Ion Counting from Explicit-Solvent Simulations and 3D-RISM

George M. Giambaşu,[†] Tyler Luchko,[†] Daniel Herschlag,[‡] Darrin M. York,[†*] and David A. Case[†*]
[†]Department of Chemistry and Chemical Biology and BioMaPS Institute, Rutgers University, Piscataway, New Jersey; and [‡]Department of Biochemistry, Stanford University, Stanford, California

ABSTRACT The ionic atmosphere around nucleic acids remains only partially understood at atomic-level detail. Ion counting (IC) experiments provide a quantitative measure of the ionic atmosphere around nucleic acids and, as such, are a natural route for testing quantitative theoretical approaches. In this article, we replicate IC experiments involving duplex DNA in $NaCl_{(aq)}$ using molecular dynamics (MD) simulation, the three-dimensional reference interaction site model (3D-RISM), and nonlinear Poisson-Boltzmann (NLPB) calculations and test against recent buffer-equilibration atomic emission spectroscopy measurements. Further, we outline the statistical mechanical basis for interpreting IC experiments and clarify the use of specific concentration scales. Near physiological concentrations, MD simulation and 3D-RISM estimates are close to experimental results, but at higher concentrations (>0.7 M), both methods underestimate the number of condensed cations and overestimate the number of excluded anions. The effect of DNA charge on ion and water atmosphere extends 20–25 Å from its surface, yielding layered density profiles. Overall, ion distributions from 3D-RISMs are relatively close to those from corresponding MD simulations, but with less Na^+ binding in grooves and tighter binding to phosphates. NLPB calculations, on the other hand, systematically underestimate the number of condensed cations at almost all concentrations and yield nearly structureless ion distributions that are qualitatively distinct from those generated by both MD simulation and 3D-RISM. These results suggest that MD simulation and 3D-RISM may be further developed to provide quantitative insight into the characterization of the ion atmosphere around nucleic acids and their effect on structure and stability.

INTRODUCTION

The ionic atmosphere around nucleic acids has a major impact on the stability of secondary and tertiary structure, modulating the binding of charged drugs and proteins and condensation and packing in cells (1). Recently, several experimental methods have been developed to examine the nature of ion atmosphere around nucleic acids through ion-counting experiments that rely on anomalous small-angle x-ray scattering (2,3), buffer equilibration atomic emission spectroscopy (BE-AES) (4,5), and titration with fluorescent dyes (6). In previous decades, ^{23}Na or ^{39}Co NMR relaxation rates have also been employed (7,8). There is a long history of theoretical models that attempt to describe ion atmosphere around nucleic acids, but there are relatively few calculations using models with atomic detail, in large part because equilibration of the water and ion distributions around nucleic acids is very slow in molecular dynamics (MD) or Monte Carlo simulations (9,10) and has only recently become feasible.

It is generally accepted that the majority of ions surrounding a nucleic acid molecule stay within a few angstroms of its surface and are able to diffuse along it or to the bulk, whereas relatively few ions occupy specific binding sites (site-bound ions), although these are no less important. The former type of binding was designated by Manning as territorial, as opposed to atmospheric, binding, a characteristic of nonpolymeric solutions (11,12). Experiments have shown that counterions associate with nucleic acids at salt concentrations much smaller than physiological concentrations, a fact attributed to the polymeric nature of nucleic acids (13). Manning's counterion condensation theory (11,12,14–16) posits that beyond a certain threshold in axial charge density, an excess of 0.7–0.8 monovalent cations are bound per phosphate unit. Further, it has been shown that this phenomenon is the result of the attempt of the ionic and solvent atmosphere to neutralize the nucleic acid through counterion condensation and coion depletion (13). Although counterion condensation is prominent at low bulk salt concentration, as the bulk salt concentration increases the contribution from anion depletion rises (17).

Theoretical approaches for studying nucleic-acid-ion interactions beyond Manning's theory include MD simulations (9,10,18–32), nonlinear Poisson-Boltzmann (NLPB) (4,17,29,33–35) calculations, and integral equation theory (IET) (36–39). MD in conjunction with molecular mechanics force fields has the advantage of providing complete atomistic detail of ion and water distributions, including dynamical information, but at larger computational expense than similar calculations using NLPB or IET. NLPB calculations avoid these computational costs but considerably oversimplify the physics of hydration and neutralization and omit ion correlation. IET, and specifically the three-dimensional reference interaction site model (3D-RISM)

Submitted October 8, 2013, and accepted for publication January 10, 2014.

*Correspondence: york@biomaps.rutgers.edu or case@biomaps.rutgers.edu

George M. Giambaşu and Tyler Luchko contributed equally to this work.

Tyler Luchko's present address is Department of Physics and Astronomy, California State University Northridge, Northridge, CA 91330.

Editor: Kathleen Hall.

0006-3495/14/02/0883/12 $2.00

http://dx.doi.org/10.1016/j.bpj.2014.01.021

(40–43), provides a third approach that makes use of explicit-solvent models and provides molecular, noncontinuum detail at intermediate computational cost. The 3D-RISM can be considered parameter-free, requiring only an approximate closure relation and the same molecular mechanics model used for MD simulations (44).

In this article, we examine both the macroscopic and microscopic pictures of how nucleic acids interact with their ionic atmosphere as generated by MD simulations, 3D-RISM, and conventional NLPB calculations. To assure a proper comparison between theory and experiment, we replicate ion-counting experiments carried out using buffer equilibration atomic emission spectroscopy (4). The primary observable that allows the connection between ion-counting experiments and theory is the number of excess ions per solute, or preferential interaction parameters (PIPs). Several approaches for computing PIPs have been used in the past and we also aim to clarify the proper use of concentration scales to be used for comparing theoretical results with experimental data. At the microscopic level, we show that structural details of density distributions of ions and water around DNA clearly distinguish between continuum and noncontinuum theories. At the macroscopic level, we reveal that MD and 3D-RISM estimates are close to experimental results near physiological concentrations, but that at concentrations >0.7 M, both methods underestimate the number of condensed cations and overestimate the number of excluded anions. NLPB calculations, on the other hand, systematically underestimate the number of condensed cations at almost all concentrations, and yield nearly structureless ion distributions that are qualitatively distinct from those generated by both MD simulations and 3D-RISM.

ION-COUNTING EXPERIMENTS

Ion-counting experiments measure the number of excess ions (Γ) or PIPs around nucleic acid molecules. These quantities can be measured experimentally in a variety of ways, most simply by dialysis experiments that count the number of ions crossing a semipermeable membrane when a nucleic acid is added (4,45). Specifically, for DNA molecules immersed in an aqueous solution containing a single type of anion (−) and cation (+),

$$\Gamma_{+/-}^{(M)} = \frac{M_{+/-}^{(1)} - M_{+/-}^{(2)}}{M_{DNA}^{(1)}}, \tag{1}$$

where M stands for the molarity of the corresponding species in the two chambers of the osmotic cell (46).

PIPs measure the preference of a solute (labeled 2) for water (labeled 1) or a cosolvent (labeled 3) (47,48). PIPs are fundamental in studying salt effects (7,49,50) and measure the integrated amount by which the local concentration of water or cosolvent around the solute differ from the corresponding values in the bulk. When computed for ionic species, PIPs are also called Donnan coefficients. In the case of this study, the cosolvent is represented by dissociated sodium chloride, although it can be any charged or neutral species in general. The definition of Γ_{23} depends on the units of concentration used (47,48): for molality (m) or molarity (M) scales, Γ_{23} is defined as

$$\Gamma_{23}^{(m)} \equiv \lim_{m_2 \to 0} \left(\frac{\partial m_3}{\partial m_2} \right)_{T,\mu_1,\mu_3} \tag{2}$$

and

$$\Gamma_{23}^{(M)} \equiv \lim_{M_2 \to 0} \left(\frac{\partial M_3}{\partial M_2} \right)_{T,\mu_1,\mu_3}, \tag{3}$$

These definitions are valid in the limit of infinite dilution of the solute and are derived in the grand canonical ensemble resembling the experimental conditions of dialysis. This corresponds to equal chemical potentials of salt and water in the two chambers of the osmosis experiment. As discussed below, the molal and molar scales lead to different values of PIPs. From the specific case of the system considered here, index 2 corresponds to the DNA, index 1 to water, and index 3 to sodium chloride, NaCl. When the cosolvent is a monovalent salt and PIPs are determined separately for monocationic (+) and monoanionic (−) species, one can show that

$$\Gamma_{2+} - \Gamma_{2-} = |Z_2| \tag{4}$$

and

$$\Gamma_{2+} + \Gamma_{2-} = |Z_2| + 2\Gamma_{23}, \tag{5}$$

where Z_2 is the charge of the solute (7,49,50). Combining Eqs. 4 and 5 leads to $\Gamma_{23} = \Gamma_{2-}$, i.e., the preferential order parameter for the anion is equal to that of the salt. Thus, $|\Gamma_{2-}|$ can be interpreted as the number of pairs of monovalent cations and anions depleted when titrating DNA in the salt solution.

Ion counting using Kirkwood-Buff integrals

Kirkwood-Buff (KB) integrals, $G_{\alpha\gamma}$, are central quantities in liquid state theory (51) and are defined as

$$G_{\alpha\gamma} = \int_V \left(g_{\alpha\gamma}(\mathbf{r}) - 1 \right) d\mathbf{r}, \tag{6}$$

where $g_{\alpha\gamma}(\mathbf{r})$ is the pair distribution function (PDF) at constant chemical potential, μ, volume, V, and temperature, T, with α and γ representing any two components of the

solution. Following Smith (48), the PIPs in the grand canonical ensemble can be calculated for molal and molar concentration units using

$$\Gamma_{2+}^{(m)} = \rho_+(G_{2+} - G_{21}), \tag{7}$$

and

$$\Gamma_{2+}^{(M)} = \rho_+ G_{2+}. \tag{8}$$

It has been argued that $G_{\alpha\gamma}$ can be estimated from MD simulation in isothermal-isobaric (NPT) or canonical (NVT) ensembles by integrating not to infinity but to a correlation radius, R_C (52–55):

$$G_{\alpha\gamma} \approx \int_0^{R_C} \left(g_{\alpha\gamma}^{NVT}(\mathbf{r}) - 1\right) d\mathbf{r} \quad st \quad g_{\alpha\gamma}^{NVT}(\mathbf{r}) = 1 \ \forall r > R_C. \tag{9}$$

The R_C, also called the radius of influence, delimits the space around the solute for which the solvent and cosolvent concentrations are different from those found in the bulk (53,55). The location and shape of the partition are arbitrary but should be large enough that the bulk region is not perturbed by the solute. The partition should be of fixed volume and shape. For well converged simulations, there is no need to estimate R, as one can carry out the integration up to the system boundary. For the case of the rigid B-DNA molecule considered here (48 nucleobases and a 46− charge), our calculations indicate that the correlation radius slightly decreases with increasing salt concentration and can be estimated to be 20–25 Å from the DNA helical axis. This means that to simulate the DNA in equilibrium with a bulk salt, one has to extend the simulation periodic box to accommodate the volume enclosed in the correlation radius plus an additional volume representing the bulk. The size of this additional volume has to be large enough to allow RDFs to properly converge (18) (see also Pollack (3)).

For 3D-RISM, to account for the long-range behavior of the ion distributions, a correction is applied using the long-range asymptotic solution to the total correlation function. The correction was originally proposed to overcome difficulties in performing the inverse fast Fourier transform when solving the RISM equation and captures the Debye-screened long-range distribution of the electrolyte (42,43).

Ion counting using the two-partition (domain) model

For methods that directly generate PDFs, such as 3D-RISM and NLPB theory, calculating PIPs with KB integrals is a natural approach. However, for MD simulations, calculating and storing 3D PDFs from a trajectory can be a computationally complex and memory-intensive procedure. One way to lessen the effort required is to use the two-partition approach, introduced first by Anderson and Record in a general context (46, 56).

Within the two-domain approach, the system of interest is partitioned into two regions (see Fig. 1): a local region (R1) surrounding the solute, for which the densities of solvent and cosolvent (ions) are different from those found in the bulk, and a bulk region (R2) for which the densities of solvent and salt are uniform. This resembles an osmotic cell that allows the free exchange of solvent and cosolvent between R1 and R2 but in which the solute is restricted to R1. It must be pointed out that during the MD simulations, the instantaneous total charge in the two partitions fluctuates, representing better the situation one would find in a grand canonical ensemble or an osmosis experiment. Since solvent and cosolvent are free to diffuse across the virtual membrane, we obtain the desired $T\mu_1\mu_3$ ensemble, similar to the KB approach (52–54).

One can further adapt Eq. 1 for the case of an MD simulation. Taking into account that MD simulations are run with a constant number of particles and constant volume $(N_x = N_x^{(1)} + N_x^{(2)}, V = V^{(1)} + V^{(2)})$ and contain a single DNA molecule, Eq. 1 becomes

$$\Gamma_{2,+}^{(m)} = N_+ - \frac{\rho_+}{\rho_{H_2O}} N_{H_2O} \tag{10}$$

$$\Gamma_{2,+}^{(M)} = N_+ - \rho_+ V. \tag{11}$$

Integrating over the volume of the simulation cell yields Eqs. 10 and 11 from Eqs. 7 and 8, respectively. It can be

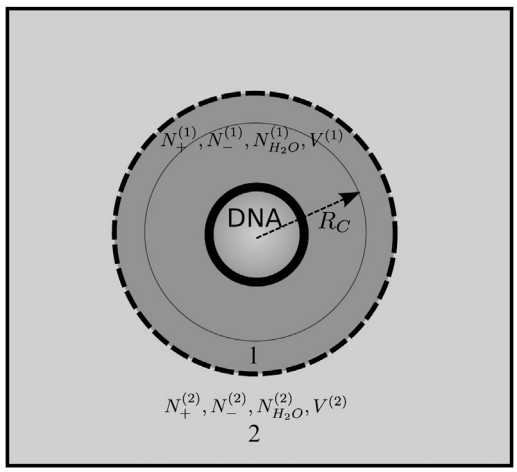

FIGURE 1 Ion counting using the two-domain approach (46), adapted from Anderson and Record (56). The simulated system is separated into two regions, dark gray (*region 1*) and light gray (*region 2*), with the boundary between the two regions shown as a dashed line. Region 1 has to be chosen so that it contains at least all the solvent and cosolvent within the correlation radius (see text) from the DNA, and region 2 so that it resembles the bulk with uniform concentration of salt. Using this approach, a formula for the PIPs can be derived that depends on the salt bulk (*region 2*) concentration. This method may be useful when the $g(r)$ cannot be easily defined.

seen from Eqs. 10 and 11 that it is sufficient to estimate the average number density of water and NaCl in the bulklike region 2 after equilibration. The partition between the inner and bulk regions is chosen so that the inner region includes all the particles located within the correlation radius.

Both the KB integral method and the two-partition method depend on a good representation of the bulk region to estimate the ion and water concentrations. With that in mind, in this work, the DNA-water-ion system used in the MD simulations was extended at least 60 Å in the direction perpendicular to the DNA main axis. Pappu and Chen (18) have shown that there is a strong interdependence between the size of the simulation system and convergence of the distribution functions of charged species around nucleic acids.

METHODS

MD simulations

Molecular mechanics force fields

The 24L B-DNA structure was prepared with Nucleic Acid Builder (NAB) (57) using the Arnott B-DNA fiber diffraction model (58). Three popular water models were used: TIP3P (59), TIP4P-Ew (60,61), and SPC/E (62). Only the SPC/E water model was used in 3D-RISM, with modified hydrogen Lennard-Jones parameters (44): $\sigma_H/2 = (\sigma_O/2) - b_{OH}$ and $\epsilon_H = 0.1\epsilon_O$, where σ and E are the standard Lennard-Jones parameters and b_{OH} is the intramolecular oxygen-hydrogen bond distance. Hydrogen Lennard-Jones parameters are necessary to converge dielectrically consistent RISM (DRISM) equations, and these parameters have been shown to reproduce the solvent polarization free energy of explicit simulations for both polypeptides (44) and ions (63) in water. Ion parameters that were optimized to reproduce solvation energies and shown to prevent the salting-out effect were taken from Joung and Cheatham (64) and were not modified for RISM calculations. The all-atom ff10 force field for nucleic acids was used (65,66). We denote the three force fields as JC:TIP3P, JC:TIP4P-Ew, JC:SPC/E. (Since the DNA is harmonically restrained and the ff10 force field improves only some torsional terms of the ff99 force field (67), the results reported here should be the same when using either force field.). All topologies and structures were generated with the AMBER 10 package (68–70). The COULOMB constant used in NAMD 2.7 71 (332.0636) was changed to correspond to that used in the AMBER suite of programs (332.0522173) (68–70).

MD simulation protocol

Simulations were performed with the NAMD simulation package (version 2.7) (71). The DNA immersed in a preequilibrated water box with hexagonal prism symmetry ($a = b \approx 127.0$ Å, $c \approx 152.0$ Å, $\alpha = 90°$, $\beta = 90°$, $\gamma = 60°$). The ionic atmosphere consisted of Na^+ and Cl^- ions that were added at random positions at least 5.0 Å away from any DNA atom to neutralize the system to reach a desired starting concentration of ~0.2 and 0.7 M. The isothermal-isobaric (NPT) or isothermal (NVT) ensemble were used at 1 atm and 300 K using the Nosé-Hoover-Langevin piston (72,73) with a decay period of 100.0 fs and a damping timescale of 50 fs and the Langevin thermostat with a collision frequency of 0.1 ps^{-1}. The smooth particle mesh Ewald method (74,75) was employed with a B-spline interpolation order of 6 and the default κ value used in NAMD. The fast Fourier transform grid points used for the lattice directions were chosen using ~ 1.0 Å spacing. Nonbonded interactions were treated using an atom-based cutoff of 9.0 Å, with no switching for nonbond potentials. Numerical integration was performed using the leap-frog Verlet algorithm with a 2 fs time step (76). Covalent bond lengths involving water hydrogens were constrained using the SHAKE algorithm (77). All heavy atoms of the DNA were restrained to their initial positions with a harmonic potential using a 2.0 kcal/mol $Å^2$ force constant. The systems were equilibrated for 35 ns in the NPT ensemble. The motions and relaxation of solvent and counter- and coions are notoriously slow to converge in nucleic acid simulations (10), and careful equilibration is critical, especially for low salt concentrations. Further details of the various simulations are given in Table S2 in the Supporting Material.

RISM

The 3D density distribution of aqueous salt solutions around a large macromolecule such as DNA can be directly computed for each atomic site of the solvent using the 3D-RISM (40–43). Like NLPB calculations, numerical solutions are iteratively solved on a 3D grid in the infinite-dilution regime, but unlike NLPB, 3D-RISM uses common molecular force fields to model both the solute and solvent and explicitly calculates an equilibrium density distribution of solvent around the solute.

The 3D-RISM is based on the Ornstein and Zernike IET (the OZ equation) (78,79), which expresses the density distribution in terms of direct and indirect correlation functionals. The OZ equation is inherently six-dimensional for polyatomic molecules due to orientational dependence on the intermolecular interactions. 3D-RISM formalism reduces this to three dimensions by orientational averaging of the solvent degrees of freedom such that the resulting solvent density distributions contain only a spatial dependence, $\rho_\gamma(\mathbf{r})$. In this manner, the distributions of atomic sites on water and salt are represented on 3D grids via the total correlation function, $h_\gamma(\mathbf{r}) = (\rho_\gamma(\mathbf{r})/\rho_\gamma^{bulk}) - 1$, and the direct correlation function, $c_\gamma(\mathbf{r})$, which are related as

$$h_\gamma(\mathbf{r}) = \sum_\alpha \int c_\alpha(\mathbf{r} - \mathbf{r}')\chi_{\alpha\gamma}(r')d\mathbf{r}'. \qquad (12)$$

$\chi_{\alpha\gamma}(r)$ is the site-site solvent susceptibility of solvent sites α and γ and contains the orientationally averaged bulk properties of the solvent. We use the DRISM integral equation (80,81) to precompute $\widehat{\chi}_{\alpha\gamma}$. In practice, the convolution integrals in Eq. 12 are performed in reciprocal space using a 3D fast Fourier transform,

$$\widehat{h}_\gamma(\mathbf{k}) = \sum_\alpha \widehat{c}_\alpha(\mathbf{k})\widehat{\chi}_{\alpha\gamma}(k), \qquad (13)$$

an approach first developed for the OZ equation (82–84) and later applied to 1D-RISM (85) and 3D-RISM (43,86,87) that considers the long-range nature of electrostatic interactions between solvent and solute.

As in all OZ-based theories, a second, so-called closure equation must be used to obtain a unique solution. Due to the complexity of the closure equation, an approximation must be used. The form of this approximation has a major impact on the convergence of calculations, as well as on resulting thermodynamic quantities and densities. The general 3D-RISM closure relation has the form

$$\begin{aligned} g_{\alpha\gamma}(\mathbf{r}) &= h_{\alpha\gamma}(\mathbf{r}) + 1 \\ &= \exp\{-\beta u_{\alpha\gamma}(\mathbf{r}) + h_{\alpha\gamma}(\mathbf{r}) - c_{\alpha\gamma}(\mathbf{r}) + b_{\alpha\gamma}(\mathbf{r})\}, \end{aligned}$$

where $g_{\alpha\gamma}(\mathbf{r})$ is the PDF and $b_{\alpha\gamma}(\mathbf{r})$ is the bridge function, which is only known as an infinite series of functionals and is always subject to some approximation (51). Many closure relations have been developed for IETs, of which the most popular are the hypernetted chain (HNC) (88) and partially linearized Kovalenko-Hirata (KH) (41) equations. In HNC 88, the bridge function is simply set to zero. HNC produces good results for ionic (38,81,89,90) and polar systems (91–93) and has an exact, closed-form expression for the excess chemical potential when coupled

with RISM theory (94). However, HNC solutions are often difficult to converge. The KH closure (41) is numerically robust and addresses this problem by linearizing regions of density greater than bulk, $g_{\alpha\gamma}(\mathbf{r})>1$. The partial series expansion of order-n (PSE-n) (95) generalizes the linearization to a Taylor series,

$$g_\alpha^{\mathrm{PSE}-n}(\mathbf{r}) = \begin{cases} \exp\left(t_\alpha^*(\mathbf{r})\right) & t_\alpha^*(\mathbf{r})<0 \\ 1+\sum_{i=1}^{n}\dfrac{t_\alpha^*(\mathbf{r})^i}{i!} & t_\alpha^*(\mathbf{r})\geq 0 \end{cases} \qquad (14)$$

$$t_\alpha^*(\mathbf{r}) = -\beta u_\alpha(\mathbf{r}) + h_\alpha(\mathbf{r}) - c_\alpha(\mathbf{r}),$$

where β is the reciprocal thermodynamic temperature. For $n = 1$, the KH closure is recovered and HNC is the limiting case as $n\to\infty$. Like HNC, KH and PSE-n closures have an exact, closed-form expression for the chemical potential.

The number of terms used in the PSE-n closure does have an impact on the physical properties of the solvent, particularly for ions (63). For all calculations presented here, we used the highest order closure that could be converged with the cSPC/E model, PSE-4. For more discussion on the choice of closure and the impact on the calculated PIPs, see the Supporting Material.

DRISM

DRISM calculations were performed using the rism1d program in the AmberTools 12 molecular modeling package (44,96) and were performed largely according to the procedure of Joung et al. (63). To obtain $\widehat{\chi}_{\alpha\gamma}(k)$ of the bulk solvent, the DRISM equation coupled with the PSE-4 closure, was iteratively solved using the modified direct inversion of the iterative subspace (97) to a residual tolerance of 10^{-12} at a temperature of 298.15 K and a dielectric constant of 78.44 for bulk water. A grid spacing of 0.025 Å was used throughout. To aid the convergence, solutions from lower-order closures, starting from KH, were iteratively used as initial guesses until the PSE-4 closure was converged. The water density for each salt concentration was interpolated from experimental measurements using cubic splines (63,98,99).

3D-RISM

3D-RISM calculations were performed using the rism3d.snglpnt program in the AmberTools 12 molecular modeling package (44,96). Equation 13, coupled with PSE-4 closures, was iteratively solved using the modified direct inversion of the iterative subspace to a residual tolerance of 10^{-6} for all ion concentrations. Similar to the DRISM calculations, rapid convergence was achieved by first performing a few iterations with KH through PSE-3 closures to provide an initial guess for PSE-4. The extent of the solvation-box grid used was selected for each concentration to obtain a precision 0.002 excess ions together with a grid spacing of 0.5 Å. (for further information, see Table S2).

NLPB model

The NLPB method has been described in detail elsewhere (100,101). Briefly, NLPB has the form

$$\nabla\cdot[\epsilon(\mathbf{x})\nabla\phi(\mathbf{x})] + 4\pi\sum_{j=1}^{m}\rho_j q_j \times \exp\left[-q_j\phi(\mathbf{x})/kT - V_j(\mathbf{x})/kT\right] = -4\pi\rho^{\mathrm{U}}(\mathbf{x}), \qquad (15)$$

where $\epsilon(\mathbf{x})$ is an inhomogeneous dielectric medium of relative permittivity, $\rho^{\mathrm{U}}(\mathbf{x})$ is the charge distribution of the solute, ρ_j is the bulk number density of the jth species of mobile ion in solution with charge q_j, $V_j(\mathbf{x})$ is the steric interaction between the mobile ion and the solute, and $\phi(\mathbf{x})$ is the resulting electrostatic potential. $\phi(\mathbf{x})$ can be numerically solved for on a 3D grid, which gives, for the PDF of ion species j, the expression

$$g_j(\mathbf{x}) = \exp\left[-q_j\phi(\mathbf{x})/kT - V_j(\mathbf{x})/kT\right]. \qquad (16)$$

NLPB calculations were performed with Adaptive Poisson Boltzmann Solver (APBS) 1.4.1 (102). The grid size and spacing were adjusted to ensure that there were at least three digits of precision in all calculated values (<1% error). Where possible, a grid spacing of 1 Å was used, even when a 2 Å spacing gave the same result within the tolerable error. However, APBS is limited to 512^3 total grid points, and we were required to use a 2 Å grid spacing for salt concentrations <0.1 M. Testing showed that even at this larger spacing, the desired precision was achieved. Grid sizes used for all reported values are summarized in Table S3. The internal dielectric of the DNA was set to 1, and 78.54 was used for the solvent, where the standard molecular surface definition was used (srfm mol) with a solvent radius of 1.4 Å. Single Debye-Hückel boundary conditions (bcfl sdh) were used. Temperature was set to 298.15 K. Ion concentrations were identical to those used for DRISM and 3D-RISM.

For all calculations, an ion radius of 1.85 Å was used to define $V_j(\mathbf{x})$ in Eq. 15. This radius was selected because a 1.85 Å ion with 1 e charge at infinite dilution in APBS, using a 66 Å × 66 Å ×66 Å grid with 0.3333 Å grid spacing, has a solvation free energy of −90.2 kcal/mol, which is reasonably close to the values from experiment and the Joung-Cheatham parameters in explicit solvent. As we do not expect Cl^- to make direct contact with the surface of the DNA, we are justified in using $V_{Na^+}(\mathbf{x}) = V_{Cl^-}(\mathbf{x})$, which is standard behavior in APBS. All other parameters are the same as the DNA calculations, and grid size and spacing were varied to ensure at least three digits of precision in the calculated value (see Table S3 for further details).

RESULTS AND DISCUSSION

Molecular distribution functions

Distribution functions are key to creating an average picture of how ions and water arrange around the DNA, as well as for extracting thermodynamic quantities such as PIPs that can be used to compare with ion-counting experiments. Here, we use three types of distribution function available from explicit-solvent MD simulations, 3D-RISM, and NLPB calculations, respectively, which we refer to as 3D, 2D, and 1D distributions. 3D and 1D distributions are commonly used to represent solvent and ion distributions around nucleic acids, and we describe them succinctly in the next paragraphs. Our approach to build 2D distributions is described in the next section.

3D distributions contain most of the information on how solvent and salt distribute around a rigid solute and are usually displayed using isosurfaces (see Fig. 2). 3D distributions are stored on 3D rectangular grids and are the main output from 3D-RISM and NLPB calculations.

1D distribution functions (see Fig. 5) are based on the fact that the approximate cylindrical shape of the B-DNA molecule works well for representing the solvent and ion distributions in cylindrical coordinates (r, θ, z) with density averaged along z and θ, where the z axis coincides with the helical axis. We call these distribution functions

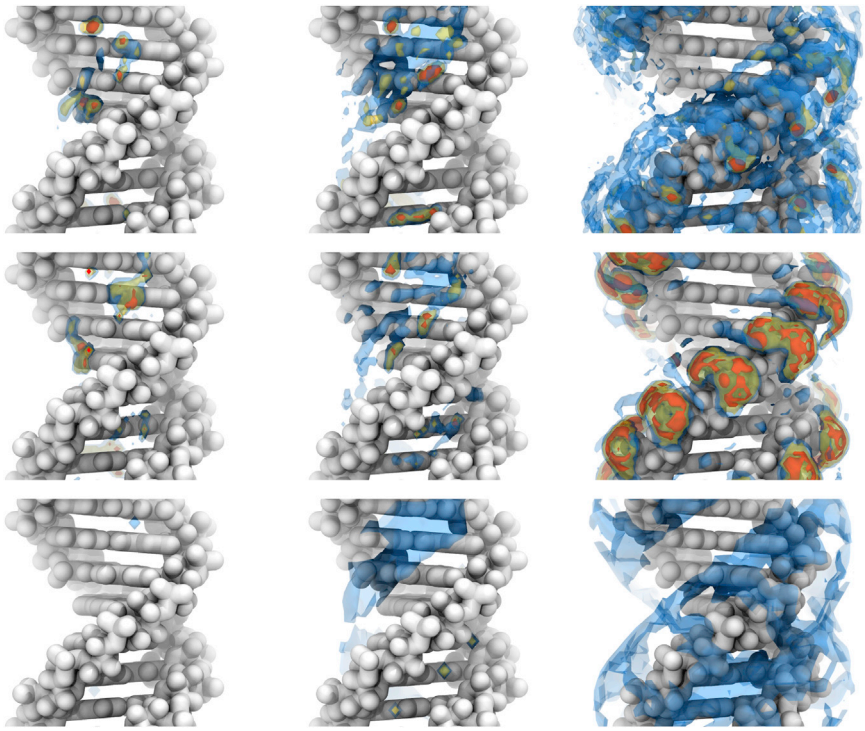

FIGURE 2 Cylindrical shells of sodium ion density around 24L DNA at 0.7452 M NaCl calculated by MD (*upper row*), 3D-RISM-PSE-4 (*middle row*), and NLPB (*lower row*). The shells isolate density extending radially from the DNA center of mass for 0–5 Å (*left*), 5–7.5 Å (*center*), and 7.5–15 Å (*right*) and correspond to regions identified in Figs. 3 and 5. Isosurfaces are at 4 (*blue*), 16 (*yellow*), and 64 times (*orange*) the bulk density. To see this figure in color, go online.

cylindrical radial distribution functions (cRDFs). cRDFs are able to capture the layered distribution of ions with respect to the DNA main axis(see Fig. 5).

In general, the overall shape of distributions does not change drastically with salt concentration, but the relative values of the density near the DNA with respect to bulk densities increase with decreasing concentration. Distributions obtained from 3D-RISM and MD simulations show a high degree of similarity, being highly structured, unlike those obtained from NLPB calculations.

2D untwisted densities

The density of ions and water in the average plane of a basepair can reveal useful information on the types of interactions in the major and minor grooves, as well as around the phosphates. This can be realized by analyzing the ion or water densities for each basepair plane, but as the number of basepairs increases, this becomes a complex task. Fig. 3 *B* shows a distribution in a plane perpendicular to the helical axis and averaged over the entire length of the DNA. Although one can roughly identify binding to the grooves and phosphates, the helical disposition of basepairs makes this type of representation hard to interpret. Here, we deconvolute this type of mapping through untwisting. First, we define a twisting rate ($\dot{\Omega}$) as the ratio between the average helical twist (Ω) and rise (Dz). Helical twist is defined as the angle between the C1′-C1′ vectors of adjacent basepairs. The rise is defined as the distance between the planes of adjacent basepairs. Second, each position that is mapped is transformed through a rotation around the helical axis (here chosen to be the z axis) that removes its corresponding twist:

$$\mathbf{r_{ut}} = R_z^{-1}\left(\dot{\Omega} z\right) \times \mathbf{r},$$

where $\mathbf{r_{ut}}$ is the untwisted position vector, $\mathbf{r}$ is the original position, and R_z is a transformation matrix that corresponds to a rotation around the the helical axis, z. For the B-DNA structure studied in this work, the average twisting rate, $\dot{\Omega}$, was calculated to be 0.18587 rad/Å. Untwisting is followed by averaging along a desired interval along the helical axis, z. An example of an untwisted Na^+ density is shown in Fig. 3 *C* overlapped with the averaged untwisted positions of all heavy atoms of the DNA molecule. The effect of untwisting can be easily monitored using the C1′ atom positions in the basepair planes (Fig. 3 *C*, *green dots*). Before untwisting, the C1′ atoms are distributed along a circle, but afterward they are localized in two well defined points located at an approximate distance of 10 Å, which is typical for a Watson-Crick basepair. Although not explicitly shown, untwisting has a minimal effect on the other nucleobase atoms. Untwisting has a larger distortion effect on the DNA backbone.

Ion bridges in the major and minor grooves

MD simulations and 3D-RISM show that both water and cations penetrate into the minor and major grooves, forming well defined interaction patterns. With water treated implicitly, NLPB is able to place ions in the grooves, albeit in

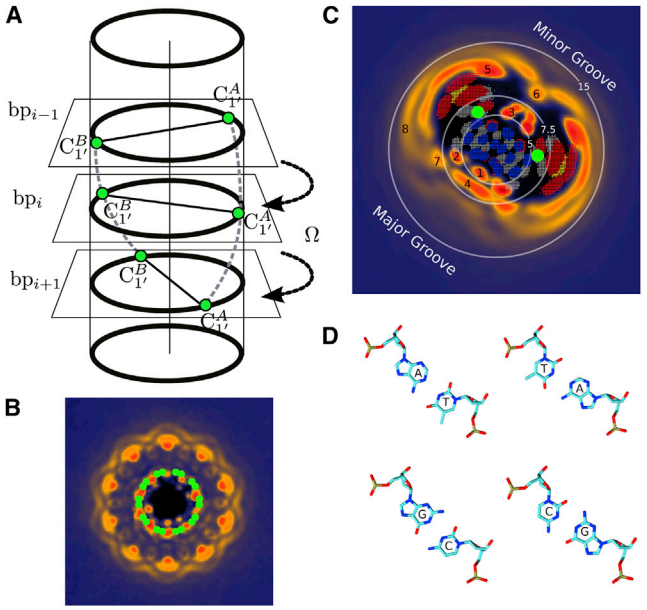

FIGURE 3 Construction of 2D untwisted density distributions. (*A* and *B*) Helical disposition of DNA basepairs (*A*) leads to a convoluted map of Na^+ ions (*B*). (*C*) Removing the helical twist angle Ω (i.e., untwisting) for each plane perpendicular to the main axis leads to a better resolved average distribution. Untwisting has a relatively minor effect on the relative distribution of atoms belonging to an average basepair, as shown by the dot clusters (*red*, oxygen; *blue*, nitrogen; *gray*, carbon; *yellow*, phosphorus). For example, although the C1′ atoms (*green circles*) are distributed nearly uniformly along a circle in the original mapping (*B*), untwisting leads to two localized positions (*C*). Numbered regions are discussed in text. (*D*) Orientation of typical basepairs. The ion densities shown here are derived from calculations using a bulk sodium chloride concentration of 0.17 M. The color scale for ion density can be found in Fig. 4.

smaller amounts and with less structure. cRDFs obtained from MD simulations and 3D-RISM possess two peaks located in the minor and major grooves (see Fig. 5, *red-* and *green-shaded areas*, respectively). The peaks are located in the interval $r < 5.0$ Å and $5.0 < r < 7.5$ Å, respectively (see Fig. 2 for the 3D density and Fig. 5 for the corresponding cRDFs). Only the second peak is present in NLPB cylindrical distribution functions. Although 3D-RISM and MD simulations show very similar sodium binding patterns in the grooves, 3D-RISM underestimates the number of bound cations for the second peak and slightly overestimates the height of the first.

Untwisted densities can isolate the types of interactions that lead up to these 1D profiles. In Fig. 3, the circles that delimit the aforementioned two regions are overlaid on untwisted distributions, and one can easily identify the corresponding types of interactions. For both MD simulations and the RISM, the peak located at $r<5$ Å corresponds to sodium ions located in the major groove. These form ion bridges (30) through inner-sphere coordination of two G:O6 atoms belonging exclusively to 5′-GC-3′/3′-CG-5′ basepair steps. Sodium ions involved in such interaction are located in the region labeled 1 in Fig. 3. Ion-bridge binding patterns have been identified previously from MD simulations of DNA and RNA duplexes with K as counter- and coions and were found to be specific to (*d*,*r*)(5′-GC-3′/3′-CG-5′) and *r*(5′-AU-3′/3′-UA-5′) basepair steps (30,32).

The peak located between 5 Å and 7.5 Å corresponds to several types of interactions between sodium ions and the minor and major grooves. MD simulations and 3D-RISM show similar binding patterns, as can be seen in Figs. 2 and 4. On the major-groove side, sodium ions in region 2 in Fig. 3 *C* make inner-sphere coordination with G:N7 and A:N7 atoms. Region 4 includes sodium ions at solvent separation from DNA, making outer-sphere contacts with hydrogen-bond acceptors in the groove. A significant water density (twice the bulk density) is located approximately between regions 1 and 2 on one side and region 4 on the other. This indicates that water is able to mediate the outer-sphere interactions of sodium ions in region 4 and compete with sodium ions in region 1 or 2 by sharing their hydrogen atoms for hydrogen-bond formations. In the minor groove, sodium ions are located in region 3, and interactions are more spatially restricted than those in the major groove. The main interacting patterns correspond to chelating inner-sphere coordination between sodium ions and N_i:O4′ and C/T_{i+1}:O2. Water is able to compete for the same sites and adopts density values 1.5–2.0 times larger than those in the bulk. Water molecules that are part of the first two solvation layers of the minor groove of B-DNA are usually associated with the term spine of hydration or hydration ribbon(s) and have been the focus of intensive and extensive experimental and theoretical research on aspects such as impact on DNA stability or sequence specificity (reviewed

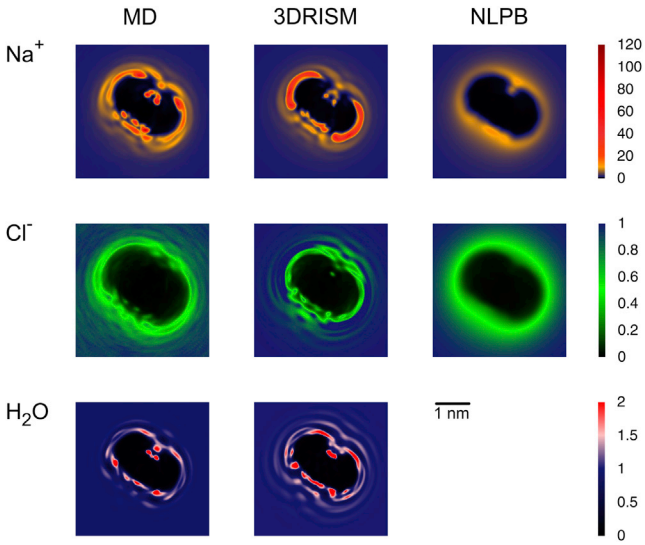

FIGURE 4 Untwisted ion and water normalized density distributions at 0.17 M NaCl. Results from MD with SPC/E (*left*), 3D-RISM-PSE-4 (*middle*), and NLPB (*right*) for sodium (*upper row*), chloride (*middle row*), and water oxygen (*lower row*). Due to the continuum treatment of the water dielectric effect, it is not possible to determine the distribution of water from NLPB theory in a straightforward manner.

in monographs such as Neidle (103), Blackburn et al. (104), and Sidel et al. (105). The most important experimentally verified findings that MD and 3D-RISM are able to capture are the ability of water to penetrate and bind specifically in the groove and the competition with monovalent ions for the aforementioned chelating interactions.

Interaction with phosphates

Interactions between cations and phosphates outside the grooves have the largest impact on the shape of the molecular distribution functions and on estimated values of PIPs (see next section). The most important interactions relevant for ion-counting measurements are those between cation phosphate oxygen atoms. Isosurfaces derived from 3D distribution functions show that the binding patterns of cations to the phosphates fits the territorial binding picture with a multitude of binding patterns, making direct contact or sitting at one or two solvation layers away from the phosphates. There are not many sources of structural experimental data that can expose the detailed distributions of both ions and water in the immediate vicinity of helical nucleic acids. Anomalous small-angle x-ray scattering (2,3), for example, can be used to model the RDF between the electronic density of the nucleic acid and the excess number of surrounding ions out of scattering profiles, but no measurement has been made on a system such as that used in this work.

Ions residing outside the grooves but in the immediate vicinity of phosphates are included in region 3 of the cRDF (see Fig. 5), exhibiting a peak at ~10 Å from the helical axis. The 3D-RISM yields the largest peak height, followed by MD simulations and finally NLPB calculations, which reveal a broad unstructured distribution. In addition, 3D-RISM distribution has the tightest shape.

Again, the untwisted densities provide insight into the interaction patterns and reveal further differences between the three methods (Fig. 4). First, distributions from MD simulations and 3D-RISM are layered and show more cationic density accumulated around phosphates. That is not the case with conventional NLPB calculations, which predict almost uniform distribution of cations around phosphates and the grooves. Furthermore, NLPB calculations fail to reveal any layering to distinguish between inner- and outer-sphere coordination between phosphates and sodiums. Second, one can explain the subtle differences between cRDFs originating from MD simulations and those from 3D-RISM (see Fig. 5). The 3D-RISM predicts a higher and more uniform accumulation of cations that make inner-sphere contacts than does MD simulation, as can be seen in the untwisted densities in region 5 of Fig. 3. MD predicts a higher density of cations making inner-sphere contacts around the minor groove. Further, for cations sitting at solvent separation from the phosphates (Fig. 3, *region 8*), MD simulation shows a higher density than 3D-RISM.

Correlation radius

The effect of the highly negatively charged nucleic acid on the cation distribution extends beyond the region where cations make inner- or outer-sphere contacts with the phosphates. In Fig. 5, one can observe that the sodium ion densities located outside of region 3 of cylindrical RDFs are different from those of the bulk for 10–20 Å more, depending on the salt concentration. As such, the correlation radius of the B-DNA molecule extends 25–35 Å away from its helical axis. As shown in Fig. 5, MD simulations, 3D-RISM, and NLPB calculations suggest that the correlation radius of the B-DNA molecule increases with decreasing monovalent salt concentration: at 0.7–0.8 M

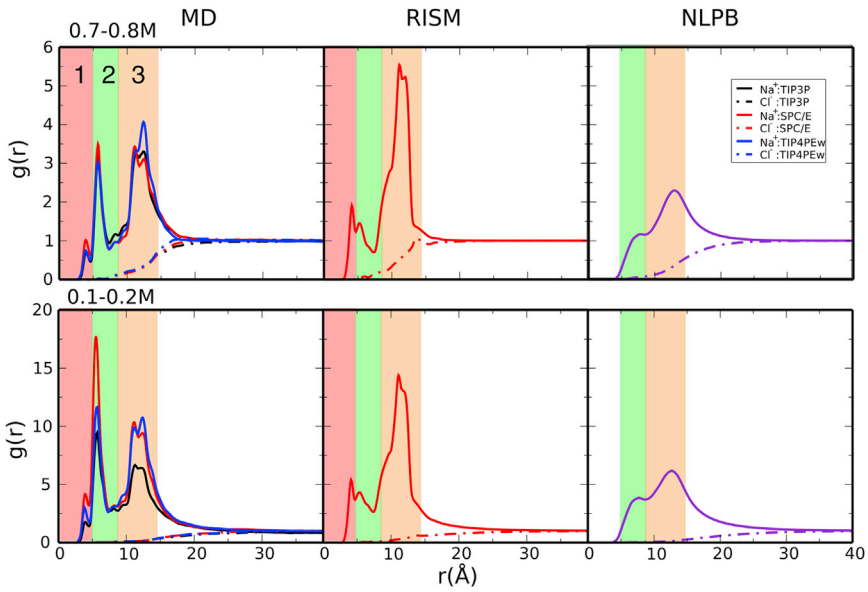

FIGURE 5 Na^+ cylindrical RDFs obtained from explicit-solvent MD simulations (*left*), 3D-RISM (*middle*), and NLPB (*right*) at ~0.2 M (*upper*) and ~0.73 M (*lower*). A cylindrical subvolume ($r = 40$ Å and $z = 60$ Å) of the simulations was used to avoid end effects. Note that the normalization factor (salt bulk concentration) is slightly different (see Table S2) for each combination of water and ion models. To see this figure in color, go online.

NaCl one can estimate a correlation radius of 25–28 Å, and at 0.1–0.2 M the value increases to 35 Å.

Preferential Interaction Parameters (PIPs)

Here we compare the estimates of ion counting from MD simulations, 3D-RISM, NLPB calculations, and experiment. Overall, we find that MD simulations and 3D-RISM are able to match results of ion-counting experiments over a range of concentrations, whereas NLPB calculations cannot (see Fig. 6 and Table S4). Good agreement between MD simulations and experiment has been found in recent studies on RNA duplexes (28,29). It is important to note in making this comparison that the estimated error in experimental results increases considerably with salt concentration. This is reflected in the net charge of the Na^+ and Cl^- (not shown), which is within error of 46 *e* but can deviate by several units.

For concentrations close to the physiological value, all MD estimates match experimental results within error but begin to fail at higher concentrations. In general, the calculations that use TIP4P-Ew and SPC/E water models are closest to experiment. TIP4P-Ew, in particular, is within or near experimental error for both Na^+ and Cl^-. SPC/E is close to experimental error for Cl^- only and TIP3P provides excess numbers that are too low. At 0.2 M, SPC/E and TIP4P-Ew are nearly indistinguishable at 0.2 M, whereas TIP3P gives a lower result. The fact that 0.2 M is near the ideal ion regime is likely a significant contributor to the success of the various models at this concentration. For 0.73 M, only TIP4P-Ew has reasonable agreement with the experimental PIPs for both Na^+ and Cl^-. SPC/E is nearly within error for the Cl^- PIP but falls well short of the Na^+ value. TIP3P greatly underestimates the number of excess Na^+ at this higher concentration. The larger van der Waals radius for Na^+ in JC:TIP3P relative to the other parameter sets is a likely contributor to this result.

The 3D-RISM has good agreement with experimental results for all but the lowest concentrations. It also agrees well with SPC/E MD simulations at 0.1842 M but not at 0.7452 M. As noted in 4.1.1, when compared with MD simulations, 3D-RISM calculations underestimate the minor- and major-groove binding of ions and instead preferentially place ions around the phosphate backbone.

The NLPB model, on the other hand, fails over almost the entire concentration range. This was not evident in some previous studies (4,29) due to the method of calculating the number of excess ions and the ion radii that were used. As discussed in 2.1, when calculating PIPs in molar concentration units, it is important to integrate over the entire volume, including the solvent-excluded region in the DNA interior, when using Eqs. 6 and 16. For example, at 1 M, a volume equivalent to the ion-excluded region of the 24L DNA fragment in our NLPB calculations contains $\sim$ 20 Na^+. Omitting the ion-excluded region from the integral gives a PIP of $\sim$ 34 rather than $\sim$ 14 calculated from the all-space integral. Previous studies using NLPB methods have generally omitted the ion-excluded volume from PIP calculations and, combined with their use of large ion radii, have achieved good fits to experimental molar data. The ion radii used were typically >4 Å, which is equivalent to having ions completely solvent-separated from the DNA surface. In contrast, MD simulations and 3D-RISM show ions in direct contact with DNA (see Fig. 4). Regardless of whether adjusting parameters would allow NLPB to obtain better agreement with ion-counting experiments, it is clear that the ion distributions using NLPB qualitatively lack structural details that are available only when correlation between solution components is considered.

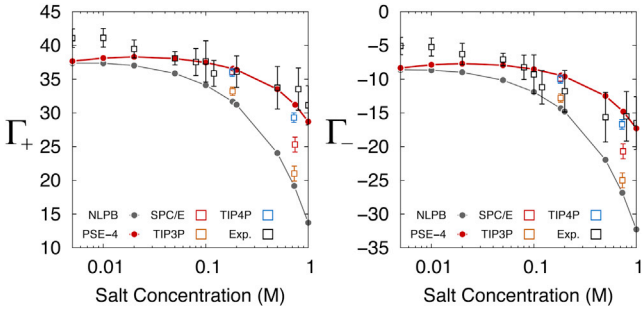

FIGURE 6 Comparison between theoretical estimates of Na^+ (*left*) and Cl^- (*right*) PIPs and experimental results from Bai et al. (4). To see this figure in color, go online.

CONCLUSIONS

The ion atmosphere around nucleic acids has tremendous importance for supporting structure, dynamics, or catalytic activity. Neutralization of the high and extended charge density of nucleic acids when they are brought into an aqueous environment is realized through two mechanisms: counterion condensation and anion depletion. Simulations can complement structural data extracted from x-ray, NMR, SAXS, and other measurements, and can provide an atomic-level description of the dynamics associated with the accumulation of ions around nucleic acids.

Herein, the results of MD simulations and 3D-RISM and NLPB calculations are analyzed and compared with available ion-counting measurements. The results for the NLPB model are markedly different from those of the two molecular-based theories. The net association of counterions is much smaller than with either the other theories or experiment at salt concentrations >0.4 M. However, regardless of the ion-counting numbers, the ion distributions themselves from NLPB are qualitatively inconsistent (at all concentrations studied here) with those from MD or 3D-RISM models, since the former all but lack any defined structure. These comparisons point to real limitations in the use of continuum solvent models in this field.

We present two methods to compute PIPs. The first approach uses KB integrals and depends on the local density of salt and water around the solute. The second approach uses two distinct spatial partitions, similar in spirit to the setup in osmosis experiments. In this case, one can estimate the PIPs by knowing the bulk concentrations of salt and water only. This approach might be well suited for irregularly shaped or flexible solutes for which distribution function might be harder to define.

At salt concentrations close to physiological values, the estimates for ion counting from MD and 3D-RISM are within error of the experimentally determined values. The ion distributions from MD simulations and 3D-RISM calculations, from which PIPs are computed, are largely compatible, clearly showing a multilayered distribution and distinct ionic binding sites in the major and minor grooves. However, there are some discrepancies in the details of these distributions, such as the propensity of 3D-RISM to overaccumulate sodiums around the phosphate backbone, its slightly reduced density maxima, and its overrepresentation of chlorides near the DNA surface. At concentrations close to 1 M, MD simulations and 3D-RISM calculations underestimate the values, with JC:TIP4-PEw and JC:SPC/E parameter sets being closer to experiment.

These results provide an overall picture of how simple helical nucleic acids are neutralized by the ion atmosphere in electrolyte solutions. The distribution of ions and water is layered outside the grooves and has well defined, sequence-dependent competing binding modes in both the minor and major grooves. The results presented here lay the ground for future efforts to improve the representation of DNA-ion interactions over larger concentration ranges, and to begin to dissect the contribution of the ion atmosphere in RNA folding.

SUPPORTING MATERIAL

Four tables and four figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(14)00091-5.

This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575. We thank Lois Pollack for many useful discussions.

This work was supported by National Institutes of Health grants GM57513, GM103297(D.A.C.), and P01GM066275(D.M.Y. and D.H.).

REFERENCES

1. Draper, D. E., D. Grilley, and A. M. Soto. 2005. Ions and RNA folding. *Annu. Rev. Biophys. Biomol. Struct.* 34:221–243.
2. Pabit, S. A., S. P. Meisburger, …, L. Pollack. 2010. Counting ions around DNA with anomalous small-angle x-ray scattering. *J. Am. Chem. Soc.* 132:16334–16336.
3. Pollack, L. 2011. SAXS studies of ion-nucleic acid interactions. *Annu. Rev. Biophys.* 40:225–242.
4. Bai, Y., M. Greenfeld, …, D. Herschlag. 2007. Quantitative and comprehensive decomposition of the ion atmosphere around nucleic acids. *J. Am. Chem. Soc.* 129:14981–14988.
5. Greenfeld, M., and D. Herschlag. 2009. Probing nucleic acid-ion interactions with buffer exchange-atomic emission spectroscopy. *Methods Enzymol.* 469:375–389.
6. Grilley, D., A. M. Soto, and D. E. Draper. 2006. Mg^{2+}-RNA interaction free energies and their relationship to the folding of RNA tertiary structures. *Proc. Natl. Acad. Sci. USA.* 103:14003–14008.
7. Bleam, M. L., C. F. Anderson, and M. T. Record, Jr. 1980. Relative binding affinities of monovalent cations for double-stranded DNA. *Proc. Natl. Acad. Sci. USA.* 77:3085–3089.
8. Braunlin, W. H., C. F. Anderson, and M. T. Record, Jr. 1987. Competitive interactions of $Co(NH3)_6^{3+}$ and Na^+ with helical B-DNA probed by ^{59}Co and ^{23}Na NMR. *Biochemistry.* 26:7724–7731.
9. Lyubartsev, A. P. 2004. Molecular simulations of DNA counterion distributions. *In* Dekker Encyclopedia of Nanoscience and Nanotechnology. J. A. Schwarz, S. E. Lyshevski, K. Putyera, and C. Contescu, editors. CRC Press, Boca Raton, FL, pp. 2131–2143.
10. Ponomarev, S. Y., K. M. Thayer, and D. L. Beveridge. 2004. Ion motions in molecular dynamics simulations on DNA. *Proc. Natl. Acad. Sci. USA.* 101:14771–14775.
11. Manning, G. S. 1979. Counterion binding in polyelectrolyte theory. *Acc. Chem. Res.* 12:443–449.
12. Manning, G. S. 1978. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.* 11:179–246.
13. Anderson, C. F., and M. T. Record, Jr. 1995. Salt-nucleic acid interactions. *Annu. Rev. Phys. Chem.* 46:657–700.
14. Manning, G. S. 2011. A counterion condensation theory for the relaxation, rise, and frequency dependence of the parallel polarization of rodlike polyelectrolytes. *Eur Phys J E Soft Matter.* 34:1–7.
15. Manning, G. S. 2007. Counterion condensation on charged spheres, cylinders, and planes. *J. Phys. Chem. B.* 111:8554–8559.
16. Manning, G. S. 2007. Electrostatic free energies of spheres, cylinders, and planes in counterion condensation theory with some applications. *Macromolecules.* 40:8071–8081.
17. Draper, D. E. 2008. RNA folding: thermodynamic and molecular descriptions of the roles of ions. *Biophys. J.* 95:5489–5495.
18. Chen, A. A., D. E. Draper, and R. V. Pappu. 2009. Molecular simulation studies of monovalent counterion-mediated interactions in a model RNA kissing loop. *J. Mol. Biol.* 390:805–819.
19. Kirmizialtin, S., and R. Elber. 2010. Computational exploration of mobile ion distributions around RNA duplex. *J. Phys. Chem. B.* 114:8207–8220.
20. Korolev, N., A. P. Lyubartsev, …, L. Nordenskiöld. 2003. A molecular dynamics simulation study of oriented DNA with polyamine and sodium counterions: diffusion and averaged binding of water and cations. *Nucleic Acids Res.* 31:5971–5981.
21. Rueda, M., E. Cubero, …, M. Orozco. 2004. Exploring the counterion atmosphere around DNA: what can be learned from molecular dynamics simulations? *Biophys. J.* 87:800–811.
22. Sen, S., D. Andreatta, …, M. A. Berg. 2009. Dynamics of water and ions near DNA: comparison of simulation to time-resolved Stokes-shift experiments. *J. Am. Chem. Soc.* 131:1724–1735.
23. Chen, A. A., M. Marucho, …, R. V. Pappu. 2009. Simulations of RNA interactions with monovalent ions. *Methods Enzymol.* 469:411–432.
24. Bonvin, A. M. 2000. Localisation and dynamics of sodium counterions around DNA in solution from molecular dynamics simulation. *Eur. Biophys. J.* 29:57–60.
25. Dai, L., Y. Mu, …, J. R. van der Maarel. 2008. Molecular dynamics simulation of multivalent-ion mediated attraction between DNA molecules. *Phys. Rev. Lett.* 100:118301.
26. Luan, B., and A. Aksimentiev. 2008. DNA attraction in monovalent and divalent electrolytes. *J. Am. Chem. Soc.* 130:15754–15755.

27. Noy, A., I. Soteras, …, M. Orozco. 2009. The impact of monovalent ion force field model in nucleic acids simulations. *Phys. Chem. Chem. Phys.* 11:10596–10607.

28. Kirmizialtin, S., S. A. Pabit, …, R. Elber. 2012. RNA and its ionic cloud: solution scattering experiments and atomically detailed simulations. *Biophys. J.* 102:819–828.

29. Kirmizialtin, S., A. R. J. Silalahi, …, M. O. Fenley. 2012. The ionic atmosphere around A-RNA: Poisson-Boltzmann and molecular dynamics simulations. *Biophys. J.* 102:829–838.

30. Auffinger, P., and E. Westhof. 2000. Water and ion binding around RNA and DNA (C,G) oligomers. *J. Mol. Biol.* 300:1113–1131.

31. Yoo, J., and A. Aksimentiev. 2012. Improved parametrization of Li^+, Na^+, K^+, and Mg^{++} ions for all-atom molecular dynamics simulations of nucleic acid systems. *J. Phys. Chem. Lett.* 3:45–50.

32. Auffinger, P., and E. Westhof. 2001. Water and ion binding around r(UpA)12 and d(TpA)12 oligomers—comparison with RNA and DNA (CpG)12 duplexes. *J. Mol. Biol.* 305:1057–1072.

33. Chu, V. B., Y. Bai, …, S. Doniach. 2007. Evaluation of ion binding to DNA duplexes using a size-modified Poisson-Boltzmann theory. *Biophys. J.* 93:3202–3209.

34. Pabit, S. A., X. Qiu, …, L. Pollack. 2009. Both helix topology and counterion distribution contribute to the more effective charge screening in dsRNA compared with dsDNA. *Nucleic Acids Res.* 37:3887–3896.

35. Bond, J. P., C. F. Anderson, and M. T. Record, Jr. 1994. Conformational transitions of duplex and triplex nucleic acid helices: thermodynamic analysis of effects of salt concentration on stability using preferential interaction coefficients. *Biophys. J.* 67:825–836.

36. Gonzales-Tovar, E., M. Lozada-Cassou, and D. Henderson. 1985. Hypernetted chain approximation for the distribution of ions around a cylindrical electrode. II. Numerical solution for a model cylindrical polyelectrolyte. *J. Chem. Phys.* 83:361–372.

37. Maruyama, Y., N. Yoshida, and F. Hirata. 2010. Revisiting the salt-induced conformational change of DNA with 3D-RISM theory. *J. Phys. Chem. B.* 114:6464–6471.

38. Howard, J. J., G. C. Lynch, and B. M. Pettitt. 2011. Ion and solvent density distributions around canonical B-DNA from integral equations. *J. Phys. Chem. B.* 115:547–556.

39. Maruyama, Y., T. Matsushita, …, F. Hirata. 2011. Solvent and salt effects on structural stability of human telomere. *J. Phys. Chem. B.* 115:2408–2416.

40. Beglov, D., and B. Roux. 1997. An integral equation to describe the solvation of polar molecules in liquid water. *J. Phys. Chem. B.* 101:7821–7826.

41. Kovalenko, A., and F. Hirata. 1999. Self-consistent description of a metal-water interface by the Kohn-Sham density functional theory and the three-dimensional reference interaction site model. *J. Chem. Phys.* 110:10095–10112.

42. Kovalenko, A., and F. Hirata. 2000. Potentials of mean force of simple ions in ambient aqueous solution. I. Three-dimensional reference interaction site model approach. *J. Chem. Phys.* 112:10391–10402.

43. Kovalenko, A. 2003. Three-dimensional RISM theory for molecular liquids and solid-liquid interfaces. *In* Molecular Theory of Solvation. F. Hirata, editor. Kluwer Academic, Norwell, MA, pp. 265–268.

44. Luchko, T., S. Gusarov, …, A. Kovalenko. 2010. Three-dimensional molecular theory of solvation coupled with molecular dynamics in Amber. *J. Chem. Theory Comput.* 6:607–624.

45. Leipply, D., D. Lambert, and D. E. Draper. 2009. Ion-RNA interactions thermodynamic analysis of the effects of mono- and divalent ions on RNA conformational equilibria. *Methods Enzymol.* 469:433–463.

46. Record, Jr., M. T., and C. F. Anderson. 1995. Interpretation of preferential interaction coefficients of nonelectrolytes and of electrolyte ions in terms of a two-domain model. *Biophys. J.* 68:786–794.

47. Shulgin, I. L., and E. Ruckenstein. 2006. A protein molecule in a mixed solvent: the preferential binding parameter via the Kirkwood-Buff theory. *Biophys. J.* 90:704–707.

48. Smith, P. E. 2006. Equilibrium dialysis data and the relationships between preferential interaction parameters for biological systems in terms of Kirkwood-Buff integrals. *J. Phys. Chem. B.* 110:2862–2868.

49. Anderson, C. F., and M. T. Record, Jr. 1990. Ion distributions around DNA and other cylindrical polyions: theoretical descriptions and physical implications. *Annu. Rev. Biophys. Biophys. Chem.* 19:423–465.

50. Anderson, C. F., and M. T. Record, Jr. 1982. Polyelectrolyte theories and their applications to DNA. *Annu. Rev. Phys. Chem.* 33:191–222.

51. Hansen, J.-P., and I. R. McDonald. 1990. Theory of Simple Liquids, 2nd ed. Academic Press, London, pp. 97–144.

52. Chitra, R., and P. E. Smith. 2001. Properties of 2,2,2-trifluoroethanol and water mixtures. *J. Chem. Phys.* 114:426–435.

53. Ben-Naim, A. 1992. Statistical Thermodynamics for Chemists and Biochemists. Plenum Press, New York.

54. Smith, P. E. 2004. Cosolvent interactions with biomolecules: relating computer simulation data to experimental thermodynamic data. *J. Phys. Chem. B.* 108:18716–18724.

55. Ben-Naim, A. 2006. Molecular Theory of Solutions. Oxford University Press, New York.

56. Record, Jr., M. T., W. Zhang, and C. F. Anderson. 1998. Analysis of effects of salts and uncharged solutes on protein and nucleic acid equilibria and processes: a practical guide to recognizing and interpreting polyelectrolyte effects, Hofmeister effects, and osmotic effects of salts. *Adv. Protein Chem.* 51:281–353.

57. Macke, T. J., and D. A. Case. 1997. Modeling unusual nucleic acid structures. *In* Molecular Modeling of Nucleic Acids. N. B. Leontis and J. SantaLucia, editors. American Chemical Society, Washington, D.C, pp. 379–393.

58. Arnott, S., and D. W. L. Hukins. 1972. Optimised parameters for A-DNA and B-DNA. *Biochem. Biophys. Res. Commun.* 47:1504–1509.

59. Jorgensen, W. L., J. Chandrasekhar, …, M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

60. Horn, H. W., W. C. Swope, …, T. Head-Gordon. 2004. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* 120:9665–9678.

61. Horn, H. W., W. C. Swope, and J. W. Pitera. 2005. Characterization of the TIP4P-Ew water model: vapor pressure and boiling point. *J. Chem. Phys.* 123:194504.

62. Berendsen, H. J. C., J. R. Grigera, and T. P. Straatsma. 1987. The missing term in effective pair potentials. *J. Phys. Chem.* 91:6269–6271.

63. Joung, I. S., T. Luchko, and D. A. Case. 2013. Simple electrolyte solutions: comparison of DRISM and molecular dynamics results for alkali halide solutions. *J. Chem. Phys.* 138:044103.

64. Joung, I. S., and T. E. Cheatham, 3rd. 2008. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B.* 112:9020–9041.

65. Pérez, A., I. Marchán, …, M. Orozco. 2007. Refinement of the AMBER force field for nucleic acids: improving the description of α/γ conformers. *Biophys. J.* 92:3817–3829.

66. Banáš, P., D. Hollas, …, M. Otyepka. 2010. Performance of molecular mechanics force fields for RNA simulations: stability of UUCG and GNRA hairpins. *J. Chem. Theory Comput.* 6:3836–3849.

67. Wang, J., P. Cieplak, and P. A. Kollman. 2000. How well does a restrained electrostatic potential (RESP) model perform in calcluating conformational energies of organic and biological molecules? *J. Comput. Chem.* 21:1049–1074.

68. Case, D. A., T. A. Darden, …, P. A. Kollman. 2002. AMBER 10. University of California. San Francisco, San Francisco.

69. Pearlman, D. A., D. A. Case, …, P. Kollman. 1995. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structure and energetic properties of molecules. *Comput. Phys. Commun.* 91:1–41.
70. Case, D. A., T. E. Cheatham, 3rd, …, R. J. Woods. 2005. The Amber biomolecular simulation programs. *J. Comput. Chem.* 26:1668–1688.
71. Phillips, J. C., R. Braun, …, K. Schulten. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781–1802.
72. Martyna, G. J., D. J. Tobias, and M. L. Klein. 1994. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* 101:4177–4189.
73. Feller, S. E., Y. Zhang, …, B. R. Brooks. 1995. Constant pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* 103:4613–4621.
74. Essmann, U., L. Perera, …, L. G. Pedersen. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.
75. Sagui, C., and T. A. Darden. 1999. Molecular dynamics simulations of biomolecules: long-range electrostatic effects. *Annu. Rev. Biophys. Biomol. Struct.* 28:155–179.
76. Allen, M. P., and D. J. Tildesley. 1987. Computer Simulation of Liquids. Clarendon Press, Oxford, United Kingdom.
77. Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* 23:327–341.
78. Ornstein, L. S., and F. Zernike. 1914. Accidental deviations of density and opalescence at the critical point of a single substance. *Proc. Akad. Sci. Amsterdam.* 17:793–806.
79. Ornstein, L. S., and F. Zernike. 1964. Accidental deviations of density and opalescence at the critical point of a single substance. *In* The Equilibrium Theory of Classical Fluids: A Lecture Note and Reprint Volume W. A. Benjamin, New York, pp. 2–16.
80. Perkyns, J. S., and B. M. Pettitt. 1992. A dielectrically consistent interaction site theory for solvent-electrolyte mixtures. *Chem. Phys. Lett.* 190:626–630.
81. Perkyns, J. S., and B. M. Pettitt. 1992. A site-site theory for finite concentration saline solutions. *J. Chem. Phys.* 97:7656–7666.
82. Springer, J. F., M. A. Pokrant, and F. A. Stevens. 1973. Integral equation solutions for the classical electron gas. *J. Chem. Phys.* 58:4863–4867.
83. Abernethy, G. M., and M. J. Gillan. 1980. A new method of solving the HNC equation for ionic liquids. *Mol. Phys.* 39:839–847.
84. Ng, K. 1974. Hypernetted chain solutions for the classical one-component plasma up to $\Gamma = 7000$. *J. Chem. Phys.* 61:2680–2689.
85. Kinoshita, M., and F. Hirata. 1996. Application of the reference interaction site model theory to analysis on surface-induced structure of water. *J. Chem. Phys.* 104:8807–8815.
86. Perkyns, J. S., G. C. Lynch, …, B. M. Pettitt. 2010. Protein solvation from theory and simulation: exact treatment of Coulomb interactions in three-dimensional theories. *J. Chem. Phys.* 132:064106.
87. Gusarov, S., B. S. Pujari, and A. Kovalenko. 2012. Efficient treatment of solvation shells in 3D molecular theory of solvation. *J. Comput. Chem.* 33:1478–1494.
88. Morita, T. 1958. Theory of classical fluids: hypernetted chain approximation. I. Formulation for a one-component system. *Prog. Theor. Phys.* 20:920–928.
89. Rasaiah, J. C., D. N. Card, and J. P. Valleau. 1972. Calculations on the "restricted primitive model" for 1–1 electrolyte solutions. *J. Chem. Phys.* 56:248–255.
90. Hansen, J. P., and I. R. McDonald. 1975. Statistical mechanics of dense ionized matter. IV. Density and charge fluctuations in a simple molten salt. *Phys. Rev. A.* 11:2111–2123.
91. Pettitt, B. M., and P. J. Rossky. 1982. Integral equation predictions of liquid state structure for waterlike intermolecular potentials. *J. Chem. Phys.* 77:1451–1457.
92. Hirata, F., and P. J. Rossky. 1981. An extended RISM equation for molecular polar fluids. *Chem. Phys. Lett.* 83:329–334.
93. Hirata, F., B. M. Pettitt, and P. J. Rossky. 1982. Application of an extended RISM equation to dipolar and quadrupolar fluids. *J. Chem. Phys.* 77:509–520.
94. Singer, S. J., and D. Chandler. 1985. Free-energy functions in the extended RISM approximation. *Mol. Phys.* 55:621–625.
95. Kast, S. M., and T. Kloss. 2008. Closed-form expressions of the chemical potential for integral equation closures with certain bridge functions. *J. Chem. Phys.* 129:236101.
96. Case, D. A., T. A. Darden, …, P. A. Kollman. 2012. AMBER 12. University of California, San Francisco, CA.
97. Kovalenko, A., S. Ten-no, and F. Hirata. 1999. Solution of three-dimensional reference interaction site model and hypernetted chain equations for simple point charge water by modified method of direct inversion in iterative subspace. *J. Comput. Chem.* 20:928–936.
98. Söhnel, O., and P. Novotný. 1985. Densities of Aqueous Solutions of Inorganic Substances. Elsevier, Munich.
99. Millero, F. J. 1970. The apparent and partial molal volume of aqueous sodium chloride solutions at various temperatures. *J. Phys. Chem.* 74:356–362.
100. Honig, B., and A. Nicholls. 1995. Classical electrostatics in biology and chemistry. *Science.* 268:1144–1149.
101. Simonson, T. 2003. Electrostatics and dynamics of proteins. *Rep. Prog. Phys.* 66:737–787.
102. Baker, N. A., D. Sept, …, J. A. McCammon. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA.* 98:10037–10041.
103. Neidle, S. 2008. Principles of Nucleic Acid Structure. Academic Press, New York.
104. Blackburn, G. M., M. J. Gait, …, D. M. Williams. 2006. Nucleic Acids in Chemistry and Biology. Oxford University Press, New York.
105. Sigel, A., H. Sigel, and R. K. O. Sigel. 2012. Interplay between Metal Ions and Nucleic Acids. Springer, Dordrecht, The Netherlands.

Ion counting from explicit solvent simulations and 3D-RISM

George M. Giambaşu,*[¶] Tyler Luchko, *[¶] Daniel Herschlag, [†] Darrin M. York, * and David A. Case *

*[1]Department of Chemistry and Chemical Biology and BioMaPS Institute, Rutgers University, Piscataway, NJ -08854 USA; [†] Department of Biochemistry, Stanford University, Stanford, CA 94305 USA; [¶] These authors contributed equally
GM Giambaşu and T Luchko contributed equally to this work.
Present address: T Luchko's present address is Department of Physics and Astronomy, California State University Northridge, Northridge, CA 91330.

Convergence of preferential interaction parameters

Preferential interaction parameters (PIPs) can be calculated using KB integrals, as done for 3D-RISM and NLPB, or using the two-partition method, which is convenient for MD. The convergence of both methods have a distance dependence that can be examined using cRDFs and applying the KB approach. The convergence of Eq 3 as a function of radial distance without long-range correction is shown in Figure S1. All three simulation methods have a similarly long-ranged g(r) and a discernible non-zero slope is still evident in $\Gamma^{+}(r)$ at r = 40Å. This gives an indication of both the extent of the solvation box required to achieve the desired precision, as used for MD and NLPB, or the need for a long-range correction, such as 17 used for 3D-RISM.

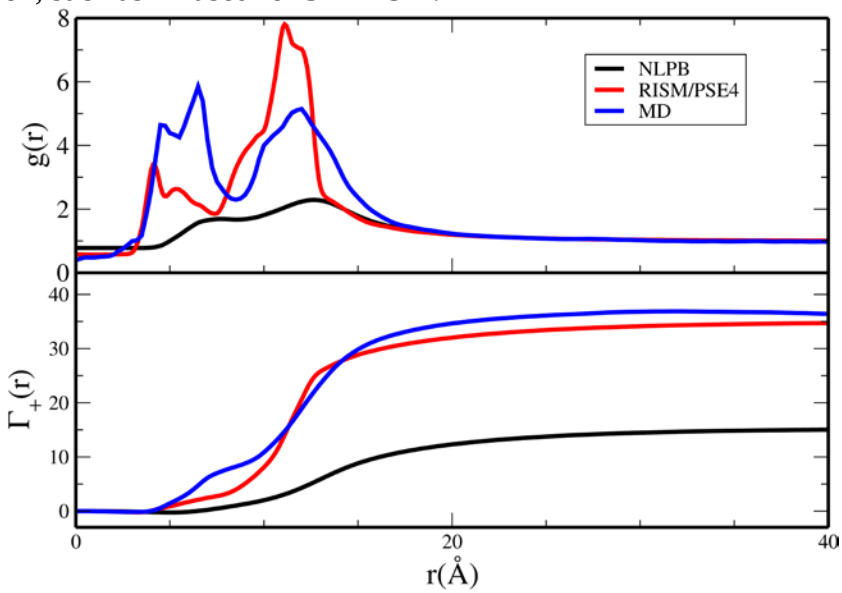

Figure S1. Cylindrical radial distribution function, g(r), (top) and its corresponding preferential interaction parameter, $\Gamma^{+}(r) = 2\pi h \int_0^r (g(r)-1) r' dr'$ (bottom) from MD, 3D-RISM and NLPB. The KB integrals reach their limiting values for r~40Å.

Numerical Precision of 3D-RISM Calculations

Numerical precision of 3D-RISM calculations depend on the grid spacing, residual tolerance and solvation box size. For a given, solvation box, a grid spacing of 0.5 Å and residual tolerance of 10-6 provide PIP precisions lower than 10-3 ions. The buffer region between the solute and nearest box edge, which determines the size of the solvent box, is then the only remaining determinate of numerical precision. The extent of the solvent buffer required to achieve the desired precision varies with salt concentration. Figure S2 shows the sodium PIP dependence on buffer size at 0.1, 0.01 and 0.005 M NaCl for 3D-RISM-KH. Differences are calculated relative to the largest buffer size tested at each concentration. The closure approximation used was found to have negligible effect on the precision. Note that the PIP for sodium is always over-estimated as the box size is reduced.

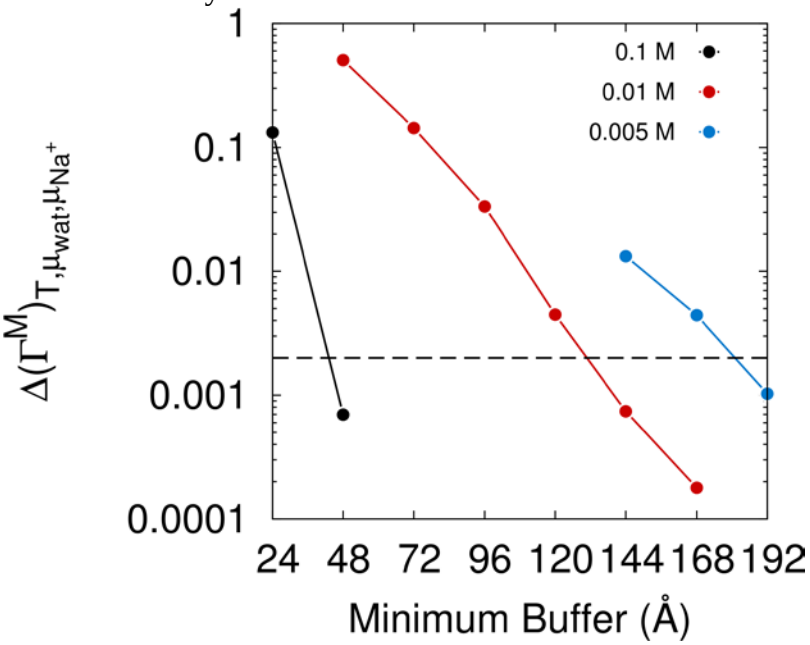

Figure S2. Difference in PIP for 3D-RISM-KH as a function of solvent box buffer size. Differences are calculated relative to references buffer sizes of 70 Å (0.1M NaCl), 192 Å (0.01M NaCl) and 216 Å (0.005M). The horizontal dashed line indicates a difference of 0.002 ions from the reference calculation.

Impact of closure relation on ion counting estimates from 3D-RISM

The PSE-4 closure relation used in the main manuscript is part of the partial series expansion of order-*n* (PSE-*n*) family of closures(1) that approximate the hypernetted-chain equation (HNC)(2) by including *n*-terms of a Taylor expansion. PSE-4 was the highest-order closure that could be converged in this study. In Figure S3 we compare the ion counting concentration profiles for lower order closures :KH or PSE-1, PSE-2 and PSE3. Irrespective of the salt concentration, the preferential interaction parameter increase with the number of terms included in the expansion. This behavior is consistent with attractive interaction being enhanced as the PSE-n closure approximation approaches the HNC.

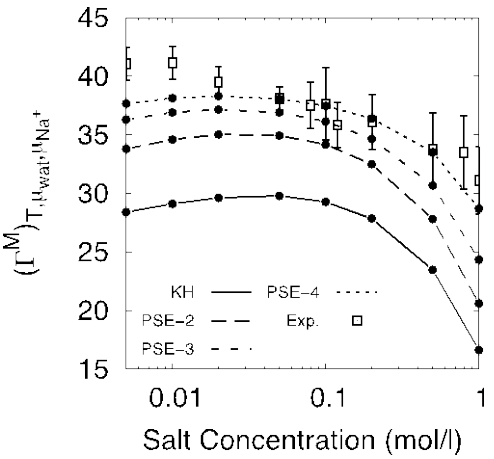

Figure S3: Closure dependence of sodium preferential interaction paramters for 3D-RISM calculations. Experimental results from (3).

Force Field	Duration (ns)	$M^{(2)}_{NaCl}$	$M^{(2)}_{H2O}$	M_{NaCl}
0.7 NaCl				
JC:TIP3P	100	0.7298	53.4867	0.7109
JC:TIP4P-Ew	150	0.7288	53.8447	0.7157
JC:SPC/E	120	0.7452	54.8205	0.7287
0.2 NaCl				
JC:TIP3P	100	0.1829	53.8635	0.1731
JC:TIP4P-Ew	100	0.1830	54.0004	0.1752
JC:SPC/E	120	0.1842	53.9492	0.1739

Table S1. Summary of the explicit solvent simulations and key properties. $M^{(2)}_{NaCl}$ and $M^{(2)}_{H2O}$ are the molarities of sodium chloride and water in the bulk; see text for discussion on determining the properties of the bulk. M_{NaCl} is the salt molarity if calculated by dividing the number of moles with the total volume of the cell.

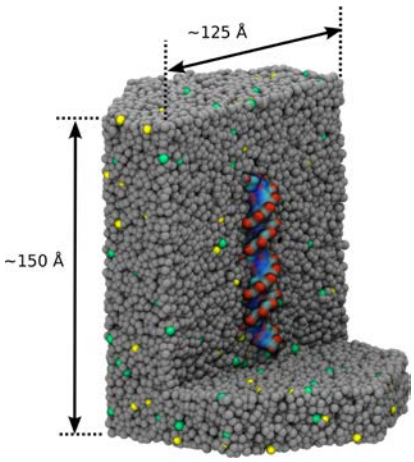

Figure S4 Section through the molecular dynamics simulations box. A single B-DNA molecule (shown in surface representation) is placed in a hexagonal prism periodic box that extends ~60 Å in the direction perpendicular on the DNA main axis and 30 Å in each direction along the main axis from the DNA. Water is represented as gray spheres, sodium ions as yellow spheres, chloride ions as green spheres.

Salt Concentration (M)	Requested Buffer (Å)	Actual Buffer (Å)
0.005	192	192.696 x 214.708 x 227.678
0.01-0.05	144	147.696 x 169.708 x 173.678
0.1-5	48	50.196 x 49.708 x 53.678

Table S2. Summary of 3D-RISM calculations. The requested buffer sizes reported here were automatically increased the actual buffer size to evenly distribute memory for parallel calculations and satisfy FFT grid size constraints. To maximize the parallel efficiency of the 3D-RISM code, the DNA molecule was oriented along the z-axis. However, due to the load balancing requirements of the parallel implementation, the buffer region was typically larger than requested by a small amount as it was automatically increased to the minimum buffer size that gave an even distribution of memory among parallel processes and ensure the grid dimensions were multiples of small prime numbers for the FFT calculation. The size of the buffer region required was found to be independent of the closure used and only depended on ion concentrations.

Salt Concentration (M)	Grid Dimensions (Å)	Grid Spacing (Å)
0.005-0.05	514 x 514 x 610	2
0.1-0.5	257 x 257 x 353	1
1-5	129 x 129 x 225	1

Table S3. Summary of NLPB calculations.

$M^{(2)}_{NaCl}$	0.18-0.2M		0.7-0.8M	
	$\Gamma^{(M)}_{Na^+}$	$\Gamma^{(M)}_{Cl^-}$	$\Gamma^{(M)}_{Na^+}$	$\Gamma^{(M)}_{Cl^-}$
Experiment	36.1(3.1)	-11.8(2.3)	33.5(3.7)	-15.5(3.1)
MD JC:TIP3P	33.2(0.6)	-12.8(0.6)	21.0(1.1)	-25.0(1.1)
MD JC:SPC/E	36.1(0.5)	-9.9(0.5)	25.3(1.1)	-20.7(1.1)
MD JC:TIP4P	36.0(0.6)	-10.0(0.6)	29.3(0.7)	-16.7(0.7)
3D-RISM JC:SPC/E	36.6	-9.4	31.2	-14.8
NLPB	31.7	-14.3	19.2	-26.8

Table S4. Estimation of preferential interaction parameters from MD, 3D-RISM and NLPB and comparison with experiment.

1. Kast, S.M., and T. Kloss. 2008. Closed-form expressions of the chemical potential for integral equation closures with certain bridge functions. J. Chem. Phys. 129: 236101.

2. Morita, T. 1958. Theory of Classical Fluids: Hyper-Netted Chain Approximation, I Formulation for a One-Component System. Prog. Theor. Phys. 20: 920–938.

3. Bai, Y., M. Greenfeld, K.J. Travers, V.B. Chu, J. Lipfert, et al. 2007. Quantitative and Comprehensive Decomposition of the Ion Atmosphere around Nucleic Acids. J. Am. Chem. Soc. 129: 14981–14988.