Published in final edited form as:

J Phys Chem B. 2011 January 27; 115(3): 547–556. doi:10.1021/jp107383s.

Ion and solvent density distributions around canonical B-DNA from integral equations

Jesse J. Howard, Gillian C. Lynch, and B. Montgomery Pettitt

Department of Chemistry, University of Houston, Houston, Texas 77204-5003

Abstract

We calculate the water and ion spatial distributions around charged oligonucleotides using a renormalized three-dimensional reference interaction site theory coupled with the HNC closure. Our goal is to understand the balance between inter-DNA strand forces and solvation forces as a function of oligonucleotide length in the short strand limit. The DNA is considered in aqueous electrolyte solutions of 1 M KCl, 0.1 M KCl or 0.1 M NaCl. The current theoretical results are compared to MD simulations and experiments. It is found that the IE theory replicates the MD and the experimental results for the base-specific hydration patterns in both the major and minor grooves. We are also able to discern characteristic structural pattern differences between Na⁺ and K⁺ ions. When compared to Poisson-Boltzmann methods the IE theory, like simulation, predicts a richly structured ion environment which is better described as multi-layer rather than double-layer.

I. INTRODUCTION

The aqueous saline environment is known to play a central role in providing the stability for different DNA conformations and configurations.¹ Diffraction experiments² and simulations³ have shown that the solvent distributions in the minor and major grooves are structured in specific ways which are dependent on the type of the basepair and the form of the DNA duplex. Counter-ions play a critical role in DNA structures by shielding the negatively charged phosphates from their complementary strands.⁴ The ions also form specific structural patterns depending on local DNA structural regions. In order to accurately model the interactions of DNA with other cellular components, such as proteins, theoretical models must accurately incorporate the effects of the solvent and ion species into the thermodynamics calculations. In this paper we show the ability of the three-dimensional integral equations (3D-IEs) to predict the solvent and ion distributions specific to DNA structures and their influence on the corresponding solvent thermodynamics.

Two options in wide use for DNA modeling are molecular dynamics (MD) simulations and Poisson-Boltzmann (PB) theory. MD simulations usually give the most accurate results for a model and can provide a wealth of information regarding the dynamics and structural details of the system. The problem with MD simulations, however, is the time needed for proper sampling, especially for an explicit solvent model with long time events, such as ion diffusion.⁵ PB calculations, on the other hand, are computationally fast but that speed comes at a cost which is the description of the distributions and concomitant fields of the solvent species of the system in a mean field or average way; the solvent is only included through a continuum dielectric screening factor and ions are primitively described as point charges. The excluded volumes of the solvent and ionic species are thus not accounted for in PB calculations.

Given the limitations in the MD and PB approaches, there is a definite need for a more accurate but less computationally expensive method for studying the effects of the surrounding environment on the conformational stability and thermodynamics of charged

molecular species like DNA. Recent improvements in theory and numerical procedures make the use of integral equation methods a practical alternative.⁶

In this paper we use integral equations to calculate the solvent and ion distributions around atomic-site models of DNA and compare them to the results from MD simulations and PB theory. This comparison confirms IEs as a predictive tool, similar to simulation, for the solvent-ion and ion-ion correlations, which are important to the thermodynamics of DNA pairing.

The current applications of 3D-IE methods to multi-site atomic models, such as the ones in this paper, rely on the use of the interaction site model (ISM) theory. The introduction of ISM theory⁷ and its extension to three-dimensional correlations⁸ has opened the field of theoretical research for solutions to a wider variety of systems. Unlike some theories for molecular liquids which rely on one-center, orientationally dependent functions⁹ and are difficult to solve for strongly anisotropic fluid models, ISM theory allows the correlation functions to be approximately calculated as orientationally averaged functions between the sites of the model which depend only on the distance between sites.

The radial or one-dimensional reference interaction site theory has enjoyed some success in its application and extension to various molecular species¹⁰. The first applications were to uncharged molecular models such as diatomics¹¹, and through renormalization it was later extended to models containing charges¹². The consequence of including electrostatic sites requires the use of methods involving the decomposition of the correlation functions into short-ranged numerically transformable functions and long-ranged analytically transformable functions unless truncations or other ways of modifying the force laws are used. The resulting extended-RISM theory has many good features but it also has some fundamental flaws.¹³ For example, the predicted dielectric constant for all RISM methods extracted from the long-ranged behavior of the correlation functions is calculated to be too small by a factor of almost four for bulk water. This can be corrected for many purposes with a dielectrically consistent reference interaction site model (DRISM) theory¹⁴. Another issue has to do with the amount of information lost in the orientationally averaged correlation functions. For profoundly anisotropic models there is little chance that a clear description of the orientational dependent distribution functions can be obtained from the 1D distributions. This limitation has been partially overcome with the introduction and application of 3D-IEs which allow for correlation information at a distance and an angle from a solute species.

The earliest application of the 3D-IEs was to anisotropic solute molecules solvated by spherical Lennard-Jones particles^{8a} and latter to molecular solvents^{8b, 15}. For solute models containing electrostatic charges, polynomial truncations and periodicity based (so called super-cell) techniques have predominately been used.^{10, 16} The 3D-IEs in conjunction with the super-cell technique have been used to study a variety of multi-site anisotropic models ranging from three-site water molecules to proteins containing hundreds of interaction sites¹⁷. This paper is focused on the use of the more exact long-range treatment of the correlation functions (analogous to the 1D equations¹⁸) to study DNA in salt solutions.

The exact treatment of the long-range part of the correlation functions developed in this lab for the 3D-IEs uses re-summation techniques⁶. The method has proven to increase the numerical stability of the converged solutions compared to other methods and has been applied to proteins in pure water and electrically charged wall solutes in electrolyte solutions using the more formally correct HNC closure and variants including a bridge function^{6, 19}. In this paper we show the utility of the new numerical treatment by applying 3D-RISM/HNC theory to electrostatically charged DNA duplex fragments in aqueous electrolyte

solutions. One of the goals is to understand the balance between inter-DNA strand forces and solvation forces as a function of oligonucleotide length in the short strand limit.

In addition to predicting the liquid structure, the 3D-IEs have been used successfully to analyze the conformational stability of biomolecules through the free energy of the system²⁰. The quality of the distribution functions and the free energies, compared to MD simulations, are strongly dependent on the choice of the closure. Two widely used closures for the 3D-IE systems are the hyper-netted chain (HNC) closure²¹ and the Kovalenko-Hirata (KH) closure¹⁰. The HNC closure gives results in better agreement with simulation for systems with electrostatic sites^{6, 22} compared to many other closures. The KH closure, which is a partially linearized form of the HNC equation, converges quickly but has been shown to underestimate the peak height features in the distributions⁶.

This paper is organized into five sections. Details of the theory and the thermodynamic equations used are presented in Section II. Section III contains the structural details and parameters for our model systems along with the computational methods used in this study. A discussion of our results is given in Section IV and our conclusions are in Section V.

II. THEORY AND EQUATIONS

In order to calculate the 3D distribution functions of the solvent and ion sites 3D-RISM IE theory was utilized⁸. For an infinitely dilute solute density the IE for multi-component fluids can be expressed in Fourier space as,

$$h_{ua}(\mathbf{k}) = \sum_b c_{ub}(\mathbf{k})(w_{ba}(\mathbf{k}) + \rho_b h_{ba}(\mathbf{k})). \quad (2.1)$$

The products in equation 2.1 represent convolution integrals in real space. The solute is fixed on a 3D grid and the solvent and ion correlations are calculated on the grid in the field of the solute. The subscripts in equation 2.1 account for the ion and solvent sites in the model. The 3D total and direct correlation functions for the solvent around the solute are represented by $h_{ua}(\mathbf{k})$ and $c_{ua}(\mathbf{k})$, respectively, $w_{ba}(\mathbf{k})$ is the site-site intra-molecular correlation function which specifies the molecular connectivity of the solvent species and ρ_b is the number density of the solvent and ion sites. The $h_{ba}(\mathbf{k})$ term is the 1D solvent-solvent correlation functions and can be calculated using DRISM theory and transferred to the 3D grid. Here, equation 2.1 is complemented with the hyper-netted chain (HNC) closure equation to form a closed set of equations. The 3D form of the HNC closure is written as,

$$c_{ua}(\mathbf{r}) = \exp(-\beta u_{ua}(\mathbf{r}) + t_{ua}(\mathbf{r})) - t_{ua}(\mathbf{r}) - 1, \quad (2.2)$$

where $u_{ua}(\mathbf{r})$ is the interaction potential between the solute molecule and solvent interaction sites, β is the inverse of the product of the system temperature and Boltzmann's constant, and t_{ua} , where $t_{ua}(\mathbf{r}) = h_{ua}(\mathbf{r}) - c_{ua}(\mathbf{r})$, is the indirect correlation function. In the study by Perkyns et. al.⁶ an additional bridge function was included in the closure equation to converge the equations. In that study the bridge function was necessary due to the strong solute-solvent interactions in the interior hydration sites. In this study, however, the same bridge function was found to be unnecessary since B- DNA does not contain any secluded, interior hydration sites.

The long-range interactions between electrostatic sites are treated with a recently developed method^{6, 19} in which the correlations are divided into two parts: (1) a short-range

numerically transformable part and (2) a long-range analytically transformable part. In our applications we re-sum the functions using the $erf(r)$ function, although other functions can be used.

Statistical mechanical methods of interaction site fluids yield convenient analytical expressions for the thermodynamic quantities of the system. The solvation free energy is readily calculated using the Morita-Hiroike HNC formula²³,

$$\Delta\mu = k_B T \sum_a \rho_a \int [\frac{1}{2} h_{ua}(\mathbf{r})^2 - c_{ua}(\mathbf{r}) - \frac{1}{2} h_{ua}(\mathbf{r}) c_{ua}(\mathbf{r})] d\mathbf{r} \quad (2.3)$$

where the total and direct correlation functions used in equation 2.3 are taken from the converged solutions to equation 2.1. The solute-solvent interaction energy is also easily calculated using,

$$\Delta\epsilon^{uv} = \sum_{\alpha} \rho_{\alpha} \int u_{\alpha}(\mathbf{r}) g_{\alpha}(\mathbf{r}) d\mathbf{r} \quad (2.4)$$

The total change in free energy between two solute states can be calculated using,

$$\Delta G = \Delta\mu + \Delta\epsilon^{uu} \quad (2.5)$$

where $\Delta\epsilon^{uu}$ is the solute-solute potential energy and $\Delta\mu$ is the change in the solvent contributions to the solvation free energy. The free energy, ΔG , is helpful in the analysis of the relative stabilities of a series of system configurations.

For comparison, we also calculated the total charge density of the ionic atmosphere using Poisson-Boltzmann theory,

$$\vec{\nabla} \left[\epsilon(\mathbf{r}) \vec{\nabla} \psi(\mathbf{r}) \right] = -4\pi \rho(\mathbf{r}) - 4\pi \sum_{\alpha} c_{\alpha}^{\infty} z_{\alpha} q \lambda(\mathbf{r}) e^{-\beta z_{\alpha} q \psi(\mathbf{r})}, \quad (2.7)$$

where $\epsilon(\mathbf{r})$ represents the spatially dependent dielectric screening factor, $\psi(\mathbf{r})$ represents the electrostatic potential, $\rho(\mathbf{r})$ represents the charge density of the solute, c_{α}^{∞} represents the bulk ion concentration of ion α , Z_{α} is the ion valence, q is equal to the charge of a proton, β is the inverse of the product of Boltzmann's constant and the temperature, and $\lambda(\mathbf{r})$ accounts for the accessibility of position $\mathbf{r}$ to the ions in solution. For completeness, we compare our IE and PB results with literature results from molecular dynamics calculations.^{3a}

III. MODEL AND METHOD

To show the ability of the IEs to predict features unique to different nucleic acid structures, we calculated the ion and water distributions around different canonical DNA homopolymer duplexes consisting of the dA_n3dT_n and dG_n3dC_n nucleotide sequences for n=1–12. The B-form of DNA is used and the CHARMM27 is adopted for the nucleic acid force field parameters.²⁴ The idealized coordinates of the DNA atomic sites are obtained using the NAB (nucleic acid builder) program²⁵. Thermodynamic results are also presented for

oligonucleotides ranging in length from one to six basepairs. The DNA duplexes are solvated in 1 M KCl, 0.1 M KCl and 0.1 M NaCl electrolyte solutions. The intermolecular site-site pair interactions among the DNA, water, and ion sites are calculated using,

$$u_{ab}(r)=4\epsilon_{ab}\left[\left(\frac{\sigma_{ab}}{r}\right)^{12}-\left(\frac{\sigma_{ab}}{r}\right)^6\right]+\frac{q_aq_b}{r}. \quad (3.1)$$

The subscripts specify the interaction sites of the molecular models and σ , ϵ and q specify the Lennard-Jones (LJ) diameter, well-depth and electrostatic charge, respectively. The SPC/E water model²⁶ is used for the solvent water parameters and the ion parameters are from an earlier simulation study of electrolyte solutions²⁷. The parameters for the water and ion sites are listed in Table 1. The traditional Lorentz-Berthelot combination rules, $\sigma_{ab} = (\sigma_{aa} + \sigma_{bb})/2$ and $\epsilon_{ab} = \sqrt{\epsilon_{aa}\epsilon_{bb}}$, are used to calculate the interaction parameters between different sites. The number density of the bulk water is 0.033314 and the solution temperature is 300 K.

The solvent-solvent and solvent-ion one-dimensional distributions are calculated using the dielectrically consistent RISM theory (DRISM) with a dielectric constant of 78.5. The radial grid for the DRISM solvent calculations consisted of 32,768 points spanning a radial distance of 100 Å ($\Delta r = 0.00305$ Å). The solvent solutions are converged at each point in the correlation functions to a relative residual error of 10^{-12} . The 3D solvent and ion distributions for the n -basepair duplexes are calculated on a 3D grid containing 256^3 points with a spatial representation spanning 120^3 Å³. The calculations involving the 12-basepair duplexes are done on a 512^3 grid spanning 240^3 Å³ to reach a relative error of 10^{-9} in the numerical solutions. A MDIIS routine^{17e, f, 18} with a subspace of 5 vectors is used to solve the set of coupled equations. The 3D Fourier transforms are calculated using the FFTW package²⁸.

IV. RESULTS AND DISCUSSION

First, we make two comparisons to show how well the results from the 3D-IEs compare to other methods. The first comparison is to MD simulations which are expected to be exact, within noise, for the model. This will give a qualitative picture of the IE results. Because there are no differences in the hydration and ion patterns between the hexamer and dodecamer duplexes, only the results of the hexamer are shown. Three salt solutions of 0.1 M NaCl, 0.1 M KCl, and 1 M KCl are used and comparisons of the results for 0.1 M aqueous solutions illustrate the differences between the Na⁺ and K⁺ distributions. For completeness, comparisons to PB theory for the same model, parameters, and grid resolution are performed to illustrate the differences between IE and PB methods.

A. 3D-RISM and MD comparison

Water structures—Figure 1 shows the 3D-IE results of the overall hydration pattern of the water-oxygen sites around the dA₆3dT₆ oligonucleotide in 1 M KCl electrolyte solution. The contour level is set at 2.5 times the bulk water density and the DNA duplex is orientated with a view into the minor groove. The hydration pattern around the dG₆3dC₆ oligonucleotide (not shown) is similar to the one for dA₆3dT₆ (Fig. 1) at the stated contour level. A characteristic feature in Fig. 1 is the hydration patterns in the minor groove and along the backbone, which are uniformly established and stable by the 2nd and $n-1$ basepairs. End effects were also minimal in the MD simulations^{3a, 29}; this will be important in the validation of our conclusions based on observations for the shorter length oligonucleotides. Experimental crystallographic data and simulation long residence times studies have identified these highly coordinated water molecule positions.^{3a} The hydration pattern around

the phosphate groups in Fig. 1 are not base specific and reveal identical “cones of hydration” along the backbones of both the AT and CG duplexes.

Planar averages of the water-oxygen site distributions are shown in Fig. 2. The averaging procedure consisted of averaging the distributions in the first set of planes, in either direction of the plane through the center of the basepair plane, for a range of ± 0.47 Å. The MD results and crystallographic hydration sites over a similar range can be found in earlier work^{3a}. Both MD and IE results show strong specific hydration patterns in the minor and major grooves. The IE results show high probabilities for solvent waters at the same relative positions as the crystallographic and MD positions. The major groove hydration sites in the IE results are more base-sequence dependent compared to the other hydration regions that coincide with the MD studies. More specifically, the CG major groove is characterized by four strong hydration sites near the guanine N7, guanine O6, cytosine N4, and cytosine C8 atoms. The IE results reliably account for the four main hydration sites in the CG major groove. These localized water molecules are an integral part of the structure and an accurate description is important to reliably probe the thermodynamic stability of the DNA model. Along the AT major groove, in the IE results, four hydration sites are found coordinated near the adenine N6 and N7, and the thymine O4 atoms; the fourth site is located between the thymine methyl and phosphate group. This is also observed in the MD results^{3a} despite some motional averaging due to conformational flexibility.

Figure 3 shows iso-surface plots of the hydration structure along the CG and AT grooves with a contour value of 2.0 times the bulk water density. The hydration patterns in the CG and AT minor grooves (Fig. 3A and 3B) are more similar in the IE results than in the MD results which allowed for structural relaxation and fluctuations and are for the AMBER force field³⁰. At the time of publication Amber force field simulations successfully maintained the B-form for DNA and the CHARMM force field did not³¹; that has subsequently been remedied³². The CG minor groove in the MD simulations widened from the canonical B-form allowing for an increased area of hydration of the guanine N2, N3, and cytosine O2 atoms. A narrower hydration pattern is found in the AT duplex minor groove for both the IE and MD results (Figure 3B). The width and shape of the hydration of the AT minor groove compares well between the two methods; the AT minor groove retains its canonical shape and size in the MD simulations.

A more stringent comparison between the IE and the MD results is with regard to the hydration patterns in the major grooves, shown in Figures 3C and 3D. In the CG major groove the water molecules appear to align with the basepair plane, while in the AT major groove the hydration pattern aligns along the direction of the groove. This distinction between the CG and AT major grooves is observed in both the IE and MD results. IE theory is definitely able to distinguish between the different bases and their characteristic hydration structure.

Here we compare the results from the HNC closure to results from the KH closure for the DNA models presented here. The KH closure has recently been used to predict the hydration patterns around DNA, however, in that study no ions were included in the solvent³⁸. For the comparison in this paper the hydration patterns for water molecules, excluding ionic species, were calculated around DNA using both the HNC and KH closure. The results of this calculation are shown in figure 4. The most striking difference between the results of the two closures is the predicted peak heights of the H₂O – oxygen sites in contact with the DNA structure. The peak heights of the results from the KH method are a fraction of the height of the peaks calculated from the HNC method, especially in the minor groove. This difference in peak height has been studied in detail in another recent publication⁶ where it

was found that the KH closure underestimates the contact peak heights in relation to the HNC closure and simulation of the water molecule distribution around the protein, BPTI.

Ion Structures—DNA strands are negatively charged polymers, yet, in electrolyte solutions come together to form stable hybridized duplexes. For this to occur the negative regions, which coincide with the backbone phosphates, must be effectively screened by counter-ions in the solvent. Figure 5 shows the potassium (counter-ion) and chloride (co-ion) ion distributions calculated from the 3D-IEs. The results show a substantial presence of counter-ions around the phosphate groups which is responsible for the thermodynamic stability of the duplex. Some presence of K^+ ions is also found in the major grooves of both CG and AT duplexes; in addition to a considerably smaller but well defined Cl^- ion probability. The ion densities in the major grooves of the duplexes in Figure 5 are base-sequence dependent which is also observed in the MD studies³³. The K^+ ion structure in the AT major groove, which is characterized by a repeating two site density pattern, is due to the strong interaction of the K^+ ions with the carbonyl oxygen on the thymine nucleotide and the N7 amino group on the adenosine nucleotide. The K^+ ion structure in the CG major groove is different and appears as a continuous density pattern which is due to the simultaneous two-fold interaction with the carbonyl oxygen (O6) and the amino (N7) group on the purine ring.

The presence of counter-ion density (K^+ ions) in the minor grooves has also been confirmed from MD simulations and NMR experiments. In addition, the low probability presence of Cl^- ions in the minor and major grooves has also been observed in MD simulations³³. Figure 6 shows the K^+ and Cl^- ion distributions in the minor grooves of the AT and CG duplexes. The IE results reveal two main differences in ion densities between the AT and CG minor grooves. The first is that there is a greater probability of cations in the AT minor groove than the CG and secondly the CG minor groove appears to accommodate a somewhat greater presence of Cl^- ions. These characteristic differences have been also previously reported in simulations from this group.³³ In the simulations longer residence times for the cations were found near the adenine N3 and thymine O2 sites which appear consistent with the IE results, although residence time is not always correlated with occupancy.³⁴ The Cl^- ions were mainly found near the guanine N2 sites and were always found paired with a cation. This same behavior is shown in Fig. 6.

Much of our analysis above has been carried out using the 1 M KCl electrolyte model. A few simulations have focused on the differences in the distributions of Na^+ and K^+ ions around DNA which prove to be useful here³⁵. Although the use of the NaCl model results in a stiffer set of equations we are able to obtain numerical solutions for a 0.1 M NaCl electrolyte solution which we compare to a 0.1 M KCl electrolyte solution to elucidate the effects of counter-ion size.

Figure 7 shows the cation distributions around the AT duplex at three different contour levels corresponding to 2, 10 and 20 cations/nm³. The binding of the cations in the minor groove are similar with only small spatial differences that arise from the different contact diameters of the two ions. When the iso-surface value is increased to 20 cations/nm³ the K^+ ions are still visible in the minor groove but are no longer visible around the phosphate groups. At the same level, however, Na^+ ions are still visible in the minor groove as well as around the phosphate groups. The stronger coordination of the Na^+ ions to the phosphate groups over the K^+ ions is a characteristic feature based on the size and consequent desolvation penalties of the ions which were also observed in the MD simulations³⁵.

We now focus on the calculated thermodynamics of nucleic acids using the 1 M KCl aqueous electrolyte solution. Table 2 shows the solute-solvent interaction energies, the

solvation free energies, and the free energy of binding for DNA oligonucleotides with basepair lengths ranging from the single basepair to the 6 basepair duplex. The change in these quantities for the duplex relative to the two single strands in solution is denoted by the $d \leftarrow s$ subscript. These calculations are for a rigid set of structures; a complete picture of DNA hybridization, especially for longer strands, would require the inclusion of flexibility and thus the entropy of the DNA strands. IEs predict a cross over to a favorable change in free energy for the hybridization of the 3mer and larger structures. 3mers such as codons are often the smallest observed forms of functioning DNA fragments.³⁶ 1mers and 2mers do not hybridize due to the solvation thermodynamics.³⁶

Here we note the trends in the interaction potential and solvation chemical potential as the length of the nucleotide increases. Instead of approaching a constant value as the number of nucleotides increases they continue to increase in value. This is due in part to our use of rigid structures which do not allow for the necessary distortions, *i.e.*, response, during nucleation of hybridization or melting of larger structures.

B. 3D-RISM and PB comparison

Poisson-Boltzmann theory has often been utilized to calculate the electrostatics of DNA and protein models. In PB calculations the total charge density of the ions can be calculated from the divergence of the electric field at each grid point. For simple hydration, the potential field is calculated from the electrostatic sources screened by the dielectric constant of the solvent. The solvent charge density is averaged in continuum models and thus zero everywhere. However, water molecules have dipoles and orient themselves in an electrical field resulting in a non-zero charge density. Figure 8 shows the charge density distribution from the water molecules in a plane through the center of an AT basepair calculated from theory. The theory predicts solvent dipole ordering in the external field of the DNA fragment. Hydrogen bonding or strong localization can be seen in the interior region of Figure 8 depicted by the dark blue regions. PB theory accounts for ordered dipoles only by the mean polarization with no structure in the solvent.

Another important model detail which cannot be ignored is excluded volume effects. PB theory, as already mentioned, only accounts for the solvent through an electrostatic screening factor, but in any aqueous electrolyte solution the solvent molecules occupy the majority of space leaving the ions with reduced access to positions along the solute interface. PB theory does not consider the size or shape of the solvent species except by exclusion of the solute leading to a lack of structure in the PB density distributions, Figure 9. Thermodynamically, the entropic effects due to excluded volume of ions within the solvent are completely neglected.

The charge density distributions calculated for the K^+ and Cl^- ions are compared in Fig. 9 for the PB and IE calculations. In Fig. 9 the concentration of the electrolyte solution is 1 M and the densities are shown in ions per nm^3 . The slice through the center of the AT basepair plane is shown. The most striking difference between the theories is the presence of density waves creating alternating positive and negative regions in the IE results. The PB distributions show only a diffuse cloud of positive average charge surrounding the negatively charged DNA strand. The lack of correlation in the ion densities can have serious repercussions when studying systems where finer details of the interactions between the solvation layers are responsible for the recognition, binding, and solvation thermodynamics of DNA³⁷ or protein/DNA systems.

V. CONCLUSION

We used our renormalized 3D-RISM/HNC theory to calculate the water and ion distributions around B-form homopolymer dA_n3dT_n and dG_n3dC_n nucleotide duplexes. Three different electrolyte solutions consisting of 1 M KCl, 0.1 M KCl and 0.1 M NaCl were used to solvate the DNA duplexes at a system temperature of 300 K. We compared our IE results to MD simulations and PB theory to evaluate how well the IEs predict the solvent and ion distributions around strong field solutes such as DNA.

Comparison of the IE and the MD simulation results showed that a significant amount of the characteristic hydration patterns of the DNA duplex are accurately predicted by the IE theory. The spine of hydration in the minor grooves of the duplexes is found to be similar in structure to the MD simulation results. For the major groove, hydrations patterns show the ability of IEs to distinguish between the different bases and their corresponding hydration patterns. The water network in the CG major groove forms horizontal layers with respect to the helical axis, while the water structure in the AT major groove is parallel to the path of the groove which is the same phenomenon observed in the MD simulations. Overall, most of the hydration sites occur near the amino and carbonyl groups.

Others have used 3D IEs to consider the distribution of water around DNA³⁸, but without salt. In this work the IEs accurately describe the cation and anion distributions in the minor grooves of the AT and CG duplexes when compared to the simulation results. The CG minor groove was found to accommodate both cations and anions in accord with MD simulations. The stronger binding of the Na^+ ions over K^+ ions to the phosphate groups predicted by IEs was also seen before in MD simulations. The thermodynamic quantities of the solute-solvent interaction energy and the solvation free energy were reported for a series of DNA oligonucleotide lengths. Interestingly, these quantities showed the 3-mer duplex to be the smallest length of a DNA duplexes at 300K with any reasonable population.

It was shown that the lack of correlations in PB theory oversimplifies the total charge distributions of the systems. The region around DNA is characterized by strong density waves creating alternating charge regions from the population and orientation of the ions and solvent dipoles. The presence of a small but definite concentration of anions near the negatively charged DNA strand is seen in both the MD simulations and the IE calculations, but no such feature is observed using PB theory.

Nucleic acids are a stringent test for statistical mechanical theories of solvation. Past integral equation methods have been limited to approximating either the correlations³⁹ or the system³⁸. The use of our method of renormalization allows a broader range of applicability of the well known HNC closure approximation, which is more accurate than most other approximations for strongly interacting systems⁶. This is particularly relevant to the case of nucleic acids. In combination with 3D-IEs^{8, 15b} a range of thermodynamic questions can be addressed that have been difficult to approach even qualitatively.

Acknowledgments

This work was supported in part by the National Institutes of Health (GM066813), the Robert A. Welch Foundation (E-1028), and a training fellowship to JJH from the Keck Center for Computational and Structural Biology of the Gulf Coast Consortia (NLM grant No. 5T15LM07093).

References

1. (a) Harmouchi M, Albiser G, Premilat S. Changes of Hydration During Conformational Transitions of DNA. *European Biophysics Journal*. 1990; 19(2):87–92. [PubMed: 2073893] (b) Westhof E.

Water - an Integral-Part of Nucleic-Acid Structure. Annual Review of Biophysics and Biophysical Chemistry. 1988; 17:125–144.

2. (a) Arai S, Chatake T, Ohhara T, Kurihara K, Tanaka I, Suzuki N, Fujimoto Z, Mizuno H, Niimura N. Complicated water orientations in the minor groove of the B-DNA decamer d(CCATTAATGG)2 observed by neutron diffraction measurements. *Nucleic Acids Res.* 2005; 33(9):3017–3024. [PubMed: 15914673] (b) Savage H, Wlodawer A. Determination of water structure around biomolecules using x-ray and neutron diffraction methods. *Methods Enzymol.* FIELD Full Journal Title:Methods in Enzymology. 1986; 127(Biomembranes, Pt. O):162–83.
3. (a) Feig M, Pettitt BM. Modeling High-resolution Hydration Patterns in Correlation with DNA Sequence and Conformation. *J Mol Biol* FIELD Full Journal Title:Journal of Molecular Biology. 1999; 286(4):1075–1095. (b) Norberg J, Nilsson L. Molecular Dynamics Applied to Nucleic Acids. *Accounts of Chemical Research.* 2002; 35(6):465–472. [PubMed: 12069632] (c) Feig M, Pettitt BM. Crystallographic water sites from a theoretical perspective. *Structure (London)* FIELD Full Journal Title:Structure (London). 1998; 6(11):1351–1354.
4. (a) Forester TR, McDonald IR. Molecular-Dynamics Studies of the Behavior of Water-Molecules and Small Ions in Concentrated-Solutions of Polymeric B-DNA. *Mol Phys.* 1991; 72(3):643–660. (b) Laughton CA, Luque FJ, Orozco M. Counterion Distribution around DNA Studied by Molecular-Dynamics and Quantum-Mechanical Simulations. *Journal of Physical Chemistry.* 1995; 99(29):11591–11599. (c) Rau DC, Parsegian VA. Direct Measurement of the Intermolecular Forces between Counterion-Condensed DNA Double Helices - Evidence for Long-Range Attractive Hydration Forces. *Biophysical Journal.* 1992; 61(1):246–259. [PubMed: 1540693]
5. Ponomarev SY, Thayer KM, Beveridge DL. Ion motions in molecular dynamics simulations on DNA. *Proc Natl Acad Sci U S A.* 2004; 101(41):14771–14775. [PubMed: 15465909]
6. Perkyns J, Lynch GG, Howard JJ, Pettitt BM. Protein solvation from theory and simulation: Exact treatment of Coulomb interactions in three-dimensional theories. *J Chem Phys.* 2010; 132(6):064106. [PubMed: 20151732]
7. (a) Chandler D, Andersen HC. Optimized cluster expansions for classical fluids. II. Theory of molecular liquids. *Journal of Chemical Physics.* 1972; 57(5):1930–7. (b) Chandler D. Derivation of an integral equation for pair correlation functions in molecular fluids. *Journal of Chemical Physics.* 1973; 59(5):2742–6.
8. (a) Beglov D, Roux B. Numerical solution of the hypernetted chain equation for a solute of arbitrary geometry in three dimensions. *Journal of Chemical Physics.* 1995; 103(1):360–4. (b) Ikeguchi M, Doi J. Direct numerical solution of the Ornstein-Zernike integral equation and spatial distribution of water around hydrophobic molecules. *J Chem Phys.* 1995; 103(12):5011–5017.
9. (a) Fries PH, Patey GN. The solution of the hypernetted-chain approximation for fluids of nonspherical particles. A general method with application to dipolar hard spheres. *Journal of Chemical Physics.* 1985; 82(1):429–40. (b) Fries PH, Patey GN. The solution of the Percus-Yevick approximation for fluids with angle-dependent pair interactions. A general method with results for dipolar hard spheres. *Journal of Chemical Physics.* 1986; 85(12):7307–11.
10. Hirata, F. *Molecular Theory of Solvation.* Vol. 24. Kluwer Academic Publishers; Dordrecht: 2003.
11. Lowden LJ, Chandler D. Solution of a new integral equation for pair correlation functions in molecular liquids. *Journal of Chemical Physics.* 1973; 59(12):6587–95.
12. (a) Springer JF, Pokrant MA, Stevens FA Jr. Integral equation solutions for the classical electron gas. *Journal of Chemical Physics.* 1973; 58(11):4863–7. (b) Ng K-C. Hypernetted chain solutions for the classical one component plasma up to $G = 7000$. *Journal of Chemical Physics.* 1974; 61(7):2680–9. (c) Hirata F, Rossky PJ. An extended RISM equation for molecular polar fluids. *Chemical Physics Letters.* 1981; 83(2):329–34. (d) Hirata F, Pettitt BM, Rossky PJ. Application of an extended RISM equation to dipolar and quadrupolar fluids. *Journal of Chemical Physics.* 1982; 77(1):509–20.
13. (a) Monson PA, Morriss GP. Recent progress in the statistical mechanics of interaction site fluids. *Adv Chem Phys* FIELD Full Journal Title:Advances in Chemical Physics. 1990; 77:451–550. (b) Chandler D, Joslin CG, Deutch JM. Calculation of the dielectric constant of polyatomic fluids with the interaction site formalism. *Molecular Physics.* 1982; 47(4):871–9.
14. (a) Perkyns J, Pettitt BM. A site-site theory for finite concentration saline solutions. *Journal of Chemical Physics.* 1992; 97(10):7656–7666. (b) Perkyns J, Pettitt BM. A dielectrically consistent

- interaction site theory for solvent-electrolyte mixtures. *Chemical Physics Letters*. 1992; 190(6): 626–630.
15. (a) Beglov D, Roux B. Integral Equation To Describe the Solvation of Polar Molecules in Liquid Water. *Journal of Physical Chemistry B*. 1997; 101(39):7821–7826. (b) Kovalenko A, Hirata F. Three-dimensional density profiles of water in contact with a solute of arbitrary shape: a RISM approach. *Chemical Physics Letters*. 1998; 290(1,2,3):237–244.
16. Beglov D, Roux B. Solvation of complex molecules in a polar liquid: An integral equation theory. *Journal of Chemical Physics*. 1996; 104(21):8678–8689.
17. (a) Imai T, Hiraoka R, Seto T, Kovalenko A, Hirata F. Three-Dimensional Distribution Function Theory for the Prediction of Protein-Ligand Binding Sites and Affinities: Application to the Binding of Noble Gases to Hen Egg-White Lysozyme in Aqueous Solution. *Journal of Physical Chemistry B*. 2007; 111(39):11585–11591. (b) Yoshida N, Imai T, Phongphanphane S, Kovalenko A, Hirata F. Molecular Recognition in Biomolecules Studied by Statistical-Mechanical Integral-Equation Theory of Liquids. *Journal of Physical Chemistry B*. 2009; 113(4):873–886. (c) Imai T, Hiraoka R, Kovalenko A, Hirata F. Locating missing water molecules in protein cavities by the three-dimensional reference interaction site model theory of molecular solvation. *Proteins: Structure, Function, and Bioinformatics*. 2007; 66(4):804–813. (d) Yoshida N, Phongphanphane S, Maruyama Y, Imai T, Hirata F. Selective Ion-Binding by Protein Probed with the 3D-RISM Theory. *Journal of the American Chemical Society*. 2006; 128(37):12042–12043. [PubMed: 16967934] (e) Kovalenko A, Hirata F. Potentials of mean force of simple ions in ambient aqueous solution. I. Three-dimensional reference interaction site model approach. *Journal of Chemical Physics*. 2000; 112(23):10391–10402. (f) Kovalenko A, Ten-No S, Hirata F. 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. *Journal of Computational Chemistry*. 1999; 20(9):928–936.
18. Kovalenko A, Hirata F. Potential of Mean Force between Two Molecular Ions in a Polar Molecular Solvent: A Study by the Three-Dimensional Reference Interaction Site Model. *The Journal of Physical Chemistry B*. 1999; 103(37):7942–7957.
19. Howard JJ, Perkyns JS, Pettitt BM. The Behavior of Ions near a Charged Walls Dependence on Ion Size, Concentration, and Surface Charge. *Journal of Physical Chemistry, B*. 2010; 114(18):6074–6083.
20. (a) Imai T, Harano Y, Kinoshita M, Kovalenko A, Hirata F. Theoretical analysis on changes in thermodynamic quantities upon protein folding: Essential role of hydration. *Journal of Chemical Physics*. 2007; 126(22):225102/1–225102/9. [PubMed: 17581082] (b) Imai T, Harano Y, Kinoshita M, Kovalenko A, Hirata F. A theoretical analysis on hydration thermodynamics of proteins. *Journal of Chemical Physics*. 2006; 125(2):024911/1–024911/7.
21. (a) Baxter RJ. Method of solution of the Percus-Yevick, hypernetted-chain, or similar equations. *Physical Review*. 1967; 154(1):170–4. (b) Watts RO. Hypernetted-chain approximation applied to a modified Lennard-Jones fluid. *Journal of Chemical Physics*. 1969; 50(3):1358–65.
22. Hansen, JP.; McDonald, IR. *Theory of Simple Liquids*. Academic Press; San Diego: 1986.
23. (a) Morita T, Hiroike K. A New Approach to the Theory of Classical Fluids. III. Progress of Theoretical Physics. 1961; 25(4):537–578. (b) Singer SJ, Chandler D. Free energy functions in the extended RISM approximation. *Mol Phys*. 1985; 55(3):621–5.
24. Foloppe N, Mackerell AD. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *Journal of Computational Chemistry*. 2000; 21(2):86–104.
25. Macke, T.; Case, DA. Modeling unusual nucleic acid structures. In: Leontes, NB.; SantaLucia, J., editors. *Molecular Modeling of Nucleic Acids*. American Chemical Society; Washington, DC: 1998.
26. Berendsen HJC, Grigera JR, Straatsma TP. The missing term in effective pair potentials. *Journal of Physical Chemistry*. 1987; 91(24):6269–71.
27. Lynden-Bell RM, Rasaiah JC. From hydrophobic to hydrophilic behaviour: A simulation study of solvation entropy and free energy of simple solutes. *The Journal of Chemical Physics*. 1997; 107(6):1981–1991.

28. Frigo M, Johnson SG. The Design and Implementation of FFTW3. *Proceedings for the IEEE*. 2005; 93(2):216–231.
29. Feig M, Pettitt BM. Experiment vs Force Fields: DNA Conformation from Molecular Dynamics Simulations. *Journal of Physical Chemistry B*. 1997; 101(38):7361–7363.
30. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society*. 1995; 117(19): 5179–5197.
31. Feig M, Pettitt BM. A molecular simulation picture of DNA hydration around A- and B-DNA. *Biopolymers*. 1998; 48(4):199–209. [PubMed: 10699840]
32. Foloppe N, Mackerell AD. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J Comput Chem FIELD Full Journal Title:Journal of Computational Chemistry*. 2000; 21(2):86–104.
33. Feig M, Pettitt BM. Sodium and chlorine ions as part of the DNA solvation shell. *Biophys J*. 1999; 77(4):1769–1781. [PubMed: 10512802]
34. Valère L, Pettitt BM. A Connected-cluster of hydration around myoglobin: Correlation between molecular dynamics simulations and experiment. *Proteins: Structure, Function, and Genetics*. 1994; 18(2):133–147.
35. Cheng Y, Korolev N, Nordenskiöld L. Similarities and differences in interaction of K⁺ and Na⁺ with condensed ordered DNA. A molecular dynamics computer simulation study. *Nucl Acids Res*. 2006; 34(2):686–696. [PubMed: 16449204]
36. Bloomfield, VA.; Crothers, DM.; Ignacio Tinoco, J. *Nucleic acids: structures, properties, and functions*. University Science Books; Sausalito: 2000.
37. Micu AM, Bagheri B, Ilin AV, Scott R, Pettitt BM. Numerical Considerations in the Computation of the Electrostatic Free Energy of Interaction within the Poisson–Boltzmann Theory. *Journal of Computational Physics*. 1997; 136(2):263–271.
38. Yonetani Y, Maruyama Y, Hirata F, Kono H. Comparison of DNA hydration patterns obtained using two distinct computational methods, molecular dynamics simulation and three-dimensional reference interaction site model theory. *Journal of Chemical Physics*. 2008; 128(18):185102/1–185102/9. [PubMed: 18532849]
39. Herman WTVV, Graciela LR, Pettitt BM. A study of model energetics and conformational properties of polynucleotide triplexes. *Biopolymers*. 1990; 30(5–6):517–532. [PubMed: 2265226]
40. Schneider B, Berman HM. Hydration of the DNA Bases Is Local. *Biophysical Journal*. 1995; 69:2661–2669. [PubMed: 8599672]

AT

Figure 1. Oxygen distributions (blue) showing the first hydration shell around the 6 base-pair homopolymers in the 1MKCl aqueous solution. The contour level cutoff is at 2.5 times the bulk water density.

Figure 2.

Volume slices through the CG (top) and AT (bottom) base pairs. The slices are averages of the planes above and below the center plane with a range of ± 0.47 Å. The color scale bar represents the density of water with respect to the bulk water density with a range up to 4 times the water density.

Figure 3.

View of the oxygen distributions showing the hydration patterns in the CG (A) and AT (B) minor grooves, and the CG (C) and AT (D) major grooves. MD results are in the left column and IE results are in the right column. The contours show water densities at 2.0 times the bulk water density. The gold spheres are the experimental results⁴⁰ for the so-called consensus waters.

HNC **KH**

Figure 4. H_2O -oxygen site distributions in the plane through the center of the AT duplex. The top figure shows the oxygen site distributions calculated using the HNC closure and the bottom figure shows the oxygen site distributions calculated for the same model using the KH closure.

AT

CG

Figure 5. Potassium (red) and chlorine (green) distributions around AT (top) and CG (bottom) homopolymers. The concentration of the ions is 1M and the surface plots are shown at a contour level of 3 ions per nm³ or 5 times the bulk density.

AT

CG

Figure 6. Potassium (red) and chlorine (green) distributions in the minor grooves of the AT (top) and CG (bottom) homopolymers in a 1M KCl electrolyte solution. The isosurface plots are shown at an ion count of 6 per nm³

Figure 7.

K^+ (red) and Na^+ (green) ion distributions around the AT duplex in the 0.1M aqueous electrolyte solutions. (Top) Shows the distributions for the two cations super imposed on the same AT duplex with the contour level set at 2 cations/nm³. (Middle) Shows the distributions with a contour level of 10 cations/nm³ and (bottom) shows the cation distributions at a contour level of 20 cations/nm³.

Figure 8.

Water contribution to the charge density from the sum of the water oxygen and hydrogen sites charge densities around an AT base-pair. Red refers to negative charge densities and blue refers to positive charge densities.

Figure 9.

The total charge densities calculated from Poisson-Boltzmann theory (bottom) and integral equation theory (top). The AT base-pair was used in this figure and total charge densities (water and ions) are shown for positive (blue) and negative (red) regions.

Table 1Water ²⁶ and ion parameters ²⁷

site	σ (Å)	ϵ (kJ/mol)	q (e)
O	3.166	0.6505	-0.8476
H	0.400	0.1926	0.4238
Na	2.583	0.5216	1.0
K	3.331	0.5216	1.0
Cl	4.401	0.5216	-1.0

Table 2

Thermodynamic quantities for different length AT duplex (kJ/mol)

n	$\Delta\mathcal{E}_{duplex}^{\mu\nu}$	$\Delta\Delta\mathcal{E}_{d\leftarrow s}^{\mu\nu}$	$\Delta\mu_{duplex}$	$\Delta\psi_{d\leftarrow s}$	$\Delta G_{d\leftarrow s}$
1	-595.95	102.08	218.71	47.97	46.39
2	-2264.29	102.13	-97.65	41.79	25.83
3	-4629.50	-42.50	-758.15	-31.40	-62.78
4	-7352.91	-355.50	-1597.29	-188.00	-233.88
5	-10366.96	-836.51	-2578.61	-427.34	-486.48
6	-13586.70	-1454.48	-3664.66	-736.04	-805.19