. Etchebest and B. Hartmann, *PLoS One*, 2012, **7**, e41704.
43 T. Dršata, A. Pérez, M. Orozco, A. V. Morozov, J. Šponer and F. Lankaš, *J. Chem. Theory Comput.*, 2013, **9**, 707–721.
44 K. M. Kosikov, A. A. Gorin, X.-J. Lu, W. K. Olson and G. S. Manning, *J. Am. Chem. Soc.*, 2002, **124**, 4838–4847.
45 A. Savelyev, C. K. Materese and G. A. Papoian, *J. Am. Chem. Soc.*, 2011, **133**, 19290–19293.
46 T. Odijk, *J. Polym. Sci., Polym. Phys. Ed.*, 1977, **15**, 477–483.
47 J. Skolnick and M. Fixman, *Macromolecules*, 1977, **10**, 944–948.
48 C. G. Baumann, S. B. Smith, V. A. Bloomfield and C. Bustamante, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 6185–6190.
49 G. S. Manning, *Biophys. J.*, 2006, **91**, 3607–3616.
50 A. Podestá, M. Indrieri, D. Brogioli, G. S. Manning, P. Milani, R. Guerra, L. Finzi and D. Dunlap, *Biophys. J.*, 2005, **89**, 2558–2563.
51 S. A. Woodson, *Curr. Opin. Chem. Biol.*, 2005, **9**, 104–109.
52 A. Savelyev and G. A. Papoian, *J. Am. Chem. Soc.*, 2006, **128**, 14506–14518.
53 X. Shen, N. A. Atamas and F. S. Zhang, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2012, **85**, 051913.
54 J. P. Peters and L. J. Maher, *Q. Rev. Biophy.*, 2010, **43**, 23–63.

55 R. Sarma and M. Sarma, *Biological Structure and Dynamics: Proceedings of the ninth Conversation in the Discipline Biomolecular Stereodynamics, held at the State University of New York at Albany, June 20–24, 1995*, Adenine Press, 1996.
56 D. A. Case, T. A. Darden, T. E. Cheatham III and P. A. Kollman, *et al.*, *AMBER 12*, University of California, San Francisco, 2012, http://ambermd.org/.
57 A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham III, C. A. Laughton and M. Orozco, *Biophys. J.*, 2007, **92**, 3817–3829.
58 J. Åaqvist, *J. Phys. Chem.*, 1990, **94**, 8021–8024.
59 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.
60 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kal and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
61 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.
62 J.-P. Ryckaert, G. Ciccotti and H. J. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.
63 J. A. Izaguirre and R. D. Skeel, *J. Chem. Phys.*, 1999, **110**, 9853.
64 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089.
65 W. G. Hoover, *Phys. Rev. A*, 1985, **31**, 1695–1697.
66 J. Šponer, P. Jurečka, I. Marchan, F. J. Luque, M. Orozco and P. Hobza, *Chem.–Eur. J.*, 2006, **12**, 2854–2865.
67 J. Šponer, J. Leszczyński and P. Hobza, *J. Comput. Chem.*, 1996, **17**, 841–850.
68 M. Kolář, K. Berka, P. Jurečka and P. Hobza, *ChemPhysChem*, 2010, **11**, 2399–2408.
69 J. Šponer, J. Leszczyński and P. Hobza, *J. Phys. Chem.*, 1996, **100**, 5590–5596.
70 J. Spiriti, H. Kamberaj, A. M. R. de Graff, M. F. Thorpe and A. van der Vaart, *J. Chem. Theory Comput.*, 2012, **8**, 2145–2156.
71 J. Curuksu, K. Zakrzewska and M. Zacharias, *Nucleic Acids Res.*, 2008, **36**, 2268–2283.
72 P. Auffinger and E. Westhof, *J. Mol. Biol.*, 2001, **305**, 1057–1072.
73 E. Sletten and N. A. Frøystein, *Met. Ions Biol. Syst.*, 1996, **32**, 397–418.
74 V. A. Buckin, B. I. Kankiya, D. Rentzeperis and L. A. Marky, *J. Am. Chem. Soc.*, 1994, **116**, 9423–9429.
75 Y. Swarnalatha and N. Yathindra, *J. Biomol. Struct. Dyn.*, 1993, **10**, 1023–1045.
76 E. N. Nikolova, G. D. Bascom, I. Andricioaei and H. M. Al-Hashimi, *Biochemistry*, 2012, **51**, 8654–8664.
77 M. Egli, *Chem. Biol.*, 2002, **9**, 277–286.
78 E. Stellwagen, J. P. Peters, L. J. Maher and N. C. Stellwagen, *Biochemistry*, 2013, **52**, 4138–4148.
79 I. Brukner, S. Susic, M. Dlakic, A. Savic and S. Pongor, *J. Mol. Biol.*, 1994, **236**, 26–32.
80 J. K. Strauss, C. Roberts, M. G. Nelson, C. Switzer and L. J. Maher, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 9515–9520.
81 J. Curuksu, M. Zacharias, R. Lavery and K. Zakrzewska, *Nucleic Acids Res.*, 2009, **37**, 3766–3773.
82 A. A. Gorin, V. B. Zhurkin and W. K. Olson, *J. Mol. Biol.*, 1995, **247**, 34–48.
83 M. S. Babcock and W. K. Olson, *J. Mol. Biol.*, 1994, **237**, 98–124.
84 D. Bhattacharyya and M. Bansal, *J. Biomol. Struct. Dyn.*, 1990, **8**, 539–572.
85 A. Pérez, F. Lankas, F. J. Luque and M. Orozco, *Nucleic Acids Res.*, 2008, **36**, 2379–2394.
86 A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.
87 T. Okonogi, S. Alley, A. Reese, P. Hopkins and B. Robinson, *Biophys. J.*, 2000, **78**, 2560–2571.
88 J. Strauss and L. Maher, *Science*, 1994, **266**, 1829–1834.
89 A. D. Mirzabekov and A. Rich, *Proc. Natl. Acad. Sci. U. S. A.*, 1979, **76**, 1118–1121.
90 T. Odijk and A. C. Houwaart, *J. Polym. Sci., Polym. Phys. Ed.*, 1978, **16**, 627–639.
91 J. M. Schurr and K. S. Schmitz, *Annu. Rev. Phys. Chem.*, 1986, **37**, 271–305.
92 R. Gurlie and K. Zakrzewska, *J. Biomol. Struct. Dyn.*, 1998, **16**, 605–618.
93 T. M. Okonogi, S. C. Alley, E. A. Harwood, P. B. Hopkins and B. H. Robinson, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 4156–4160.
94 K. R. Hagerman and P. J. Hagerman, *J. Mol. Biol.*, 1996, **260**, 207–223.
95 L. D. Williams and L. J. Maher III, *Annu. Rev. Biophys. Biomol. Struct.*, 2000, **29**, 497–521.
96 R. Baron, W. F. van Gunsteren and P. H. Hünenberger, *Trends Phys. Chem.*, 2006, **11**, 87–122.
97 R. Baron, P. H. Hünenberger and J. A. McCammon, *J. Chem. Theory Comput.*, 2009, **5**, 3150–3160.
98 S. A. Harris, E. Gavathiotis, M. S. Searle, M. Orozco and C. A. Laughton, *J. Am. Chem. Soc.*, 2001, **123**, 12658–12663.
99 A. Noy, A. Pérez, F. Lankas, F. J. Luque and M. Orozco, *J. Mol. Biol.*, 2004, **343**, 627–638.
100 A. Pérez, J. R. Blas, M. Rueda, J. M. López-Bes, X. de la Cruz and M. Orozco, *J. Chem. Theory Comput.*, 2005, **1**, 790–800.
101 F. Lankaš, J. Šponer, P. Hobza and J. Langowski, *J. Mol. Biol.*, 2000, **299**, 695–709.
102 F. Lankaš, J. Šponer, J. Langowski and T. E. Cheatham, *Biophys. J.*, 2003, **85**, 2872–2883.