

# Competition between stacking and divalent cation mediated electrostatic interactions determines the conformations of short DNA sequences

Balaka Mondal,<sup>†</sup> Debayan Chakraborty,<sup>†</sup> Naoto Hori,<sup>‡</sup> Hung T. Nguyen,<sup>†</sup> and  
D. Thirumalai<sup>\*,†,¶</sup>

<sup>†</sup>*Department of Chemistry, The University of Texas, Austin, TX 78712, USA*

<sup>‡</sup>*School of Pharmacy, University of Nottingham, Nottingham, NG7 2RD, United Kingdom*

<sup>¶</sup>*Department of Physics, The University of Texas, Austin, TX 78712, USA*

E-mail: dave.thirumalai@gmail.com

## Abstract

Interplay between divalent cations ( $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$ ) and single stranded DNA (ssDNA) and double stranded DNA (dsDNA), as well as stacking interactions, are important in nucleosome stability and phase separation in nucleic acids. Quantitative techniques accounting for ion-DNA interactions are needed to obtain insights into these and related problems. Towards this end, we created a sequence dependent computational TIS-ION model that explicitly accounts for monovalent and divalent ions. Simulations of the rigid 24 base-pair (bp) dsDNA and flexible ssDNA sequences, dT<sub>30</sub> and dA<sub>30</sub>, with varying amounts of the divalent cations, show that the calculated excess number of ions around the dsDNA and ssDNA *agree quantitatively with ion-counting experiments*. Using an ensemble of all-atom structures generated from coarse-grained simulations,

we calculated the Small Angle X-ray Scattering (SAXS) profiles, which are in excellent agreement with experiments. Although ion counting experiments mask the differences between  $Mg^{2+}$  and  $Ca^{2+}$ , we find that  $Mg^{2+}$  binds to the minor grooves and phosphate groups whereas  $Ca^{2+}$  binds specifically to the minor groove. Both  $Mg^{2+}$  and  $Ca^{2+}$  exhibit a tendency to bind to the minor groove of DNA as opposed to the major groove. The  $dA_{30}$  conformations are dominated by stacking interactions, resulting in structures with considerable helical order. The near cancellation of the favorable stacking and unfavorable electrostatic interactions leads to  $dT_{30}$  populating an ensemble of heterogeneous conformations. The successful applications of the TIS-ION model poises to confront many problems in DNA biophysics.

## Introduction

Single and double stranded DNA are polyanions because each phosphate group carries a negative charge. As a result, their conformations are often viewed solely through the lens of electrostatic interactions, which are sensitive to the size, shape and valence of the counterions. Indeed, interactions between backbone the phosphate groups and cations, especially multivalent ions, determine the flexibility of DNAs.<sup>1,5</sup> Specifically, cation-mediated changes in the DNA conformations control their propensities to interact with other biomolecules, which in turn are involved in transcription and organization of chromosomes on long length scales.<sup>2-4</sup> Ion counting experiments, Small Angle X-ray Scattering (SAXS) measurements, as well as Fluorescence Resonance Energy Transfer (FRET) experiments have provided insights into the interactions of divalent cations with nucleic acids (NAs).<sup>6-10</sup> Besides electrostatic interactions, the stability of dsDNA should also depend on both Watson-Crick base-pairing ( $U_{HB}$ ) and base stacking interactions ( $U_{ST}$ ). The relative contributions of  $U_{HB}$  and  $U_{ST}$ , measured in dsDNA by introducing a single nick as a function of temperature and monovalent salt concentration<sup>11</sup> found that the stability of dsDNA was predominantly determined by stacking interactions.

In the 24-base pair (bp) dsDNA (Figure 1A), studied in ion-counting experiments,<sup>6</sup> which is used as a case study in this work, Watson-Crick base pairs are fully formed. As a result, the dsDNA is rigid, with the persistence length ( $\approx 50$  nm) being much greater than the contour length ( $\approx 7.6$  nm). Combination of molecular dynamics simulations and integral equation theories successfully reproduced the experimental data from ion counting experiments for a 24-nucleotide dsDNA sequence only in monovalent (NaCl) salt solution.<sup>12</sup> Are there differences between  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  in the interactions with dsDNA?

Both ion counting and SAXS experiments of short homopolymeric single-stranded DNAs (ssDNAs), poly dA and poly dT, suggested that their ion-atmospheres are similar.<sup>10,13</sup> A plausible implication is that the differences in their conformational ensembles must arise from stacking interactions. The brief survey<sup>10,13</sup> illustrates that experiments cannot resolve the relative importance of  $U_{ST}$ ,  $U_{HB}$  and ion mediated interactions in modulating the structures of DNA. Single-molecule pulling experiments and theoretical analysis<sup>14,15</sup> have shown that the bare persistence length of ssDNAs is  $\approx 1$  nm, which is much smaller than the contour length of even the relatively short sequences (dA<sub>30</sub> and dT<sub>30</sub>) investigated here. Consequently, cations would affect their conformations to a greater extent than dsDNA, raising the question, how do the conformations of these two ssDNA sequences change as the concentrations of  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  are varied? Because purine stacks are preferred over pyrimidine stacks, stacking interactions involving poly dA is more favorable than stacking of poly dT. Although it is clear that competition of stacking and electrostatic interactions should determine the ensembles of states, a quantitative assessment of their relative contributions has not been made. The brief survey<sup>10,13</sup> illustrates that experiments alone cannot resolve the relative importance of  $U_{ST}$ ,  $U_{HB}$  and ion mediated interactions in modulating the structures of DNA. Thus, there is a need to develop well calibrated computational models that not only can take sequence effects into account but also must include key interaction in DNA as well as the role of cations.

Theoretical models, such as, Poisson-Boltzmann (PB) equation,<sup>16-19</sup> simulations using

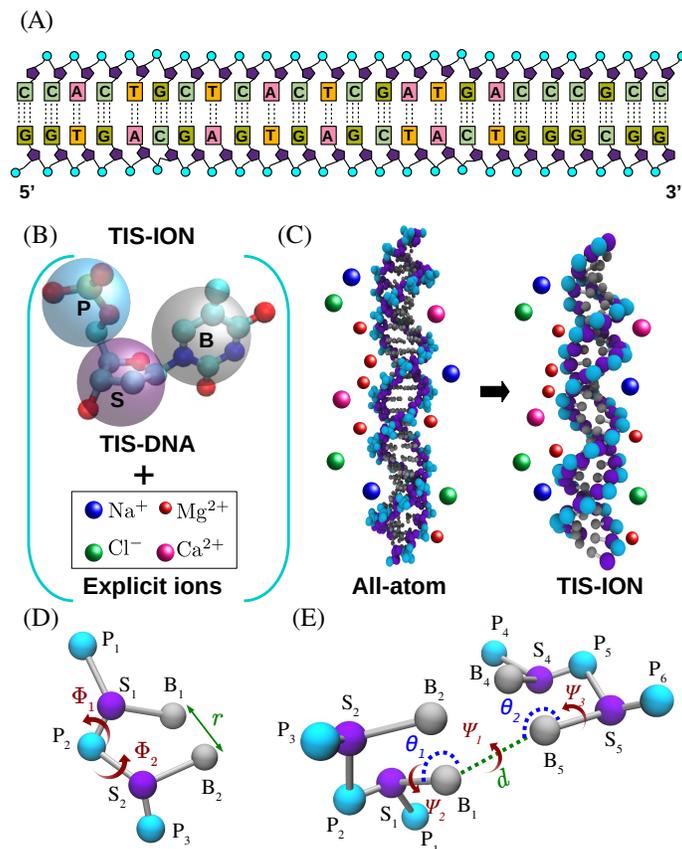


Figure 1: **Coarse-grained DNA with explicit ions.** (A) Sequence of the 24-bp dsDNA. (B) Coarse-grained representation of a single nucleotide. In TIS-DNA model, each nucleotide is described using three interaction sites, placed at the center of phosphate (P), sugar (S) and base (B) groups, shown in cyan, purple, and grey, respectively. Ions are modeled explicitly using spherical beads of appropriate radii. Na<sup>+</sup>, Mg<sup>2+</sup>, Ca<sup>2+</sup> and Cl<sup>-</sup> are shown in blue, red, magenta and green, respectively. (C) All-atom and coarse-grained representations of the dsDNA in Na<sup>+</sup>, Mg<sup>2+</sup>, Ca<sup>2+</sup> and Cl<sup>-</sup> ions. Similar color schemes are adapted in all-atom and coarse-grained representations. The three-dimensional (3D) structures of dsDNA were generated using the Nucleic Acid Builder (NAB)<sup>64</sup> software by adopting the parameters for B-DNA. The TIS-DNA<sup>22</sup> coarse-grained structures were generated from the all-atom structures using the visual molecular dynamics (VMD)<sup>65</sup> software. (D) Illustration of the structural features associated with base-stacking (Eq. 3). The distance between adjacent bases B<sub>1</sub> and B<sub>2</sub> is  $r$ ,  $\phi_1(P_1, S_1, P_2, S_2)$  and  $\phi_2(P_3, S_2, P_2, S_1)$  are the dihedral angles around sugar-phosphate bonds. (E) Geometry associated with hydrogen bond in Eq. 4. Hydrogen bonding distance  $d$  is measured between bases B<sub>1</sub> and B<sub>5</sub>;  $\theta_1(S_1, B_1, B_5)$  and  $\theta_2(S_5, B_5, B_1)$  are the angles;  $\psi_1(S_1, B_1, B_5, S_5)$ ,  $\psi_2(P_2, S_1, B_1, B_5)$  and  $\psi_3(P_6, S_5, B_5, B_1)$  are the dihedral angles.

Debye-Hückel theory,<sup>20-23</sup> or Manning condensation<sup>24</sup> have mimicked the influence of ions on NAs in monovalent salts. However, these models are not as successful in describing DNA interactions with divalent or multivalent cations, where ion-ion correlations are important. Progress has been made in using all-atom (AA) simulations<sup>12,25-27</sup> to reproduce a few experimental measurements. Excess  $\text{Na}^+$  around the 24-nucleotide dsDNA, calculated using AA simulations in a fixed background of ( $\sim 5\text{mM}$ )  $\text{Mg}^{2+}$ ,<sup>28</sup> is in very good agreement with ion counting experiments. These simulations were performed by keeping the dsDNA rigid, which could be justified because the short sequence is rigid. These studies did not compare the simulation results with experiments (ion counting results and SAXS profiles) over a wide range of divalent ion concentrations. One of the previous coarse-grained (CG) studies, built in the spirit of the TIS model,<sup>29</sup> used surrogate small molecules (dimethylphosphate and magnesium hexahydrate), as done in AA simulations<sup>28</sup> to calculate energetic parameters involving monovalent, divalent, and multivalent cations. The calculated persistent length for  $\text{dA}_{68}$  as a function of  $\text{Na}^+$  is in reasonable agreement with experiments. Several earlier models<sup>12,25-29</sup> have enjoyed modest success in reproducing a few mechanical and thermodynamic properties of dsDNA and ssDNA sequences. However, a unified model that would simultaneously reproduce a variety of experimental measurements for both dsDNA and ssDNA, in both monovalent and divalent cations, over a broad concentration range, is still missing.

In order to overcome the difficulties associated with the previous models, we created a new coarse-grained model that is based on the TIS model for NAs.<sup>30,31</sup> Each nucleotide is represented as three spherical beads that are located at the centers of mass of phosphate (P, in cyan), sugar (S, in purple) and base (B, in grey) groups (Figure 1B). In some of our previous studies on NAs, electrostatic interactions between the phosphate (P) groups were modeled using the Debye-Hückel potential with a reduced charge on P in order to account for counter ion condensation.<sup>21,22</sup> The TIS model, with the implicit treatment of ions, is moderately successful and is the basis on which other three site models for NAs were developed.<sup>32-34</sup> However, not including monovalent and divalent cations explicitly makes it impossible to

investigate myriad of problems involving DNA biophysics. To overcome this problem, we combine the TIS model for DNA<sup>22</sup> and explicit treatment of monovalent and divalent cations to create the three interaction site TIS-ION model. To develop the new model, we followed the strategy used in the context of RNA folding<sup>35-37</sup> in which the effects of  $\text{Mg}^{2+}$ , and  $\text{Ca}^{2+}$ , monovalent cations were modeled explicitly. The resulting TIS-ION model, which also accounts for sequence effects in DNA, is a significant advance, as we demonstrate here by quantitative comparison with experiments, *without using a single adjustable parameter* to fit any of the data. The simulations quantitatively reproduce the experimental findings for the dependence of the number of ions around short DNA constructs as a function of  $[\text{Mg}^{2+}]$ . The corresponding results for  $\text{Ca}^{2+}$  are predictions, which can be readily tested. We also find that the agreement between the calculated and experimental SAXS profiles, for a 25-nucleotide dsDNA, dT<sub>30</sub> and dA<sub>30</sub> are excellent. Strikingly, simulations show that in these short DNA sequences (dsDNA and dA<sub>30</sub>), the conformations are determined by stacking interactions, with electrostatic interactions playing a minor role.

## Methods

**TIS-ION model:** Our previous coarse-grained TIS-DNA model,<sup>22</sup> used the Debye-Hückel (DH) theory to account for the electrostatic interactions. The incentive to develop an explicit ion model compatible with DNA arises from the following considerations: 1. DH theory is expected to be accurate only at low monovalent ion concentrations ( $< 10^{-3}$  moles/liter). 2. Experiments have shown that divalent cations such as  $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$  play a key role in stabilizing nucleic acids.<sup>4</sup> Thus, one has to go beyond the DH theory to predict the outcomes of DNA-ion interactions. 3. Numerous experiments<sup>5,49,60,61</sup> have shown that ion-specific and size-specific effects on the structure and dynamics of nucleic acids, which cannot be accounted for using the DH theory. The TIS-ION model, which treats monovalent and divalent cations explicitly, **represents a significant advance** towards the modeling of ion-nucleic acid

interactions, thus allowing us to deal with myriads of problems of biophysical interest.

The Supporting Information (SI) includes an overview of the parametrization of the model. It also lists the parameter values. The parameters for different ions are adopted from the TIS-RNA model with explicit ions<sup>35</sup> which accurately predicts the structure, thermodynamics and folding/unfolding transitions in a series of folded RNA sequences.<sup>35-37</sup> Surprisingly, none of the force-field parameters in the TIS-ION model required any adjustments to recapitulate the experimental results associated with rigid dsDNA or flexible ssDNA sequences.

The energy function in the TIS-ION model is,

$$U_{TIS-ION} = U_B + U_{Ang} + U_{EV} + U_{ST} + U_{HB} + U_{EL}. \quad (1)$$

The potentials for bond stretch ( $U_B$ ) and bond angle ( $U_{Ang}$ ) are given by,  $U_B = k_b(r - r_0)^2$  and  $U_{Ang} = k_\theta(\theta - \theta_0)^2$ , where  $r_0$  and  $\theta_0$  are the equilibrium bond length and bond angle, respectively, which were taken from an ideal B-form DNA helix. Excluded volume repulsion between sites  $i$  and  $j$  (DNA sites or ions) separated by a distance,  $r$  (in nm), is modeled as,

$$U_{EV} = \sum_{i,j} \epsilon_{ij} \left[ \left( \frac{0.16}{r + 0.16 - D_{ij}} \right)^{12} - 2 \left( \frac{0.16}{r + 0.16 - D_{ij}} \right)^6 + 1 \right], \quad r \leq D_{ij}, \quad (2)$$

where  $D_{ij} = R_i + R_j$  and  $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$ . When both the interacting sites belong to DNA,  $D_{ij}$  is taken to be 0.32 nm.<sup>21</sup>

**Stacking Potential:** The potential for stacking interactions,  $U_{ST}$ , between two consecutive nucleotides,  $i$  and  $i + 1$ , is taken to be,<sup>21,22</sup>

$$U_{ST} = \frac{U_{ST}^0}{1 + k_r(r - r_0)^2 + \sum_{i=1}^2 k_\phi(\phi_i - \phi_{i,0})^2}, \quad (3)$$

where  $r$  is the distance between the bases  $B_i$  and  $B_{i+1}$ , and  $\phi_1$  ( $P_i, S_i, P_{i+1}, S_{i+1}$ ) and  $\phi_2$  ( $P_{i+2}, S_{i+1}, P_{i+1}, S_i$ ) are the dihedral angles around the sugar-phosphate bonds (see Fig-

ure 1D). The equilibrium stacking distance  $r_0$ , and the backbone dihedral angles  $\phi_{1,0}$  and  $\phi_{2,0}$  are calculated from the coarse-grained representation of an ideal B-form DNA helix.  $U_{ST}^0$  gives an estimate of stacking energy between two consecutive nucleotides in an ideal B-form DNA helix, and is  $\approx -4.19$  kcal/mol for a T-dimer and  $-5.77$  kcal/mol for an A-dimer, at  $T = 277$  K.<sup>22</sup> It is worth noting that the  $U_{ST}^0$  values are not fit parameters, but calibrated in our previous work<sup>22</sup> to accurately reproduce the thermodynamics of DNA duplexes, following the nearest neighbor model.<sup>38,39</sup>

**Hydrogen bond interactions,  $U_{HB}$ :** Hydrogen bonds are considered between canonical base pairs (Watson-Crick), belonging to the complementary strands. The  $U_{HB}$  potential is,

$$U_{HB} = \frac{U_{HB}^0}{1 + k_d(d - d_0)^2 + \sum_{i=1}^2 k_\theta(\theta_i - \theta_{i,0})^2 + \sum_{i=1}^3 k_\psi(\psi_i - \psi_{i,0})^2},$$

where  $d$ ,  $\theta_1$ ,  $\theta_2$ ,  $\psi_1$ ,  $\psi_2$  and  $\psi_3$  are described in Figure 1E. Hydrogen bond interactions are absent in the homopolymeric ssDNA sequences.

**Electrostatic interactions,  $U_{EL}$ :** The interaction between two charged groups with charges  $Q_i$  and  $Q_j$ , separated by a distance,  $r_{ij}$ , is given by the Coulomb potential,

$$U_{EL} = \frac{e^2}{4\pi\epsilon_0\epsilon(T)} \sum_{i,j} \frac{Q_i Q_j}{r_{ij}}, \quad (4)$$

where  $\epsilon_0$  is the vacuum permittivity, and the temperature-dependent dielectric constant,  $\epsilon(T) = 87.74 - 0.4008T + 9.398 \times 10^{-4}T^2 - 1.410 \times 10^{-6}T^3$ , is taken from,<sup>40</sup> where the temperature,  $T$ , is reported in celcius. The charge on the phosphate group is  $Q_P = -1e$  (unit of electron charge). The force-field parameters are listed in the SI.

**Simulations:** We used OpenMM,<sup>41</sup> a high-performance toolkit for molecular simulations, to simulate ssDNA and dsDNA in mixed salt solutions of monovalent and divalent cations, on a Graphics Processing Unit (GPU). We used the Particle-Mesh-Ewald (PME) algorithm,<sup>42</sup> as implemented in OpenMM to account for the long-range Coulomb interactions. In order to enhance conformational sampling, the simulations were performed by integrating

the Langevin equation in the low friction limit.<sup>43</sup> The equation of motion for a bead or ion, with coordinate  $\vec{r}_i$ , is,

$$m_i \ddot{\vec{r}}_i = -m_i \zeta \dot{\vec{r}}_i + \vec{F}_c + \vec{\Gamma}, \quad (5)$$

where  $m_i$  is the mass of the bead,  $\zeta$  is the friction coefficient,  $\vec{F}_c = -\frac{\partial U_{TIS-ION}(\{\mathbf{r}\})}{\partial \vec{r}_i}$ , and  $\vec{\Gamma}$  is the random force with a white noise spectrum. The autocorrelation function of the random force in the discretized form is  $\langle \Gamma(t) \Gamma(t + nh) \rangle = \frac{2\zeta m_i k_B T}{h} \delta_{0,n}$ , where  $n = 0, 1, \dots$  and  $\delta_{0,n}$  is the Kronecker delta function. The equations of motion were integrated using Langevin leap-frog method,<sup>44</sup> as implemented in OpenMM, using  $\zeta = 0.01$  picosecond<sup>-1</sup>. The integration time step is 2 femtoseconds. We performed simulations at  $T = 277$  K to compare the results with the ion-counting experiments.<sup>6,10</sup> Additional simulations were performed at  $T = 298$  K to compare with the SAXS experiments.<sup>7,45</sup>

The systems consist of a single DNA in a cubic box of length  $L$  whose value depends on the ion concentration. To minimize finite size effects, we used periodic boundary conditions. Neutrality of the sample is maintained by adding an appropriate number of Cl<sup>-</sup> ions.

**Counting ions in the vicinity of NAs:** The experimentally measurable  $\Gamma_i$ , which is an estimate of the excess number of cations  $i$ , in the vicinity of NAs, is calculated using the Kirkwood-Buff integral,<sup>46,47</sup>

$$\Gamma_i = \int_V [C_i(r) - C_i^B] dV = C_i^B \int_V [g_i(r) - 1] dV, \quad (6)$$

where  $V$  is the volume over which the integral is evaluated. The concentration of ions  $i$  in the vicinity of the polyanion,  $C_i^B$ , differs from the bulk concentration  $C_i^S = \frac{N}{V_S}$ , where  $N$  is the total number of ions  $i$ , and  $V_S (=L^3)$  is the volume of the simulation box. We calculated  $C_i^B$  for the  $i^{th}$  ionic species after the concentration profile reaches a plateau at large separation from the DNA (see Supplementary Information (SI), Figure S1). In practice, the integral in Eq. 6 is carried out over a finite volume because the density fluctuations are predominantly localized near the NA surface.

For dsDNA, the volume integral is,

$$\Gamma_i = C_i^B 2\pi h \int_{r=0}^{r_c} [g_i(r) - 1] r dr, \quad (7)$$

where  $h$  is the approximate height of the cylinder encompassing the dsDNA along the  $z$  axis (helical axis or the principal axis),  $r$  is the radial distance from the principal axis of the dsDNA and  $r_c = \frac{L}{2}$ ,  $L$  is the simulation box length (see Figure S2). We varied  $h$  to obtain convergence in the  $\Gamma_i$  values and chose  $h = \frac{L}{2}$  (see Figure S3). The size of the smallest simulation box is  $L \approx 23$  nm which is significantly larger than the length of dsDNA  $\approx 7.6$  nm. For ssDNA, the volume integral is,

$$\Gamma_i = C_i^B 4\pi \int_{r=0}^{r_{cs}} [g_i(r) - 1] r^2 dr, \quad (8)$$

where  $r_{cs} = \frac{L}{2}$ , and  $r$  is the distance of the ion  $i$  from the center of mass of ssDNA (Figure S2B in the SI).

**Calculation of  $\Gamma_i$  from simulations:** We performed Langevin dynamics simulations of the 24 bp dsDNA and ssDNA dA<sub>30</sub> and dT<sub>30</sub>, in a solution containing both monovalent and divalent ions ( $T = 277$  K). First, we performed simulations for each DNA construct at different concentration of Mg<sup>2+</sup> or Ca<sup>2+</sup> by varying the box length from  $\approx 23$  nm to 83 nm. We used a constant Na<sup>+</sup> concentration (= 20 mM), and varied [Mg<sup>2+</sup>] or [Ca<sup>2+</sup>] from 0.1 mM to 20 mM (in case of ssDNA) or 50 mM (in case of dsDNA) to closely mimic the experiments.<sup>6,10</sup> We calculated the  $\Gamma_i$  values for each ionic species by averaging over all the sampled conformations after equilibrium has been reached using either Eq. 7 (dsDNA) or Eq. 8 (ssDNA).

**Stacking in ssDNA:** We describe stacking in ssDNA by introducing an order parameter,  $p_S$  – a measure that is related to the helix content<sup>48</sup> in a biomolecule. We assume that each nucleotide is either in a stacked ( $S$ ) state or in a random coil state ( $C$ ). To measure the stacking propensity of the  $i^{th}$  nucleotide, we calculated the sugar-phosphate dihedral angles,

$\psi(P_i, S_i, P_{i+1}, S_{i+1})$  and  $\phi(S_i, P_{i+1}, S_{i+1}, P_{i+2})$ , and compared them with the typical  $\psi_B$  and  $\phi_B$  values in an ideal B-form helix. If the deviations  $\delta\psi(= |\psi - \psi_B|)$  and  $\delta\phi(= |\phi - \phi_B|)$  are within 25% of the ideal helix parameters, we consider the nucleotide to be in the  $S$  state. For a given ssDNA conformation,  $p_S$  is defined as,

$$p_S = \frac{\sum_n^{n_S} N_n \Theta(N_n - 3)}{N_S^I}, \quad (9)$$

where  $n_S$  is the number of consecutive stacks along the ssDNA contour,  $N_n$  is the length of the  $n^{\text{th}}$  stack and  $N_S^I$  is the length of the persistent stacks in completely stacked ideal B helix and  $\Theta(N_n - 3)$  is unity only if  $N_n \geq 3$ . A minimum of three consecutive nucleotides ( $N_n \geq 3$ ) should be in the state  $S$  to calculate  $p_S$ . For instance, in the oligomer  $CCSCC\overline{SSSS}CCC\overline{SSSS}CCSSC$ , there are two persistent stacks ( $n_S = 2$ ) of length 4 and 3, ( $N_1 = 4, N_2 = 3$ ) respectively, yielding  $p_S = \frac{4+3}{20}$ , which means that there are 7 nucleotides that are persistently stacked. In the example, the lengths of the longest stretch of stacked and unstacked bases are,  $N_S^{\text{max}} = 4$  and  $N_U^{\text{max}} = 3$ , respectively.

## Results

**Comparison with ion counting experiments:** We compare in Figures 2A and B the calculated and experimental<sup>6</sup>  $\Gamma_{\text{Na}^+}$  (blue) and  $\Gamma_{\text{Mg}^{2+}}$  (red) or  $\Gamma_{\text{Ca}^{2+}}$  (magenta) for dsDNA as a function of  $[\text{Mg}^{2+}]$  (or  $[\text{Ca}^{2+}]$ ). The agreement is excellent for both monovalent and divalent cations over the entire concentration range. Because multivalent cations are more effective in screening the phosphate charges, monovalent cations ( $\text{Na}^+$ ) are replaced by divalent cations ( $\text{X}^{2+}$ ) as the concentration of  $\text{Mg}^{2+}$  or  $\text{Ca}^{2+}$  is increased. This is a consequence of the entropically favored counterion release mechanism.<sup>49-53</sup>

The simulation results for  $\Gamma_i$ , as a function of  $[\text{Mg}^{2+}]$ , for dT<sub>30</sub> and dA<sub>30</sub>. is also in very good agreement with experiments<sup>10</sup> for both dT<sub>30</sub> and dA<sub>30</sub> (Figure 3). Interestingly, there is negligible difference in  $\Gamma_i$  for dT<sub>30</sub> and dA<sub>30</sub>, which implies that the preferential ion

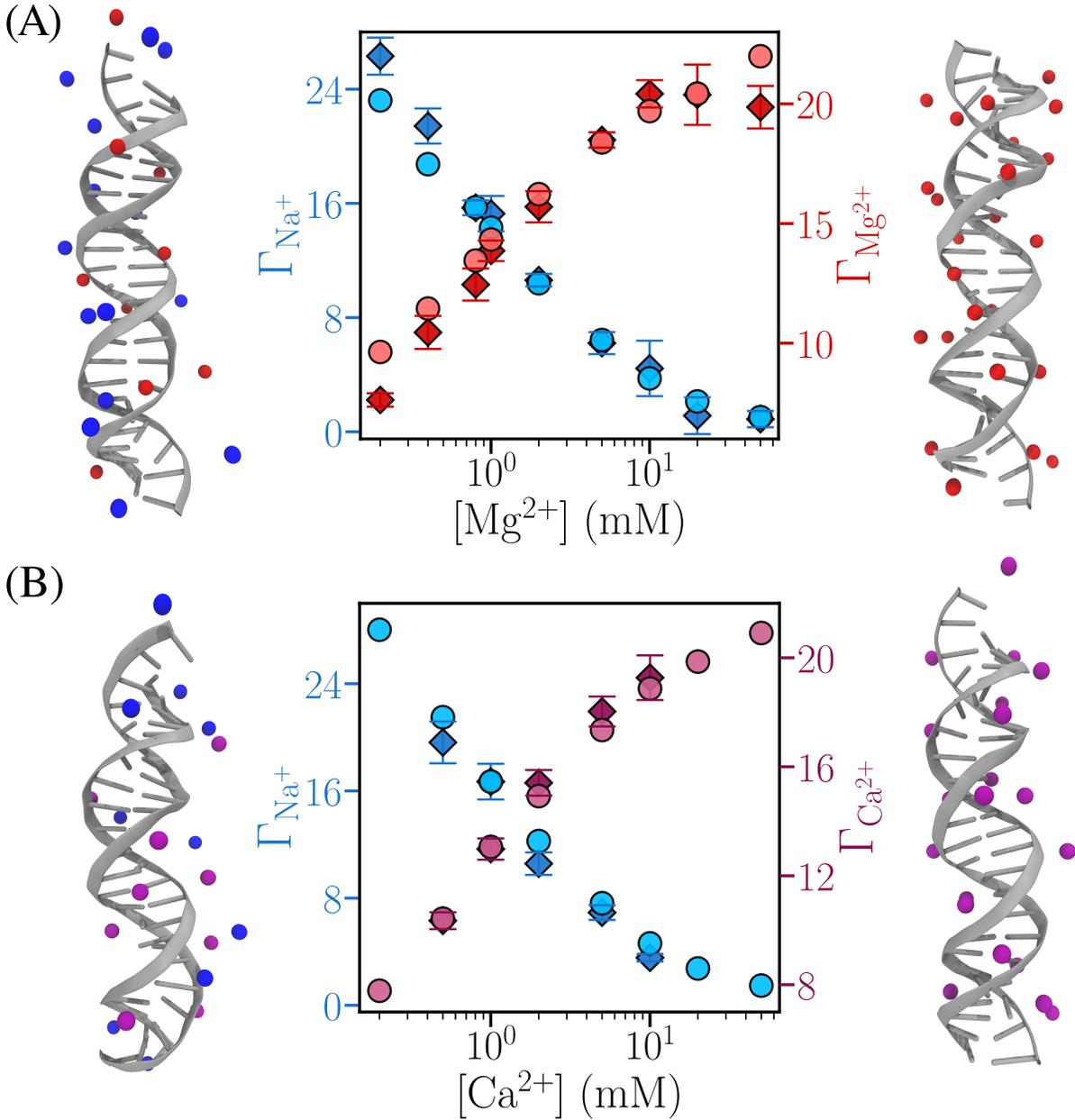


Figure 2: **Ions around dsDNA.** (A) Excess number of ions,  $\Gamma_{\text{Na}^+}$  (blue) and  $\Gamma_{\text{Mg}^{2+}}$  (red), in the vicinity of dsDNA as a function of  $[\text{Mg}^{2+}]$ . Experimental results<sup>6</sup> in diamonds in dark shade, and the simulation results are in circles in light shade. Reconstructed all-atom structures, highlighting  $\text{Na}^+$  and  $\text{Mg}^{2+}$  ions within 10 Å of the phosphate groups, are shown for  $[\text{Mg}^{2+}] = 0.2$  mM and 20 mM. (B) Same as (A), except the results are shown for  $\text{Ca}^{2+}$  ions (in magenta). In all the panels,  $[\text{Na}^+] = 20$  mM.

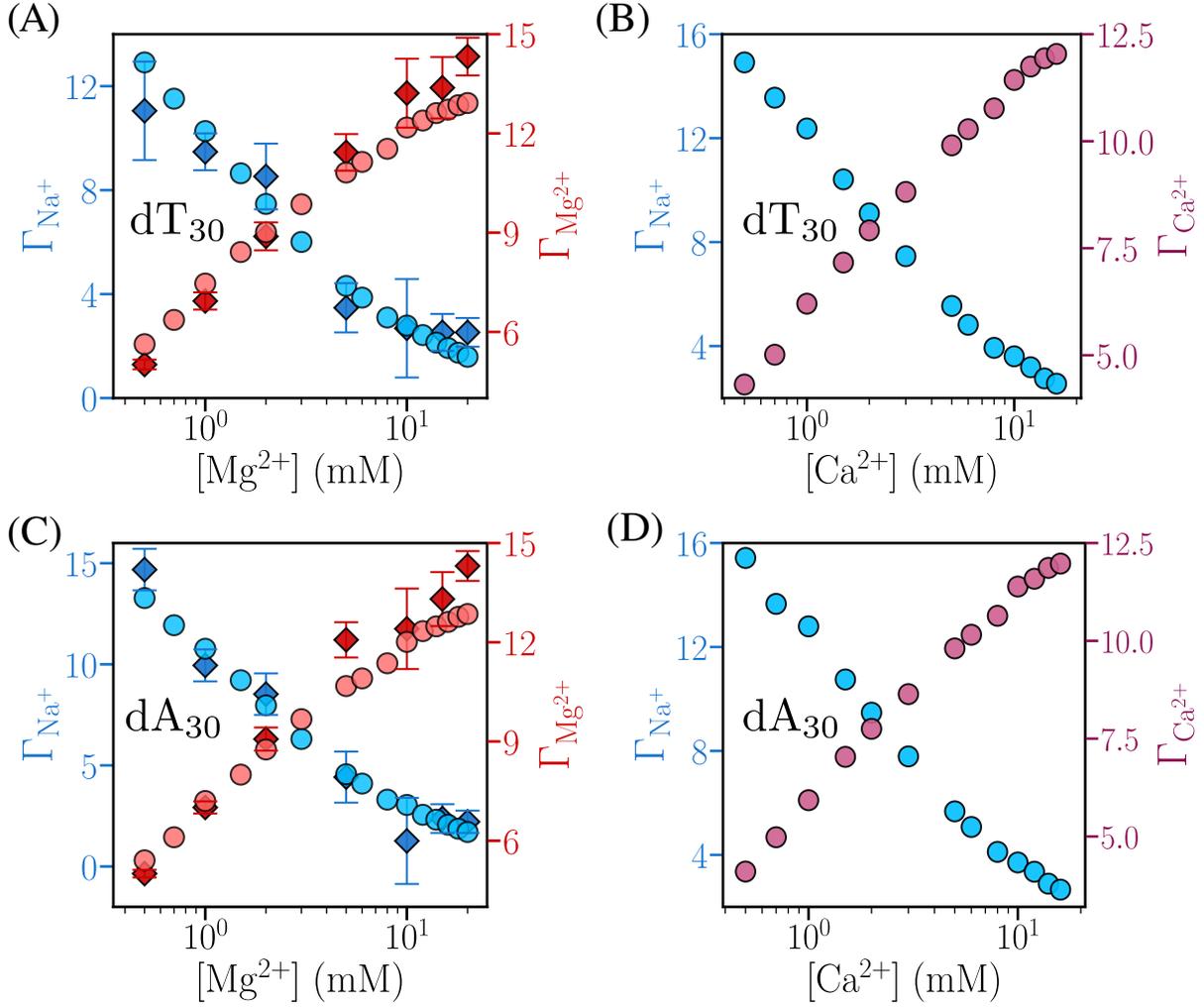


Figure 3: **Ions around ssDNA.** (A) Excess number of ions,  $\Gamma_{\text{Na}^+}$  (in blue) and  $\Gamma_{\text{Mg}^{2+}}$  (in red), in the vicinity of ssDNA sequence dT<sub>30</sub> as a function of  $[\text{Mg}^{2+}]$ . Experimental results<sup>10</sup> are in the dark shade diamonds. Simulation results are in circles in light shade. (B) Same as (A), except the results are for  $\text{Ca}^{2+}$  ions (in magenta). (C) Same as (A), except results are for the dA<sub>30</sub> sequence. (D) Same as (C), except results are for  $\text{Ca}^{2+}$  ions.  $[\text{Na}^+] = 20$  mM in the panels.

coefficient cannot be used to determine sequence effects, at least in short homopolymeric ssDNA sequences. As demonstrated below, notable distinctions exist in the conformational characteristics of dT<sub>30</sub> and dA<sub>30</sub>.

Comparison of the results for Mg<sup>2+</sup> localization around dsDNA, dT<sub>30</sub> and dA<sub>30</sub> allows us to infer the following observations: (1) The Mg<sup>2+</sup> concentration, at which Na<sup>+</sup> is released is  $\approx 1$  mM for dsDNA, whereas it is  $\approx 2.5$  mM for dT<sub>30</sub> and dA<sub>30</sub>. This is likely due to the large difference in flexibility between dsDNA and ssDNA. (2) Interestingly, the number of cations around dT<sub>30</sub> and dA<sub>30</sub>, as a function of [Mg<sup>2+</sup>] is almost the same although there is a significant difference in their conformational fluctuations (see below). (3) Somewhat surprisingly, our simulations as well as experiments,<sup>6</sup> predict that  $\Gamma_{\text{Mg}^{2+}} \approx \Gamma_{\text{Ca}^{2+}}$ , at all values of the divalent ion concentration even though their charge densities are different.

**Agreement between calculated and experimental SAXS profiles is excellent:**

We calculated the SAXS profiles,  $I(q)$ , for a 25 bp dsDNA (see Figure 4A for the sequence), dT<sub>30</sub> and dA<sub>30</sub> at  $T = 298$  K, as a function of the wave vector  $q$ , using simulations in order to compare with experiments.<sup>7,10,45</sup> The experimental SAXS profiles<sup>10</sup> were accessed either from the publicly available SAXS database SASDBD (dT<sub>30</sub>: SASDBD6, dA<sub>30</sub>: SASDBE6, in 20 mM NaCl solution) or extracted using the WebPlotDigitizer<sup>54</sup> from the published data (dT<sub>30</sub> in salt solution containing 20 mM Mg<sup>2+</sup> and 20 mM NaCl,<sup>7</sup> dsDNA in 3 mM and 16 mM Mg<sup>2+</sup>, and 0.4 mM NaCl<sup>45</sup>). We calculated  $I(q)$  using at least 5,000 randomly chosen conformations using the TIS2AA software<sup>55</sup> and CRY SOL.<sup>56</sup> Before calculating  $I(q)$ , we converted the structures generated in the coarse-grained simulations to all-atom conformations using an in-house code. The resulting ensemble of all-atom structures proves to be convenient in the calculation of SAXS profiles.

**dsDNA:** We first calculated  $I(q)$  as a function of  $q$  at 3 mM and 16 mM [Mg<sup>2+</sup>], for the 25 bp dsDNA for which experimental<sup>45</sup> data is available. The results in Figures 4B and 4C show that the simulations reproduce the experimental results almost quantitatively. There are only minor deviations at  $q > 2 \text{ nm}^{-1}$ . Few of the reconstructed all-atom structures used

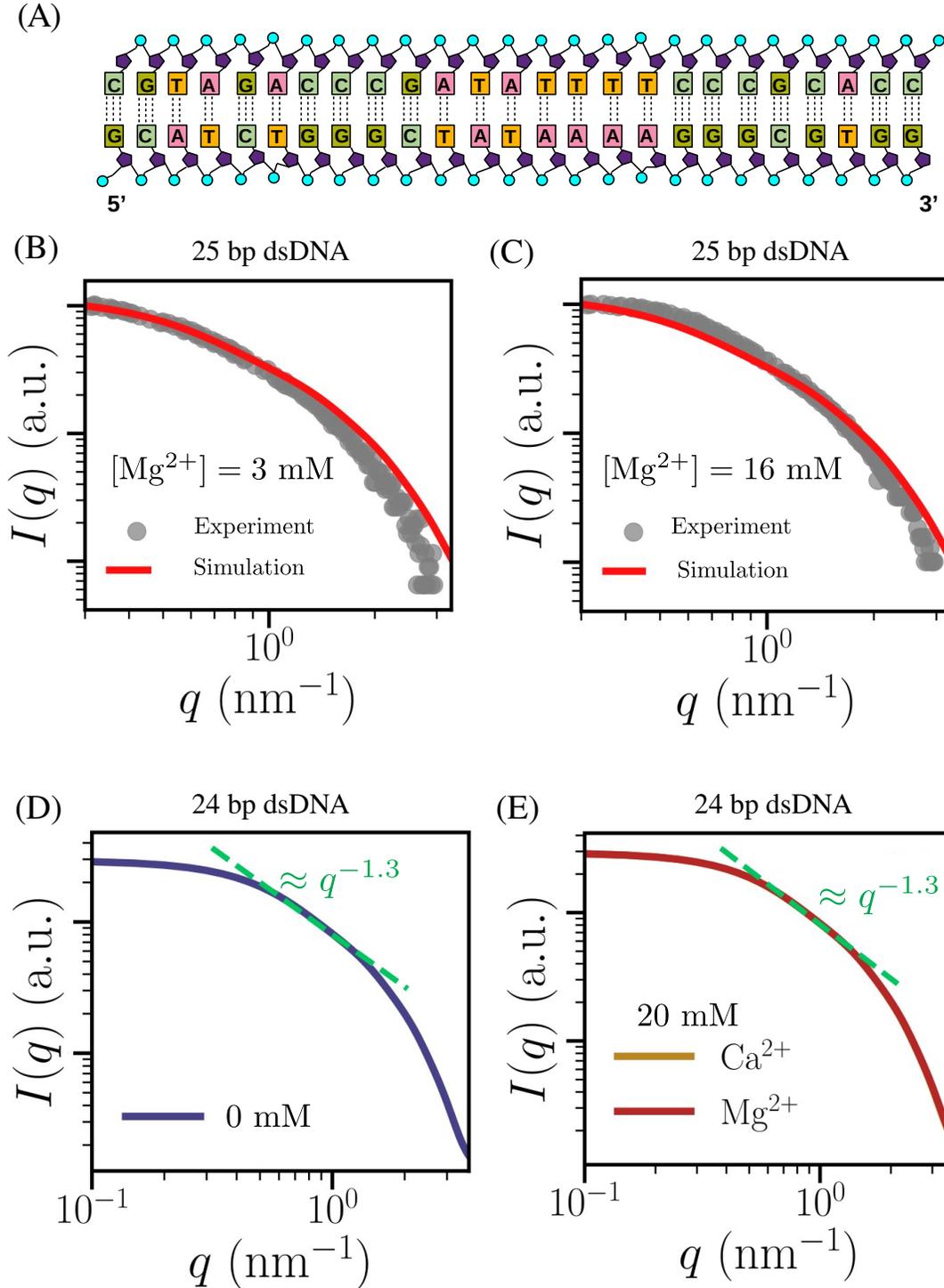


Figure 4: **Simulated SAXS profiles of dsDNA.** (A) Sequence of the 25-bp dsDNA. (B) Comparison of experimental<sup>45</sup> and simulated SAXS profiles in  $[\text{Mg}^{2+}] = 3 \text{ mM}$  and  $[\text{Na}^+] = 0.4 \text{ mM}$  at  $T = 298 \text{ K}$ . (C) Same as (B), except the results are for  $[\text{Mg}^{2+}] = 16 \text{ mM}$ . (D) SAXS profile of the 24 bp dsDNA (Figure 1(A)) in  $[\text{Na}^+] = 20 \text{ mM}$  at  $T = 277 \text{ K}$ . (E) Comparison of SAXS profiles of the same dsDNA sequence, in  $[\text{Mg}^{2+}] = 20 \text{ mM}$  and  $[\text{Ca}^{2+}] = 20 \text{ mM}$ , in  $[\text{Na}^+] = 20 \text{ mM}$ .

to calculate the SAXS profiles for the 25 bp dsDNA are shown in Figure S4 in the SI.

Measurements of  $I(q)$  are not available for the dsDNA sequence studied in the ion counting experiments. Nevertheless, the accuracy of our simulations allows us to predict the outcomes and compare the results for  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$ . From Figures 4D and 4E, which show  $I(q)$  as a function of  $q$  at 0 mM and 20 mM divalent cation concentrations, it is clear that the SAXS profiles are almost identical for both the ions. The length of the dsDNA (24 bp) is  $\approx 6$  times smaller than the bare persistence length of a typical dsDNA (150 bp). The dominance of the bare persistence length in the short dsDNA, due to WC base pairing, implies that the 24-nucleotide sequence should be impervious to cations, thus explaining the  $I(q)$  results in Figure 4. Scaling of  $I(q)$  with  $q$  are known for various polymer shapes. For example,  $I(q) \propto q^{-x}$ , where  $x = 2$  (Gaussian chains),  $x = 1$  (rods),  $x = 5/3$  (polymer in good solvent), and  $x = 4$  (for globules).<sup>57</sup> The dependence of  $I(q) \sim q^{-1.3}$  at the intermediate  $q$  values (Figure 4D, 4E), suggests that the dsDNA is not a rod, but behaves like a stiff polymer.<sup>57</sup>

**SAXS profiles for ssDNAs show scale-dependent structural changes:** Figure 5, with comparisons to experiments in the insets, displays the calculated SAXS profiles for dT<sub>30</sub> and dA<sub>30</sub>. The simulated and the experimental SAXS profiles are in excellent agreement. The variations of  $I(q)$  with  $q$  reveals details about the changes in the structure of the ssDNA sequences at different length scales. At  $[\text{X}^{2+}] = 0$  mM, in the range  $q \sim 0.8 - 1.5 \text{ nm}^{-1}$ ,  $x \approx 1.4$  for dT<sub>30</sub>. In contrast, for dA<sub>30</sub>,  $x \approx 1.6$ , which is close to the value expected for polymers in good solvents. The value of  $x$  increases for dT<sub>30</sub> ( $x \approx 1.8$ ), which implies that there is change in conformation from a stiff to more flexible polymer behavior as  $[\text{X}^{2+}]$  increases. In contrast,  $x \approx 1.6$  is roughly the same for dA<sub>30</sub> at the elevated  $[\text{X}^{2+}]$ . Examples of the reconstructed all-atom structures, used to calculate SAXS profiles for dT<sub>30</sub> and dA<sub>30</sub>, are shown in Figures 5E, S5 and S6, respectively.

**Stacking versus electrostatic interactions:** It can be argued that the differences in  $x$  values in dT<sub>30</sub> and dA<sub>30</sub> as  $[\text{X}^{2+}]$  increases should be a consequence of the interplay between

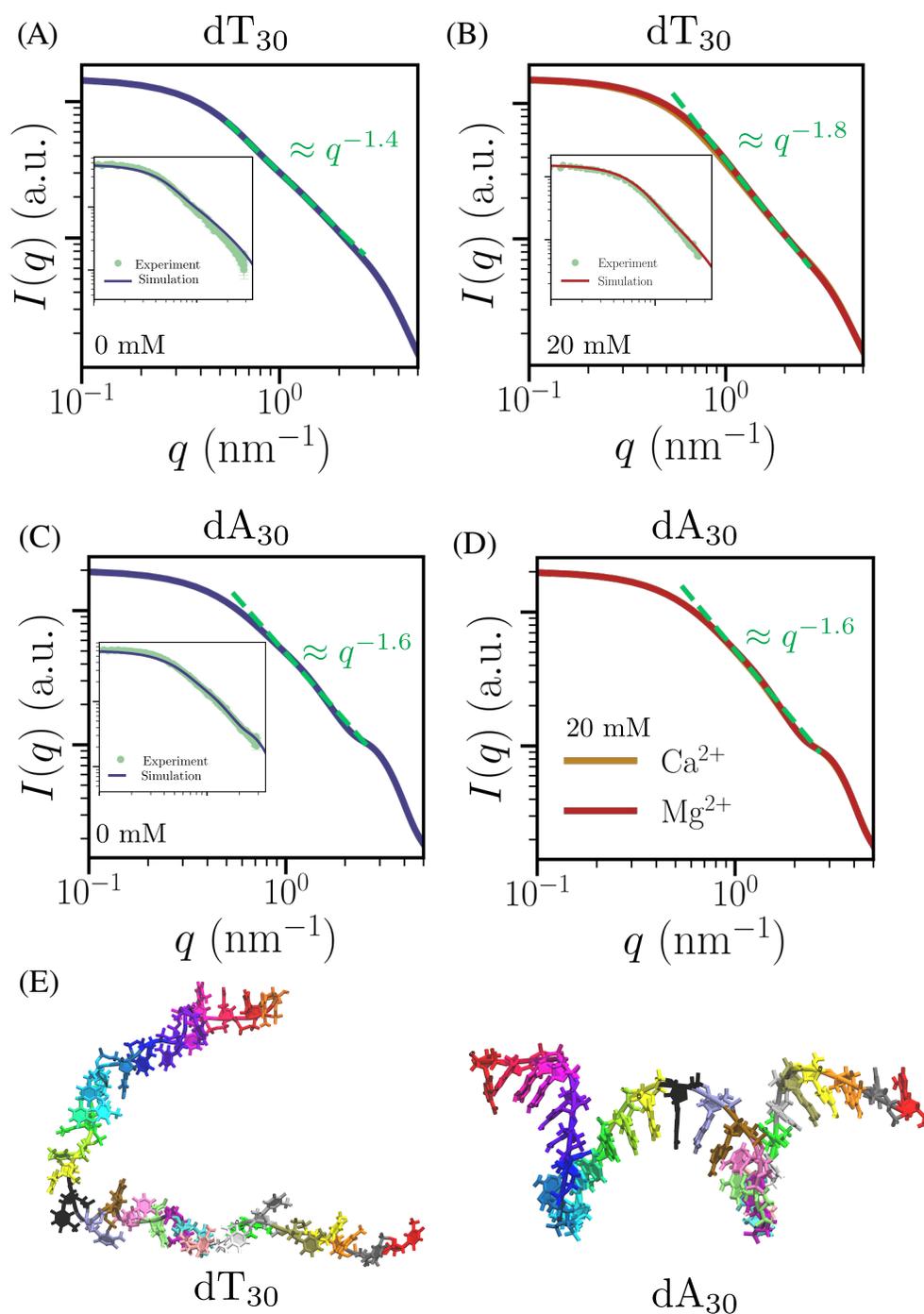


Figure 5: **Comparison of simulated and experimental SAXS profiles for ssDNA.** (A-D) SAXS profiles are computed for dT<sub>30</sub> and dA<sub>30</sub>, in [Mg<sup>2+</sup>] or [Ca<sup>2+</sup>], at 0 mM and 20 mM, in [Na<sup>+</sup>] = 20 mM at  $T = 277$  K. Inset compares the simulated and experimental SAXS profiles with error bars, when available, at  $T = 298$  K. (E) Representative all-atom structures, reconstructed from coarse-grained simulations, are shown at [Mg<sup>2+</sup>] = 20 mM. To aid visualization, nucleotides are assigned different colors based on their sequence number.

favorable stacking and cation-mediated electrostatic repulsion between the phosphate groups. In dA<sub>30</sub> the stacking interactions dominate, and essentially determine the conformations of the ssDNA. Indeed, the favorable stacking interaction ( $\langle U_{ST} \rangle$ ), is more stabilizing than the effective repulsion between the phosphate groups ( $\langle U_{EL}^P \rangle$ ) in dA<sub>30</sub> (see Figure S7 in SI). Thus,  $\langle U_{ST} \rangle$ , with  $\langle U_{EL}^P \rangle$  playing a minor role, determines the structural ensemble of dA<sub>30</sub>, which explains the constancy of  $x$  with increasing [Mg<sup>2+</sup>]. In contrast, the values of  $\langle U_{ST} \rangle$  and  $\langle U_{EL}^P \rangle$  are comparable in dT<sub>30</sub>. As a result, thermal fluctuations play an important role in shaping the conformations of dT<sub>30</sub>. At low divalent concentrations, there is a preference to populate stiff conformations, which become more flexible at the highest divalent concentrations. The enhanced preference for stacking in Adenine compared to Thymine rationalizes the simulation results. A natural prediction is that balance between  $\langle U_{ST} \rangle$  and  $\langle U_{EL}^P \rangle$  may be altered in dT<sub>30</sub> at lower temperatures.

**Distribution of  $R_g$ :** The results in Figure 3 show that there are little variations in  $\Gamma_{X^{2+}}$  as a function of  $[X^{2+}]$ , despite the known differences on the stacking interactions between A and T.<sup>9</sup> The sequence-dependent variations are reflected in the distributions  $P(R_g)$  of the radius of gyration,  $R_g$  (Figures 6A-D). The dispersion in  $P(R_g)$  changes considerably as [Mg<sup>2+</sup>] or [Ca<sup>2+</sup>] is increased. At  $[X^{2+}] = 20$  mM, dT<sub>30</sub> samples a broad range of conformations (Figures 6A and 6B). Although qualitatively similar results are found for dA<sub>30</sub> (Figures 6C and 6D), the changes in the width of  $P(R_g)$  with increase in  $[X^{2+}]$  is less pronounced.

Comparison of the calculated and extracted mean  $\langle R_g \rangle$  from experimental SAXS profiles<sup>7</sup> for dT<sub>30</sub> at  $T = 298$  K (see inset in Figure 6E), shows very good agreement at all values of  $[X^{2+}]$ . Although the simulations predict a somewhat larger degree of compaction, the differences between experiments and simulations are small. The  $\langle R_g \rangle$  values in Ca<sup>2+</sup> are higher than in Mg<sup>2+</sup> for both dT<sub>30</sub> and dA<sub>30</sub> (Figures 6E and 6F), which in effect means that Mg<sup>2+</sup> induces more compaction because of the higher charge density.

**Effective Flory exponent:** In order to elucidate the polymer properties of the ssDNAs,

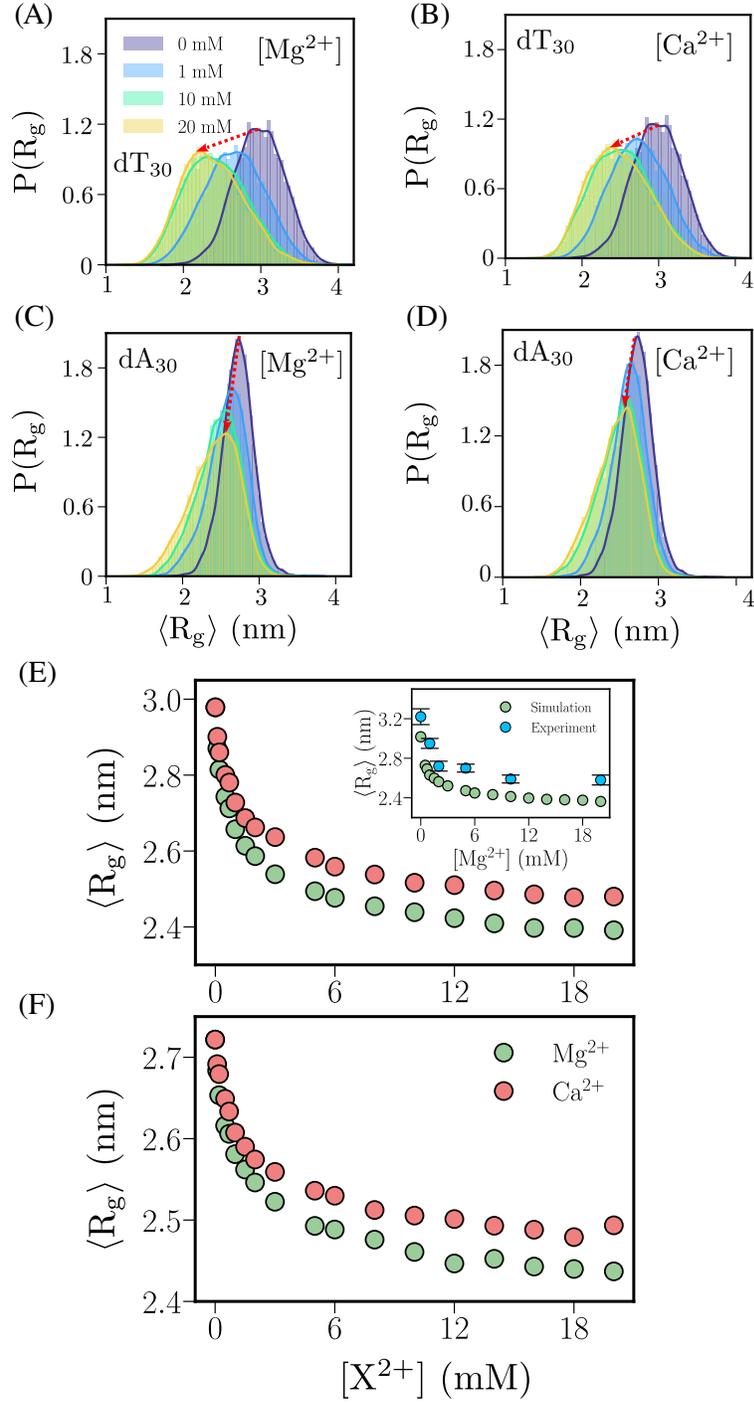


Figure 6:  $[X^{2+}]$  **dependent changes in the radius of gyration.** (A-D) Normalized distribution of  $R_g$  for dT<sub>30</sub> and dA<sub>30</sub>, as a function of  $[Mg^{2+}]$  and  $[Ca^{2+}]$ . Movement of the peak of the distributions are indicated with arrows. (E) Mean value,  $\langle R_g \rangle$  of dT<sub>30</sub> as a function of  $[X^{2+}]$ , where  $X^{2+} = Mg^{2+}$  or  $Ca^{2+}$ , in  $[Na^+] = 20$  mM at  $T = 277$  K. Inset compares the calculated and experimentally measured<sup>7</sup>  $\langle R_g \rangle$  for dT<sub>30</sub> at 298 K in 20 mM NaCl solution. (F) Same as (E), except the results are for dA<sub>30</sub> sequence.

we fit the calculated mean distance  $R_{ij}$  between the phosphate groups  $i$  and  $j$ , to the equation,

$$R_{ij} = R_0|i - j|^{\nu_{eff}}, \quad (10)$$

where  $R_0$  is the prefactor, and  $\nu_{eff}$  is the effective Flory exponent. Because of the limited range of  $|i - j|$  (Figure 7), the extracted values of  $\nu_{eff}$  and  $R_0$  might not be accurate. Nevertheless, Figures 7A-7D show that  $\ln R_{ij} \approx \nu_{eff} \ln |i - j|$  at all  $[X^{2+}]$  is linear, thus allowing to extract  $\nu_{eff}$  as a function of  $[Mg^{2+}]$  (Figure 7E) and  $[Ca^{2+}]$  (Figure 7F). The plots clearly show the influence of electrostatic interactions.

A few observations pertaining to the internal structures follow from Figures 7E and 7F. (1) The values of  $\nu_{eff}$  decrease linearly as  $[X^{2+}]$  increases in both the sequences. At all  $[X^{2+}]$ ,  $\nu_{eff}$  for dA<sub>30</sub> is significantly larger than dT<sub>30</sub> in both the divalent cations. (2) At higher  $[X^{2+}]$  ( $= 20$  mM),  $\nu_{eff} \approx 0.6$  for dT<sub>30</sub>, which is the characteristic value for a homopolymer in a good solvent. Interestingly, only at  $\approx 1$  M monovalent salt concentration,  $\nu_{eff} \approx 0.6$ <sup>22,58</sup> for poly dT. (3) Because  $\nu_{eff}$  for dA<sub>30</sub> exceeds 0.6 at all  $[X^{2+}]$ , it follows that poly dA sequences are likely to behave as semiflexible chains. The stiffness of dA<sub>30</sub> can be attributed to the propensity of adenosine bases to form stable stacks. This is consistent with the minimal changes observed in SAXS scaling relations ( $x = 1.6$ ) for dA<sub>30</sub>, as divalent cation concentrations are increased from 0 to 20 mM. (4) The prefactor  $R_0$  for dT<sub>30</sub> in  $Mg^{2+}$  varies from 0.7 nm - 0.8 nm in the concentration range (0 - 20) mM (Table S5 in the SI). The values of  $R_0 \approx (0.5 - 0.6)$  nm is smaller for dA<sub>30</sub> (Table S5 in SI). The prefactor is a measure of the mean distance between the nucleobases. The smaller value of  $R_0$  in dA<sub>30</sub> is due to the more favorable stacking interactions in Adenine compared to Thymine. A similar trend is observed in presence of  $Ca^{2+}$  as well (see Table S6 in SI).

**Stacking interactions stiffen dA<sub>30</sub>:** The procedure used to calculate  $\langle p_S \rangle$  is illustrated in Figures 8A and 8B. The distributions of the longest stacked ( $N_S^{max}$ ) and unstacked ( $N_U^{max}$ ) bases in  $[Mg^{2+}] = 20$  mM, are shown in Figures 8C and 8D, respectively. For dT<sub>30</sub>,  $\langle N_S^{max} \rangle$

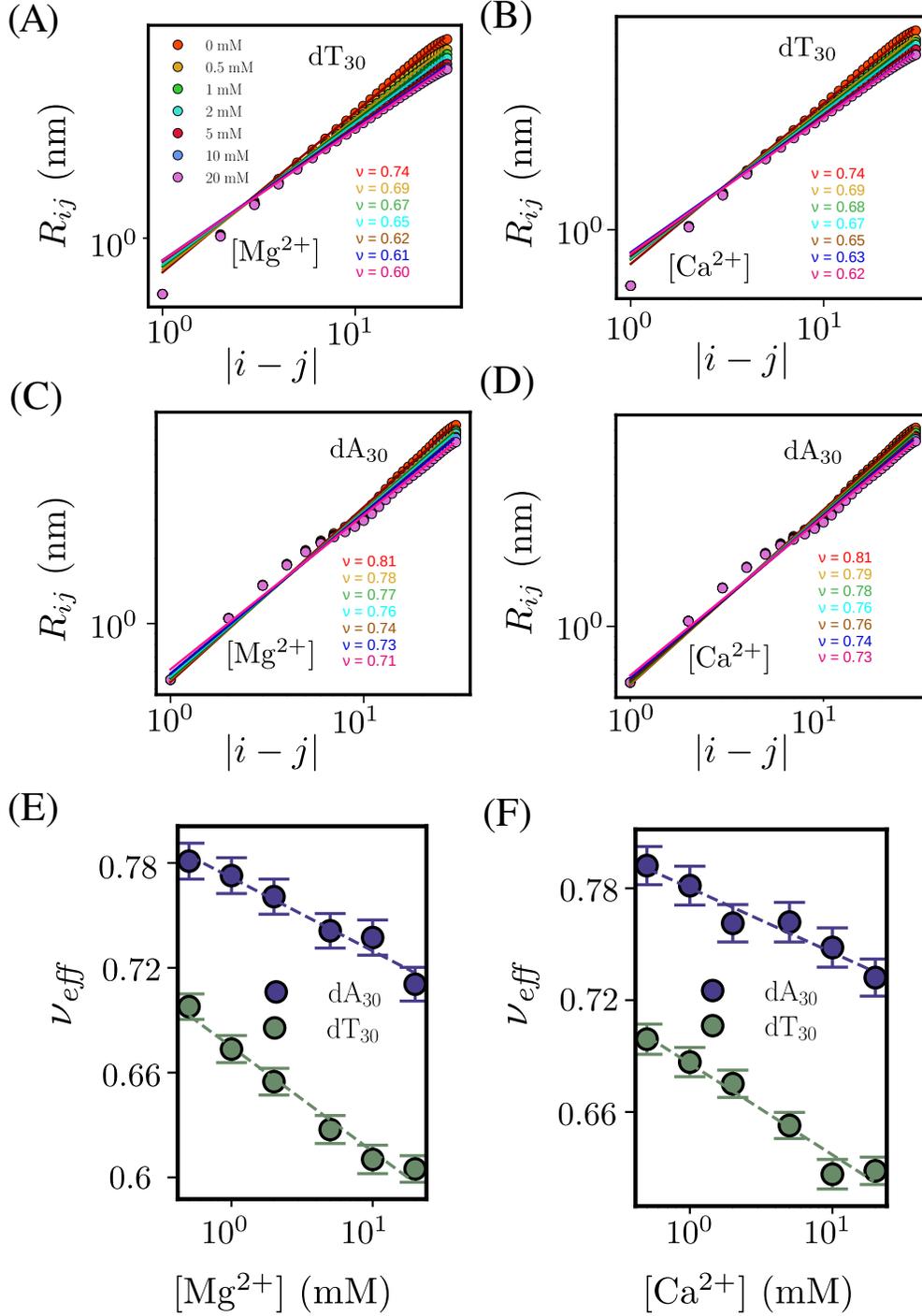


Figure 7: **Effective Flory exponent.** (A) Distance between the phosphate groups of dT<sub>30</sub>,  $R_{ij}$ , as a function of nucleotide separation,  $|i-j|$ , along the contour, at different [Mg<sup>2+</sup>], in [Na<sup>+</sup>] = 20 mM. (B) Same as (A), except the results are for [Ca<sup>2+</sup>]. (C) Same as (A), except the results are for dA<sub>30</sub> sequence. (D) Same as (C), except the results are for [Ca<sup>2+</sup>]. (E) Effective Flory exponent,  $\nu_{eff}$  as a function of [Mg<sup>2+</sup>], for dT<sub>30</sub> and dA<sub>30</sub>. (F) Same as (E), except the results are for [Ca<sup>2+</sup>].

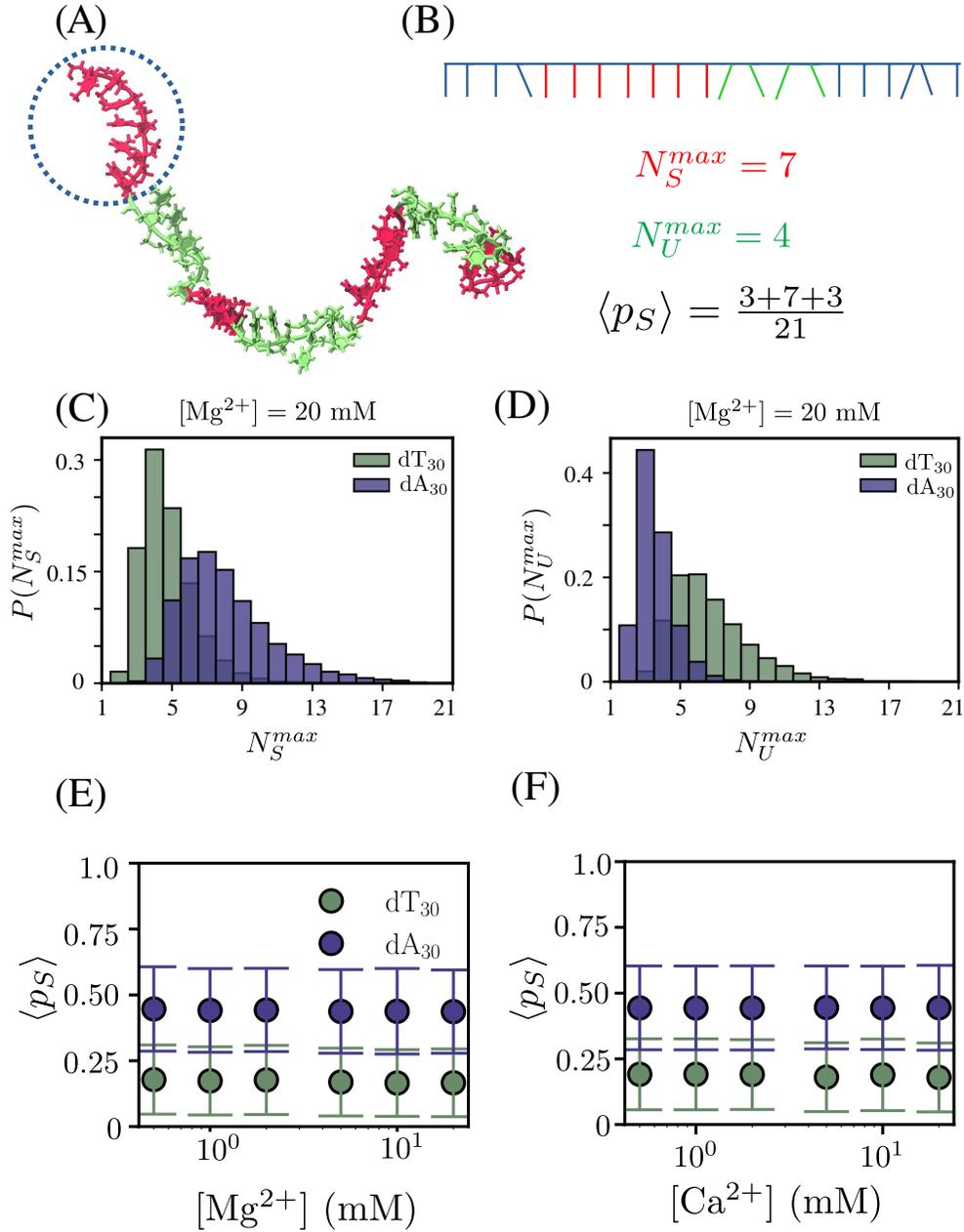


Figure 8: **Stacking propensities in dT<sub>30</sub> and dA<sub>30</sub>.** (A) A simulation snapshot of dT<sub>30</sub> chain highlighting a region of persistent stacks composed of six nucleotides. Stacks are shown in red. Unstacked regions are in green. (B) Schematic of a ssDNA segment depicting a stretch of the longest stacked bases,  $N_S^{max}$  (in red), stretch of longest unstacked bases,  $N_U^{max}$  (in green), and  $\langle p_S \rangle$ . (C) Normalized distributions of  $N_S^{max}$  in dT<sub>30</sub> and dA<sub>30</sub> at  $[Mg^{2+}] = 20$  mM. (D) Same as (C), except the distributions correspond to  $N_U^{max}$ . (E) Ensemble averaged values of the fraction of persistent stacks,  $\langle p_S \rangle$ , as a function of  $[Mg^{2+}]$ . (F) Same as (E), except the results are for  $[Ca^{2+}]$ .

$\approx 4$  nucleotides (nt), whereas for dA<sub>30</sub>, the peak appears at  $N_S^{max} \approx 7$  nt, with a tail that extends to  $N_S^{max} \approx 16$  nt. The longest unstacked segments in dA<sub>30</sub> has a narrower distribution with  $\langle N_U^{max} \rangle \approx 3$  nt compared to dT<sub>30</sub>, for which the distribution extends to  $\approx 16$  nt. Long helical stacks are interrupted by short unstacked bases in dA<sub>30</sub>, whereas shorter helical stacks are interrupted by long-stretches of unstacked segments in dT<sub>30</sub>. For instance, the probability of finding dT<sub>30</sub> conformations with  $N_U^{max} = 10$  nt, which is the periodicity of a B-DNA helix, is  $\approx 5\%$  whereas it is only  $0.02\%$  (a decrease by a factor of 250) in dA<sub>30</sub>. Figure S8 compares the distribution  $P(N_S^{max})$  and  $P(N_U^{max})$  at  $[\text{Mg}^{2+}] = 0$  mM and 20 mM, in dT<sub>30</sub> and dA<sub>30</sub>. Figures 8E and 8F show that  $\langle p_S^{dA_{30}} \rangle \approx 2\langle p_S^{dT_{30}} \rangle$  at all  $[\text{X}^{2+}]$ , and is independent of the nature of the cation. This finding explains the increased stiffness of dA<sub>30</sub> relative to dT<sub>30</sub>, which accords well with the observation that Adenine bases participate in stronger base stacking compared to Thymine.<sup>7,9,10,59</sup> The independence of  $\langle p_S \rangle$  on  $[\text{X}^{2+}]$  finds support in the magnetic tweezer experiments<sup>9</sup> that report nearly constant free energy difference between stacked poly(dA) and unstacked polypyrimidine over three orders of magnitude in NaCl concentration.

**Spatial distribution of ions around dsDNA:** The finding that  $\Gamma_{\text{X}^{2+}}$  is the same for  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  (Figure 2) is surprising because one would expect that the number of smaller  $\text{Mg}^{2+}$  around the dsDNA should be larger than  $\text{Ca}^{2+}$ . Hence, it is natural to ask if there are any discernible differences in the distributions of  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  ions around stiff dsDNAs at the nucleotide level which are masked in ion-counting and SAXS experiments. To answer this question we projected the ensemble averaged occupancy of ions onto the frontal plane (see Figure 9A) passing through the dsDNA at  $[\text{X}^{2+}] = 20$  mM. The two-dimensional projection of  $\text{Mg}^{2+}$  occupancy (Figure 9B) shows the cylindrical arrangement of  $\text{Mg}^{2+}$  ions in multiple layers surrounding the dsDNA. The inner-most layer contains a nearly uniform density of ions at a distance  $\approx 0.85$  nm from the center line of dsDNA, indicating non-specific binding, while ions are excluded from the backbone due to the finite sizes of DNA and ions. Unlike  $\text{Mg}^{2+}$ , a relatively sparse layer of  $\text{Ca}^{2+}$  ions is localized at  $\approx 0.95$  nm from the center

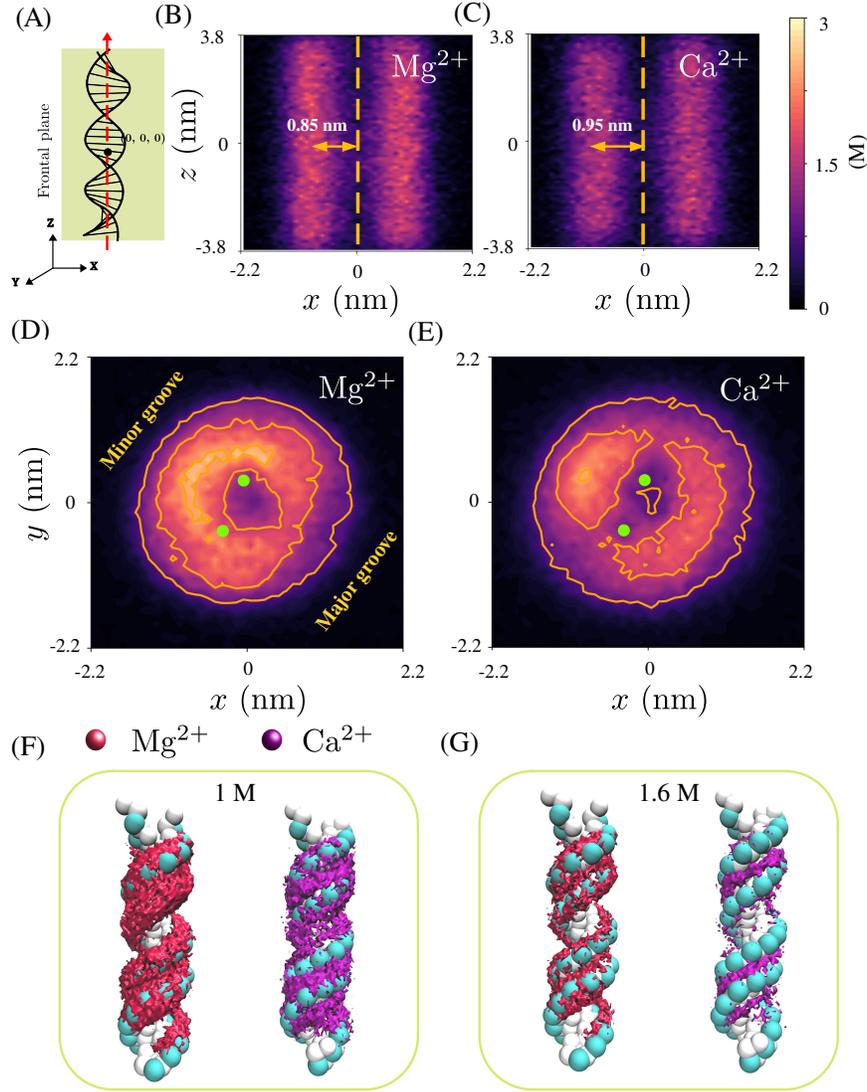


Figure 9: **Two-dimensional spatial distribution of divalent cations around dsDNA.**

(A) Illustration of the frontal plane passing through the dsDNA onto which ion occupancy is projected. Frontal plane is projected along the major principal axis,  $z$ , and the minor principal axis,  $x$ . The major principal axis is shown in red dashed line, passing through the origin. (B) Spatial distribution profile for  $\text{Mg}^{2+}$  ions projected onto the frontal plane passing through the dsDNA. (C) Same as (B), except the results are for  $\text{Ca}^{2+}$  ions. The local concentration scales are shown on the right. (D) 2D untwisted density of  $\text{Mg}^{2+}$  ions projected onto an average base-pair plane obtained by removing the helical twist,<sup>12</sup> in  $[\text{Mg}^{2+}] = 20 \text{ mM}$ . Averaged locations of the phosphate groups are shown in green. Major and minor grooves are highlighted. (E) Same as (D), except results are shown for  $\text{Ca}^{2+}$  ions. (F) Ensemble average occupancy of  $\text{Mg}^{2+}$  (in red) and  $\text{Ca}^{2+}$  (in magenta) around dsDNA, in  $[\text{X}^{2+}] = 20 \text{ mM}$ . Isosurfaces are for grids with  $\rho_i = 1 \text{ M}$ . Phosphate beads are highlighted in cyan. (G) Same as (F), except isosurfaces are for grids with  $\rho_i = 1.6 \text{ M}$ .

of line of dsDNA. In addition,  $\text{Ca}^{2+}$  ions are excluded to a greater extent from the dsDNA backbone compared to  $\text{Mg}^{2+}$  because of their larger size. Local ion concentrations  $C_{\text{Na}^+}(r)$ ,  $C_{\text{Mg}^{2+}}(r)$  and  $C_{\text{Ca}^{2+}}(r)$ , calculated using Eq. S1, are shown in the Figure S9 in the SI, also recapitulates these differences in ion binding. In agreement with the ion occupancy results, the highest  $[\text{Mg}^{2+}]$  in the vicinity of the dsDNA is 520 mM at  $r \approx 0.85$  nm, whereas the highest  $[\text{Ca}^{2+}] \approx 450$  mM at  $r \approx 0.95$  nm.

The differences in spatial binding of  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  with phosphates, may be rationalized using the size differences alone. The size of  $\text{Ca}^{2+}$  ion ( $R \approx 0.17$  nm) is larger than  $\text{Mg}^{2+}$  ions ( $R \approx 0.08$  nm). As a result, the volume excluded by  $\text{Ca}^{2+}$  is eight times larger than  $\text{Mg}^{2+}$ . The larger size also impacts the surface charge density of  $\text{Ca}^{2+}$ , which is 4.5 times smaller than  $\text{Mg}^{2+}$ , resulting in considerably weaker binding with the phosphate groups. Unlike the divalent cations, which show  $\approx 20$  fold increase in local concentration around dsDNA compared to bulk, maximum  $[\text{Na}^+] \approx 45$  mM in  $\text{Mg}^{2+}$  solution, and  $\approx 55$  mM in  $\text{Ca}^{2+}$  solution at  $r \approx 1.15$  nm. This is a consequence of the lower charge density of the monovalent cations.

**Binding patterns of  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  to the grooves of dsDNA:** To discern the groove specific binding of the divalent cations, the density of  $\text{X}^{2+}$  ions is projected onto  $xy$  plane of a base-pair. This is achieved by removing the helical twist angle for each base-pair plane, that is perpendicular to the helical axis ( $z$ ), following a strategy discussed previously.<sup>12</sup> Such a representation successfully differentiates between the binding of the ions to the major and minor grooves. The resulting 2D densities of  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  ions are shown in Figures 9D and 9E, respectively, in  $[\text{X}^{2+}] = 20$  mM. Figure 9D vividly shows that  $\text{Mg}^{2+}$  binds to both the minor groove and the phosphates, while  $\text{Ca}^{2+}$  binds specifically to the minor groove. This binding pattern remains unchanged at all other  $[\text{X}^{2+}]$  (Figure S10).

We also calculated the spatial occupancy of  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  ions surrounding DNA using three-dimensional (3D) cubic grid of length 1 Å. The number density corresponding to a grid  $i$  is computed as,  $\rho_i = \frac{\langle N_i \rangle}{V_i}$ , where  $\langle N_i \rangle$  is the average number of  $\text{X}^{2+}$  ions occupying

the  $i^{\text{th}}$  grid with volume  $V_i$ . Isosurfaces are drawn for  $\text{Mg}^{2+}$  (in red) and  $\text{Ca}^{2+}$  (in magenta) ions. Figure 9F and 9G show two isosurfaces highlighting grids ( $i$ ) with  $\rho_i = 1$  M and 1.6 M, respectively. At  $\rho_i = 1$  M (Figure 9F), binding of  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  is almost uniform along the dsDNA cylinder. In contrast, at  $\rho_i = 1.6$  M,  $\text{Ca}^{2+}$  ions are preferentially localized near the minor groove of DNA, whereas  $\text{Mg}^{2+}$  ions bind directly with the phosphate groups (Figure 9G), in addition to the minor grooves.

The binding modes are determined by the excluded volume of the divalent cations. The distance between adjacent phosphate groups belonging to the same strand is  $\approx 0.6$  nm. Due to smaller size of  $\text{Mg}^{2+}$  ions, it can directly interact with two phosphate groups (-1 e charge each), unlike  $\text{Ca}^{2+}$  ions. On the other hand,  $\text{Ca}^{2+}$  ions binding to the minor groove is dictated by the local phosphate ion concentration and the groove dimensions. In a typical B-DNA, the width of minor and major grooves are  $\approx 1.2$  nm and  $\approx 2.2$  nm, respectively. Due to close proximity of phosphate groups from two strands in the minor grooves, negative charge density is higher in the minor grooves compared to the major grooves. In addition, the width of the minor groove ( $\approx 1.2$  nm) is sufficient to accommodate multiple  $\text{Ca}^{2+}$  ions which can neutralize phosphates from complementary strands simultaneously (Figure 9G).

**Spatial distribution of ions around ssDNA:** To elucidate the spatial organization of ions along the ssDNA contour, we calculated a joint distribution function of two order parameters,  $r_z$  and  $r_{\text{X}^{2+}\text{-P}}^{\text{Min}}$ ,  $P(r_z, r_{\text{X}^{2+}\text{-P}}^{\text{Min}})$ , where  $r_{\text{X}^{2+}\text{-P}}^{\text{Min}}$  is the distance of closest approach between  $\text{X}^{2+}$  ions and the phosphate beads, and  $r_z$  is the  $z$ -component of the distance of the  $\text{X}^{2+}$  ion from the center of the major principal axis (presumed to lie along the  $z$  axis) of the ssDNA (see Figure 10A). To compare the spatial arrangement of  $\text{Mg}^{2+}$  ions around dT<sub>30</sub> and dA<sub>30</sub>, we computed the difference in the joint distribution functions,  $\Delta_{\text{Mg}^{2+}}^{\text{dT}_{30}\text{-dA}_{30}} (= P_{\text{dT}_{30}}(r_z, r_{\text{Mg}^{2+}\text{-P}}^{\text{Min}}) - P_{\text{dA}_{30}}(r_z, r_{\text{Mg}^{2+}\text{-P}}^{\text{Min}}))$  at 20 mM  $\text{Mg}^{2+}$  (Figure 10B, S11). As is evident,  $\text{Mg}^{2+}$  ions preferentially accumulate along the contour of dT<sub>30</sub> as compared to dA<sub>30</sub>. Lack of persistent structural order increases the local phosphate charge density in dT<sub>30</sub>, as opposed to dA<sub>30</sub>, which is rich in helical motifs (Figures 6, 7). A similar trend is observed for  $\text{Ca}^{2+}$

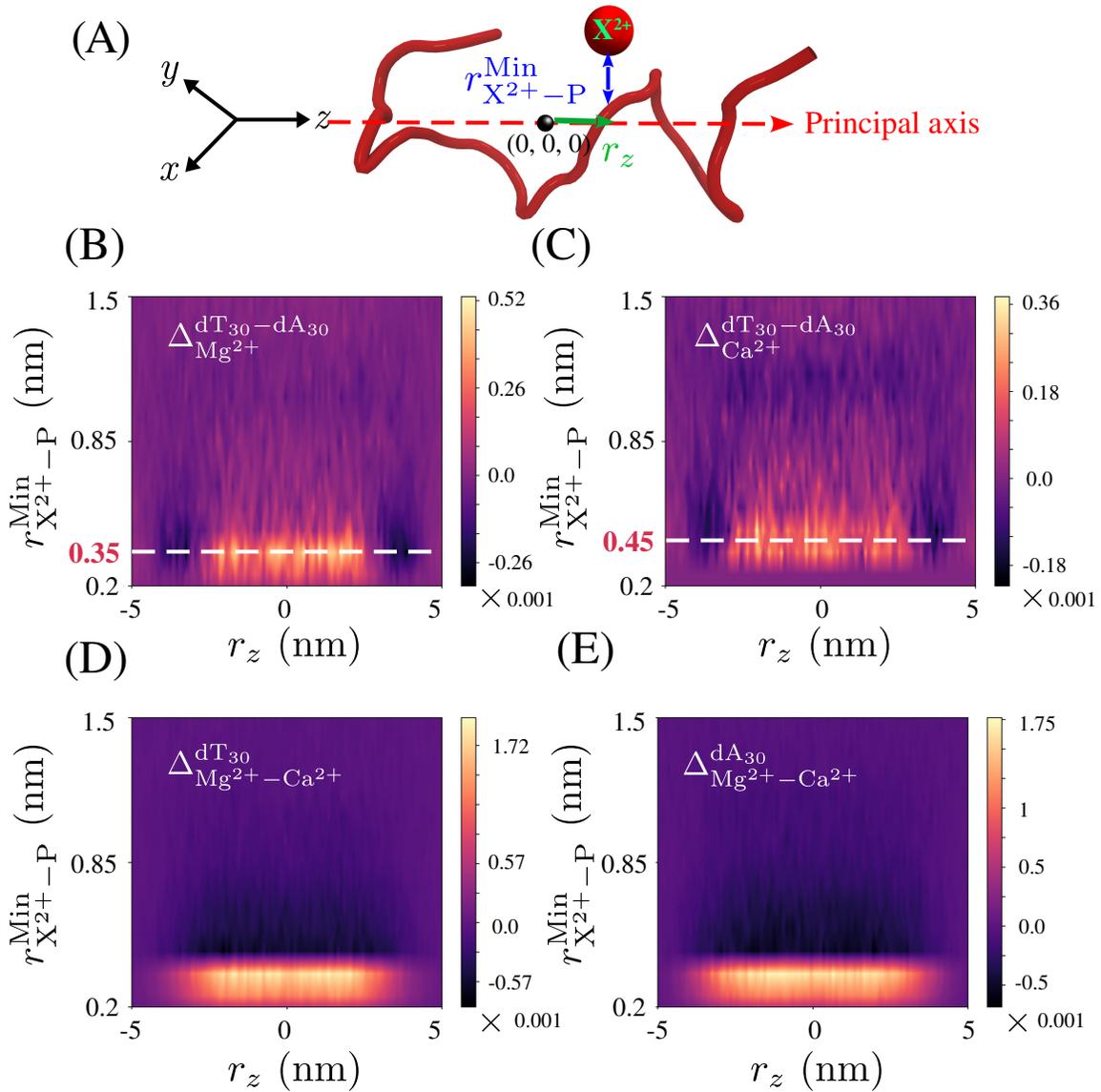


Figure 10: **Two-dimensional spatial distribution of divalent cations around ssDNA.** (A) Definition of  $r_{X^{2+}-P}^{\text{Min}}$  and  $r_z$ , using a typical ssDNA construct. Major principal axis (red dashed line) is aligned along the  $z$  axis. Ion  $X^{2+}$  is in red,  $r_{X^{2+}-P}^{\text{Min}}$  is the distance of closest approach between the phosphate and  $X^{2+}$ , and  $r_z$  is the  $z$ -component of distance of  $X^{2+}$  from the center of the major principal axis of ssDNA. (B) Difference in the normalized joint distribution functions,  $P(r_z, r_{X^{2+}-P}^{\text{Min}})$ , for dT<sub>30</sub> and dA<sub>30</sub>,  $\Delta_{\text{Mg}^{2+}}^{\text{dT}_{30}-\text{dA}_{30}}$ , in  $[\text{Mg}^{2+}] = 20$  mM. (C) Same as (B), except the results correspond to are  $\text{Ca}^{2+}$  ions. (D) Difference in  $P(r_z, r_{X^{2+}-P}^{\text{Min}})$ , for dT<sub>30</sub>, in  $[\text{Mg}^{2+}] = 20$  mM and  $[\text{Ca}^{2+}] = 20$  mM,  $\Delta_{\text{Mg}^{2+}-\text{Ca}^{2+}}^{\text{dT}_{30}}$ . (E) Same as (D), except the plot is for dA<sub>30</sub>.

ions as well (Figure 10C), except that the highest  $\text{Mg}^{2+}$  occupancy is found at  $r_{\text{Mg}^{2+}-\text{P}}^{\text{Min}} \approx 0.35$  nm, whereas for  $\text{Ca}^{2+}$ , it is located at  $r_{\text{Ca}^{2+}-\text{P}}^{\text{Min}} \approx 0.45$  nm, due to its larger size.

We compared the relative population of  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  ions, around  $\text{dT}_{30}$ , by computing the difference in the joint distribution functions,  $\Delta_{\text{Mg}^{2+}-\text{Ca}^{2+}}^{\text{dT}_{30}} = P_{\text{dT}_{30}}(r_z, r_{\text{Mg}^{2+}-\text{P}}^{\text{Min}}) - P_{\text{dT}_{30}}(r_z, r_{\text{Ca}^{2+}-\text{P}}^{\text{Min}})$  (Figure 10D). The population of  $\text{Ca}^{2+}$  ions is reduced in the immediate vicinity of  $\text{dT}_{30}$ , compared to  $\text{Mg}^{2+}$ , indicating weaker interactions and more diffuse binding. Figure 10E illustrates the results for  $\text{dA}_{30}$ , where qualitatively similar differences are found between  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  ions.

To quantify the local ion condensation, we calculated the concentration of ion  $\text{X}^{2+}$  in the vicinity of individual phosphate groups  $i$ ,  $C_{i,\text{X}^{2+}}^*$ , using Eq. S3 in the SI, within a spherical volume of radius set by the Bjerrum length  $l_B$  (Figure S12). For  $\text{dT}_{30}$ ,  $C_{\text{Mg}^{2+}}^* \approx 780$  mM, is larger than  $C_{\text{Ca}^{2+}}^* \approx 540$  mM, which is a consequence of higher charge density and smaller size of  $\text{Mg}^{2+}$ . A similar trend is found in  $\text{dA}_{30}$ , where  $C_{\text{Mg}^{2+}}^* \approx 660$  mM, whereas  $C_{\text{Ca}^{2+}}^* \approx 480$  mM. In both the cases, there are non-specific interactions between the phosphate groups and  $\text{X}^{2+}$  ions, which may be attributed to the lack of tertiary structures in the homopolymeric sequences. This non-specificity is evident when the location of  $\text{X}^{2+}$  is analyzed around each nucleotide (direct interactions only) in individual conformations (Figures S13, S14).

## Discussion:

**$\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  sense major and minor grooves in dsDNA differently:** Neither the dependence of  $\Gamma_{\text{X}^{2+}}$  on  $[\text{X}^{2+}]$  (Figures 2 and 3) nor the scattering profiles (Figures 4 and 5) show significant differences between  $[\text{Mg}^{2+}]$  and  $[\text{Ca}^{2+}]$ . Therefore, the conformations of the short dsDNA sequences are determined predominantly by the stacking and base-pairing interactions, with electrostatic interactions playing a sub-dominant role. Divalent cations merely decrease the effective charge on the phosphate groups. Thus,  $\Gamma_{\text{X}^{2+}}$  and  $I(q)$  are impervious to the nature of the divalent cations. However, differences between  $[\text{Mg}^{2+}]$  and

$[\text{Ca}^{2+}]$  emerge when the details of their binding to dsDNA are examined at a finer scale:  $[\text{Mg}^{2+}]$  binds predominantly to the phosphate groups and minor grooves, whereas  $[\text{Ca}^{2+}]$  preferentially binds to the minor groove, and exhibits little or no direct interactions with the phosphate groups. The distinct binding modes of  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  to dsDNA could affect the differential interaction of proteins with DNA,<sup>60</sup> and possibly certain enzymatic activities where  $\text{Ca}^{2+}$  can not substitute for  $\text{Mg}^{2+}$ .<sup>61,62</sup>

**Stacking versus electrostatic interactions in ssDNA:** The contrast in the extent of heterogeneity between  $\text{dT}_{30}$  and  $\text{dA}_{30}$  is striking. The stability of ssDNA should result due to the interplay between favorable stacking interactions ( $U_{ST}$ ) and unfavorable electrostatic interactions ( $U_{EL}$ ). In  $\text{dT}_{30}$ , these two interactions nearly cancel each other (Figure S7). As a result,  $\text{dT}_{30}$  behaves like a random coil, especially at high  $[\text{X}^{2+}]$  (Figures 7E, F show  $\nu_{eff} \approx 0.6$  in 20 mM  $\text{Mg}^{2+}$  or  $\text{Ca}^{2+}$ ). In contrast, the stability of  $\text{dA}_{30}$  is determined by the favorable stacking interactions, which are strong enough to overcome the electrostatic repulsion (figure S7). As a result,  $\text{dA}_{30}$  forms helical structures at both  $[\text{X}^{2+}] = 0$  mM and 20 mM (Figures 7E, F).

Although we have focused on  $\text{dT}_{30}$  and  $\text{dA}_{30}$ , our result could be applicable to  $\text{dC}_n$  and  $\text{dG}_n$ . We will assume that  $n$  is not too large because as  $n$  increases there is a high probability that the ssDNA chain can fold onto itself, stabilized by the formation of intramolecular stacks. It is thought (see for example<sup>63</sup>) that the order of stacking interactions, from the most to the favorable may be arranged as  $U_{ST}^{GG} > U_{ST}^{AA} > U_{ST}^{TT} > U_{ST}^{CC}$ . It follows from our results that  $\text{dG}_{30}$  is likely to form more stable and stiff helix over the range of ion concentrations examined here, whereas the conformations of  $\text{dC}_{30}$  would be determined almost exclusively by  $U_{EL}$ . In other words,  $\text{dC}_{30}$  is, for practical purposes, should be viewed as a short polyelectrolyte. It should be noted that the reported results and predictions made in this work apply only to DNA sequences. Comparison between DNA and RNA is not feasible because there are significant differences between DNA and RNA strands with even identical base compositions. These variations extend to interactions with ions,<sup>13</sup> and

propensity for condensate formation.<sup>66</sup>

## Conclusions:

Using a transferable computational TIS-ION model that includes monovalent and divalent cations explicitly, we probed the interplay between stacking, base-pairing (dsDNA), and electrostatic interactions in modulating the conformations of dsDNA and ssDNA sequences. The energy function combines the coarse-grained TIS model for nucleic acids with explicit models for the spherical cations ( $\text{Na}^+$ ,  $\text{Mg}^{2+}$ , and  $\text{Ca}^{2+}$ ). Calculations of the number of monovalent and divalent cations in the vicinity of dsDNA and ssDNA sequences, over a wide range of cation concentrations, are in excellent agreement with experiments<sup>6,10</sup> as are the SAXS profiles for dsDNA, dT<sub>30</sub> and dA<sub>30</sub>. Remarkably, the accurate predictions were made *without fitting* a single parameter to experimental data, thus attesting to the robustness and transferability of the model.

Strikingly, we find that neither SAXS nor ion-counting experiments distinguishes between  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$ , in terms of phosphate binding, but the fingerprints of preferential binding patterns are discernible at the molecular scale. While  $\text{Mg}^{2+}$  binds to both the phosphate groups, and DNA minor groove, the larger sized  $\text{Ca}^{2+}$  are exclusively localized near the minor groove. The interplay between stacking and electrostatic interactions is most evident in the nature of conformations explored by the ssDNAs, dT<sub>30</sub> and dA<sub>30</sub>. In the former, these two interactions nearly cancel each other, which makes dT<sub>30</sub> behave as a homopolymer in a good solvent, at sufficiently high divalent ion concentrations. On the other hand, because stacking dominates in dA<sub>30</sub>, the conformations have considerable helical order.

**Data Availability** Simulations were performed using an in-house code built on the OpenMM platform. The code is available at <https://github.com/balaka92/TIS-DNA-Ex-Ion>.

**Funding:** This work was supported by the National Science Foundation through CHE 2320256. Additional support from the Welch Foundation (F-0019) through the Collie-Welch

Chair is gratefully acknowledged.

**Acknowledgements:** We thank the Texas Advanced Computing Center and Pittsburgh Supercomputing Center for providing generous computational resources.

**Supporting Information:** The Supporting Information is available free of charge at <https://pubs.acs.org>.

- Detailed description of the TIS-ION force-field parameters with the tabulated values; Definition of local and bulk ion concentration; Procedure to obtain SAXS profiles from coarse-grained structures; Reconstructed all atom structures of double and single-stranded DNA; Stacking and electrostatic components of the potential energy; Variation in stacking pattern as a function of divalent cation concentration; Spatial dependence of ion concentrations; 2D spatial densities of divalent cations onto an averaged base-pair plane of the dsDNA; 2D spatial distribution of divalent cations around ssDNA; Structural heterogeneity in the conformations of ssDNA sequences; Tabulated values of Flory prefactors;

## References

- (1) Rouzina, I.; Bloomfield, V. A. DNA bending by small, mobile multivalent cations. *Biophys. J.* **1998**, *74*, 3152–3164
- (2) Soutlanas, P.; Dillingham, M. S.; Velankar, S. S.; Wigley, D. B. DNA binding mediates conformational changes and metal ion coordination in the active site of PcrA helicase. *J. Mol. Biol.* **1999**, *290*, 137–148
- (3) Grigoryev, S. A.; Arya, G.; Correll, S.; Woodcock, C. L.; Schlick, T. Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13317–13322
- (4) Wong, G. C.; Pollack, L. Electrostatics of strongly charged biological polymers: Ion-mediated interactions and self-organization in nucleic acids and proteins. *Annu. Rev. Phys. Chem.* **2010**, *61*, 171–189
- (5) Baumann, C. G.; Smith, S. B.; Bloomfield, V. A.; Bustamante, C. Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 6185–6190
- (6) Bai, Y.; Greenfeld, M.; Travers, K. J.; Chu, V. B.; Lipfert, J.; Doniach, S.; Herschlag, D. Quantitative and comprehensive decomposition of the ion atmosphere around nucleic acids. *J. Am. Chem. Soc.* **2007**, *129*, 14981–14988
- (7) Meisburger, S. P.; Sutton, J. L.; Chen, H.; Pabit, S. A.; Kirmizialtin, S.; Elber, R.; Pollack, L. Polyelectrolyte properties of single stranded DNA measured using SAXS and single-molecule FRET: Beyond the wormlike chain model. *Biopolymers* **2013**, *99*, 1032–1045
- (8) Murphy, M. C.; Rasnik, I.; Cheng, W.; Lohman, T. M.; Ha, T. Probing Single-Stranded DNA Conformational Flexibility Using Fluorescence Spectroscopy. *Biophys. J.* **2004**, *86*, 2530–2537

- (9) McIntosh, D. B.; Duggan, G.; Gouil, Q.; Saleh, O. A. Sequence-dependent elasticity and electrostatics of single-stranded DNA: Signatures of base-stacking. *Biophys. J.* **2014**, *106*, 659–666
- (10) Plumridge, A.; Meisburger, S. P.; Andresen, K.; Pollack, L. The impact of base stacking on the conformations and electrostatics of single-stranded DNA. *Nucleic Acids Res.* **2017**, *45*, 3932–3943
- (11) Yakovchuk, P.; Protozanova, E.; Frank-Kamenetskii, M. D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res.* **2006**, *34*, 564–574
- (12) Giambasu, G. M.; Luchko, T.; Herschlag, D.; York, D. M.; Case, D. A. Ion counting from explicit-solvent simulations and 3D-RISM. *Biophys. J.* **2014**, *106*, 883–894
- (13) Plumridge, A.; Meisburger, S. P.; Pollack, L. Visualizing single-stranded nucleic acids in solution. *Nucleic Acids Res.* **2017**, *45*, e66
- (14) Saleh, O. A.; McIntosh, D. B.; Pincus, P.; Ribbeck, N. Nonlinear low-force elasticity of single-stranded DNA molecules. *Phys. Rev. Lett.* **2009**, *102*, 068301
- (15) Toan, N. M.; Thirumalai, D. Theory of biopolymer stretching at high forces. *Macromolecules* **2010**, *43*, 4394–4400
- (16) Shkel, I. A.; Tsodikov, O. V.; Record Jr, M. T. Asymptotic solution of the cylindrical nonlinear Poisson–Boltzmann equation at low salt concentration: Analytic expressions for surface potential and preferential interaction coefficient. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2597–2602
- (17) Kirmizialtin, S.; Silalahi, A. R.; Elber, R.; Fenley, M. O. The ionic atmosphere around A-RNA: Poisson-Boltzmann and molecular dynamics simulations. *Biophys. J.* **2012**, *102*, 829–838

- (18) Lamm, G.; Pack, G. R. Local dielectric constants and Poisson–Boltzmann calculations of DNA counterion distributions. *Int. J. Quantum Chem.* **1997**, *65*, 1087–1093
- (19) Chu, V. B.; Bai, Y.; Lipfert, J.; Herschlag, D.; Doniach, S. Evaluation of ion binding to DNA duplexes using a size-modified Poisson-Boltzmann theory. *Biophys. J.* **2007**, *93*, 3202–3209
- (20) Vologodskii, A.; Cozzarelli, N. Modeling of long-range electrostatic interactions in DNA. *Biopolymers* **1995**, *35*, 289–296
- (21) Denesyuk, N. A.; Thirumalai, D. Coarse-grained model for predicting RNA folding thermodynamics. *J. Phys. Chem. B* **2013**, *117*, 4901–4911
- (22) Chakraborty, D.; Hori, N.; Thirumalai, D. Sequence-Dependent Three Interaction Site Model for Single- and Double-Stranded DNA. *J. Chem. Theory Comput.* **2018**, *14*, 3763–3779
- (23) Nguyen, H. T.; Hori, N.; Thirumalai, D. Theory and simulations for RNA folding in mixtures of monovalent and divalent cations. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 21022–21030
- (24) Sun, L. Z.; Qian, J. L.; Cai, P.; Xu, X. Mutual effects between single-stranded DNA conformation and  $\text{Na}^+$ - $\text{Mg}^{2+}$  ion competition in mixed salt solutions. *Phys. Chem. Chem. Phys.* **2022**, *24*, 20867–20881
- (25) Savelyev, A.; MacKerell Jr, A. D. Competition among  $\text{Li}^+$ ,  $\text{Na}^+$ ,  $\text{K}^+$ , and  $\text{Rb}^+$  monovalent ions for DNA in molecular dynamics simulations using the additive CHARMM36 and drude polarizable force fields. *J. Phys. Chem. B* **2015**, *119*, 4428–4440
- (26) Yates, L. A.; Aramayo, R. J.; Pokhrel, N.; Caldwell, C. C.; Kaplan, J. A.; Perera, R. L.; Spies, M.; Antony, E.; Zhang, X. A structural and dynamic model for the assembly of Replication Protein A on single-stranded DNA. *Nat. Commun.* **2018**, *9*

- (27) Cruz-León, S.; Grotz, K. K.; Schwierz, N. Extended magnesium and calcium force field parameters for accurate ion–nucleic acid interactions in biomolecular simulations. *J. Chem. Phys.* **2021**, *154*, 171102
- (28) Yoo, J.; Aksimentiev, A. Competitive binding of cations to duplex DNA revealed through molecular dynamics simulations. *J. Phys. Chem. B* **2012**, *116*, 12946–12954
- (29) Hinckley, D. M.; J. de Pablo, Coarse-grained ions for nucleic acid modeling. *J. Chem. Theory Comput.* **2015**, *11*, 5436–5446
- (30) Hyeon, C.; Thirumalai, D. Mechanical unfolding of RNA hairpins. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6789–6794
- (31) Hyeon, C.; Thirumalai, D. Capturing the essence of folding and functions of biomolecules using coarse-grained models. *Nat. Commun.* **2011**, *2*, 487
- (32) Ouldridge, T. E.; Louis, A. A.; Doye, J. P. Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *J. Chem. Phys.* **2011**, *134*, 085101
- (33) Hinckley, D. M.; Freeman, G. S.; Whitmer, J. K.; J. de Pablo, An experimentally-informed coarse-grained 3-site-per-nucleotide model of DNA: Structure, thermodynamics, and dynamics of hybridization. *J. Chem. Phys.* **2013**, *139*, 144903
- (34) Sun, L. Z.; Qian, J. L.; Cai, P.; Hu, H. X.; Xu, X.; Luo, M. B. Mg<sup>2+</sup> effects on the single-stranded DNA conformations and nanopore translocation dynamics. *Polymer* **2022**, *250*, 124895
- (35) Denesyuk, N. A.; Thirumalai, D. How do metal ions direct ribozyme folding? *NatChem* **2015**, *7*, 793–801
- (36) Hori, N.; Denesyuk, N. A.; Thirumalai, D. Shape changes and cooperativity in the folding of the central domain of the 16S ribosomal RNA. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2020837118

- (37) Hori, N.; Thirumalai, D. Watching ion-driven kinetics of ribozyme folding and misfolding caused by energetic and topological frustration one molecule at a time. *Nucleic Acids Res.* **2023**, *51*, gkad755
- (38) SantaLucia, J.; Allawi, H. T.; Seneviratne, P. A. Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry* **1996**, *35*, 3555–3562
- (39) SantaLucia, J.; Hicks, D. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 415–440
- (40) Hasted, J. Water: A Comprehensive Treatise. *The Physics and Physical Chemistry of Water* **1972**, *1*, 255–305
- (41) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L. P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLOS Biol.* **2017**, *13*, e1005659
- (42) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089–10092
- (43) Honeycutt, J. D.; Thirumalai, D. The nature of folded states of globular proteins. *Biopolymers* **1992**, *32*, 695–709
- (44) Hockney, R. W.; Eastwood, J. W. *Computer simulation using particles*; CRC Press, 2021
- (45) Pabit, S. A.; Qiu, X.; Lamb, J. S.; Li, L.; Meisburger, S. P.; Pollack, L. Both helix topology and counterion distribution contribute to the more effective charge screening in dsRNA compared with dsDNA. *Nucleic Acids Res.* **2009**, *37*, 3887–3896

- (46) Smith, P. Equilibrium dialysis data and the relationships between preferential interaction parameters for biological systems in terms of Kirkwood-Buff integrals. *J. Phys. Chem. B* **2006**, *110*, 2862–2868
- (47) Schurr, J.; Rangel, D.; Aragon, S. A contribution to the theory of preferential interaction coefficients. *Biophys. J.* **2005**, *89*, 2258–2276
- (48) Klimov, D. K.; Betancourt, M. R.; Thirumalai, D. Virtual atom representation of hydrogen bonds in minimal off-lattice models of alpha helices: Effect on stability, cooperativity and kinetics. *Folding and Design* **1998**, *3*, 481–496
- (49) Bloomfield, V. A. Condensation of DNA by multivalent cations: Considerations on mechanism. *Biopolymers* **1991**, *31*, 1471–1481
- (50) Arscott, P.; MA, C.; Wenner, J.; Bloomfield, V. DNA Condensation by Cobalt Hexaammine(III) in alcohol-water mixtures - dielectric-constant and other solvent effects. *Biopolymers* **1995**, *36*, 345–364
- (51) Ma, C.; Bloomfield, V. A. Gel electrophoresis measurement of counterion condensation on DNA. *Biopolymers* **1995**, *35*, 211–216
- (52) Wilson, R. W.; Bloomfield, V. A. Counterion-induced condensation of deoxyribonucleic acid. A light-scattering study. *Biochemistry* **1979**, *18*, 2192–2196
- (53) Bloomfield, V. A. DNA condensation by multivalent cations. *Biopolymers* **1997**, *44*, 269–282
- (54) Rohatgi, A. Webplotdigitizer: Version 4.6. 2022; <https://automeris.io/WebPlotDigitizer> (accessed September 21, 2021)
- (55) Hori, N. TIS2AA. 2017; <https://github.com/naotohori/TIS2AA/tree/v1.0> (accessed March 6, 2022)

- (56) Svergun, D.; Barberato, C.; Koch, M. H. CRY SOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **1995**, *28*, 768–773
- (57) Higgins, J. S.; Benoit, H. C. *Polymers and neutron scattering*; Oxford University Press: London, 1994
- (58) Sim, A. Y.; Lipfert, J.; Herschlag, D.; Doniach, S. Salt dependence of the radius of gyration and flexibility of single stranded DNA in solution probed by Small-Angle X-ray Scattering. *Phys. Rev. E* **2012**, *86*, 021901
- (59) Ke, C.; Humeniuk, M.; Hanna, S.; Marszalek, P. E. Direct measurements of base stacking interactions in DNA by single-molecule atomic-force spectroscopy. *Phys. Rev. Lett.* **2007**, *99*, 018302
- (60) Osawa, M.; Dace, A.; Tong, K. I.; Valiveti, A.; Ikura, M.; Ames, J. B. Mg<sup>2+</sup> and Ca<sup>2+</sup> differentially regulate DNA binding and dimerization of DREAM. *J. Biol. Chem.* **2005**, *280*, 18008–18014
- (61) Bellamy, S. R.; Kovacheva, Y. S.; Zulkipli, I. H.; Halford, S. E. Differences between Ca<sup>2+</sup> and Mg<sup>2+</sup> in DNA binding and release by the SfiI restriction endonuclease: implications for DNA looping. *Nucleic Acids Res.* **2009**, *37*, 5443–5453
- (62) Pingoud, A.; Fuxreiter, M.; Pingoud, V.; Wende, W. Type II restriction endonucleases: structure and mechanism. *Cell. Mol. Life Sci.* **2005**, *62*, 685–707
- (63) Brown, R. F.; Andrews, C. T.; Elcock, A. H. Stacking free energies of all DNA and RNA nucleoside pairs and dinucleoside-monophosphates computed using recently revised AMBER parameters and compared with experiment. *J. Chem. Theory Comput.* **2015**, *11*, 2315–2328

- (64) Macke, T. J.; Case, D. A. Modeling unusual nucleic acid structures. *ACS Publications* **1998**, *682*, 379–393
- (65) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38
- (66) Tom, J. K.; Onuchic, P. L; Deniz, A. A. Short polyA RNA homopolymers undergo  $\text{Mg}^{2+}$ -mediated kinetically arrested condensation. *J. Phys. Chem. B*, **2022**, *126*, 9715–9725

# Graphical TOC Entry

