

## Article

# Determining the Locations of Ions and Water around DNA from X-Ray Scattering Measurements

Steve P. Meisburger,<sup>1</sup> Suzette A. Pabit,<sup>1</sup> and Lois Pollack<sup>1,\*</sup><sup>1</sup>School of Applied and Engineering Physics, Cornell University, Ithaca, New York

**ABSTRACT** Nucleic acids carry a negative charge, attracting salt ions and water. Interactions with these components of the solvent drive DNA to condense, RNA to fold, and proteins to bind. To understand these biological processes, knowledge of solvent structure around the nucleic acids is critical. Yet, because they are often disordered, ions and water evade detection by x-ray crystallography and other high-resolution methods. Small-angle x-ray scattering (SAXS) is uniquely sensitive to the spatial correlations between solutes and the surrounding solvent. Thus, SAXS provides an experimental constraint to guide or test emerging solvation theories. However, the interpretation of SAXS profiles is nontrivial because of the difficulty in separating the scattering signals of each component: the macromolecule, ions, and hydration water. Here, we demonstrate methods for robustly deconvoluting these signals, facilitating a more straightforward comparison with theory. Using SAXS data collected on an absolute intensity scale for short DNA duplexes in solution with Na<sup>+</sup>, K<sup>+</sup>, Rb<sup>+</sup>, or Cs<sup>+</sup> counterions, we mathematically decompose the scattering profiles into components (DNA, water, and ions) and validate the decomposition using anomalous scattering measurements. In addition, we generate a library of physically motivated ion atmosphere models and rank them by agreement with the scattering data. The best-fit models have relatively compact ion atmospheres when compared to predictions from the mean-field Poisson-Boltzmann theory of electrostatics. Thus, the x-ray scattering methods presented here provide a valuable measurement of the global structure of the ion atmosphere that can be used to test electrostatics theories that go beyond the mean-field approximation.

## INTRODUCTION

Charged biomolecules interact with the surrounding environment. In aqueous solutions, these highly charged molecules attract an atmosphere of ions and water that is responsible for screening electrostatic interactions between molecules. One striking example of the consequences of electrostatic screening is the salt-dependent collective behavior of nucleic acids: DNA molecules strongly repel each other in low-salt solutions (1) but form condensed phases when certain multivalent cations are added (2–4). In biological systems, ion atmospheres are necessarily altered as nucleic acids carry out their varied roles in the cell: processing the genetic code, regulating gene expression, and assembling complexes with protein and small-molecule partners. In these cases, changes in the ion atmosphere may contribute significantly to the total free-energy difference that drives these biologically important structural transitions. Beyond its biological importance, the ion atmosphere is also integral to DNA-based nanotechnologies. Applications such as the electronic detection of specific DNA sequences in solution (5–8) and electronic control of hybridization (9) rely on devices that use an electric field to influence the ion atmospheres around surface-tethered molecules. Knowledge of the precise arrangement

of ions and water around DNA will facilitate the efficient design and operation of such biologically inspired devices.

Efforts to model the solvent distribution around charged biomolecules are not new: many theoretical perspectives ranging in scope from mean-field models to all-atom simulations with explicit solvent have been applied. The former approach is commonly used by experimental scientists who wish to gain insight into solvation effects with minimal computational expense, and it includes the Poisson-Boltzmann (PB) theory, which provides a mean-field treatment of the ion atmosphere in a uniform dielectric medium. The application of this theory to macromolecular systems leads to a partial differential equation, the nonlinear PB equation (NPBE), whose solution prescribes the average counterion density surrounding the macromolecule (10). It is important to note that NPBE-based approaches accurately predict the colligative properties of DNA. Currently, all-atom grid-based solvers provide exact solutions to the NPBE and are widely used for macromolecular electrostatics calculations (11–13).

Despite the appeal of this straightforward-to-apply method, PB has several simplifying assumptions that break down in the modeling of highly charged molecules such as DNA (10,14–17): ions are modeled as point charges, ion-ion interactions are neglected, and the solvent is approximated as a uniform dielectric medium. More sophisticated models have emerged to address some or all of these concerns.

Submitted December 10, 2014, and accepted for publication May 11, 2015.

\*Correspondence: [lp26@cornell.edu](mailto:lp26@cornell.edu)

Editor: James Cole.

© 2015 by the Biophysical Society  
0006-3495/15/06/2886/10 \$2.00

<http://dx.doi.org/10.1016/j.bpj.2015.05.006>



Notable examples include the tightly bound ion model (18), molecular dynamics (MD) (17), and the three-dimensional reference interaction site model (19). These more sophisticated models differ significantly from PB in their predictions for the spatial arrangement of the ions, as exemplified by comparisons between PB and MD (20). Here, by careful measurement and NPBE-based modeling, we highlight a key difference between PB models and measurement that is consistent with the breakdown of PB near the DNA surface.

Our structural studies rely on small-angle x-ray scattering (SAXS), a technique uniquely sensitive to the spatial distribution of all statistically associated ions (15,21,22). X-ray scattering signals contain contributions from the ions, macromolecule, and hydration water. To separate the scattering contribution from the ions, several studies have taken advantage of anomalous (or resonant) scattering (15,23), where the strength of scattering from only one component (e.g., the ions) is varied by tuning the x-ray energy near an absorption edge. Scattering profiles are typically measured at two energies near the edge, and the difference signal reports the spatial correlation between the ions and the DNA. This method, called anomalous SAXS (ASAXS) has been particularly useful for comparison with theories describing ion distributions around duplexes (15,20,24).

Using absolute-intensity calibrations, we recently extended ASAXS methods to count the total number of statistically associated ions around RNA and DNA (25). Although ion counting has been demonstrated using the Donnan effect in equilibrium dialysis experiments (26,27) or with ion-binding fluorescent dyes (28), ASAXS is the only method that provides both the number and distribution of ions around nucleic acids, which is essential for the most accurate comparison to theory. However, the achievable signal/noise ratio in ASAXS measurements is limited by the small difference signals (typically <10%) and sensitivity of macromolecule solutions to radiation damage (29). Furthermore, both the macromolecule and its hydration layer contribute to the anomalous difference signal; these cross-terms must be taken into account when comparing with theory.

Here, we describe a straightforward extension of the absolute-intensity ASAXS method in which heavy atom isomorphous replacement is used to vary the contrast (scattering strength) of the ion atmosphere. Relative to anomalous scattering, changes in ion contrast by heavy-atom substitution can be large, providing a greater signal/noise ratio and more precise information about the locations of ions around the nucleic acid. Unlike previous heavy-atom replacement studies (15,21,22,30,31), we decompose the scattering data into its three components (DNA alone, ions alone, and the DNA-ion cross-term) to extract greater detail about the ion environment. A comparison of absolutely calibrated SAXS profiles to predictions based on the NPBE suggests that standard PB theory underestimates the number of

ions present near the DNA surface. Finally, these absolute-intensity measurements provide the number of waters in close proximity to the DNA. These data should be valuable for testing new computational methods that build on PB theory by including additional solvent degrees of freedom (19,32).

## MATERIALS AND METHODS

### Sample preparation

DNA samples were prepared as described in the [Supporting Material](#). Briefly, a mixed-sequence, 25 basepair DNA duplex (charge  $-48e$ ) was formed from two synthetic oligonucleotides by annealing and buffer-exchanged in four different solutions containing 100 mM of 1:1 salt (NaCl, KCl, RbCl, and CsCl) and 1 mM Na-MOPS, pH 7.0. The DNA concentration was determined by ultraviolet absorbance at 260 nm, and stocks were diluted using the matching buffer to a concentration of 50  $\mu$ M for SAXS and 150  $\mu$ M for ASAXS.

### SAXS data collection

SAXS data were collected at the Cornell High Energy Synchrotron Source (CHESS) beamline G1, which has a 49-pole wiggler source and multilayer optics. The monochromator was tuned to an x-ray energy of 10.53 keV, and the beam size was defined using slit dimensions 300  $\mu$ m(H)  $\times$  250  $\mu$ m(V). Parasitic scattering from the beam-defining slits was blocked using a set of guard slits, and the entire x-ray path between the beam-defining slits and the detector was evacuated ( $\sim 10$  mTorr) to reduce the background. The photocurrent from a PIN diode integrated into the beamstop was used to correct for variations in the x-ray intensity over time. Samples were held in an in-vacuum quartz capillary with a 2 mm diameter and 10  $\mu$ m wall thickness (Hampton Research, Aliso Viejo, CA). The sample was oscillated through the beam using a computer-controlled syringe pump to mitigate radiation damage (33). Scattering patterns were recorded with a photon-counting area detector (Pilatus 100K, Dectris, Baden, Switzerland) at a distance of 1.689 m from the sample position in four 60 s exposures.

### ASAXS data collection

ASAXS experiments were performed at the CHESS C1 beamline using bend-magnet radiation and a double-bounce Si(110) monochromator. The sample was held in a 3-mm-diameter quartz capillary with 10  $\mu$ m wall thickness and oscillated during x-ray exposure. An evacuated flight tube with a Si<sub>3</sub>N<sub>4</sub> window on the upstream end was positioned between the sample capillary and detector. Images were recorded on the area detector (Pilatus 100K, Dectris) at a distance of 0.957 m from the sample. A semi-transparent beamstop consisting of a stack of Mo foils (Goodfellow Corp., Coraopolis, PA) with a 550  $\mu$ m nominal thickness was mounted inside the flight tube. The attenuated beam was recorded on the detector and its centroid was determined at each x-ray energy, defining  $q = 0$  for azimuthal integration. X-rays scattered elastically by the beamstop were intercepted by a Si drift detector (X-flash, Roentec, Berlin, Germany) and counts were integrated during each exposure for normalization. Scattering data were acquired at two energies below the Rb<sup>+</sup> K-edge, 15.093 keV (lo) and 15.191 keV (on). The monochromator cycled between the energies every 2 min, and the total exposure time per sample was 32 min.

### X-ray data processing

SAXS and ASAXS data were processed using in-house routines written in MATLAB (The MathWorks, Natick, MA). The detector coordinates were

converted to momentum transfer ( $q = 4\pi\sin(2\theta/2)/\lambda$ , where  $2\theta$  is the scattering angle and  $\lambda$  the wavelength) using powder diffraction from silver behenate for calibration purposes (34). Pixels occluded by the beamstop were masked, and the one-dimensional intensity was computed from each two-dimensional image by binning pixels according to the magnitude of  $q$ , with constant bin widths of  $0.002 \text{ \AA}^{-1}$  (SAXS) and  $0.004 \text{ \AA}^{-1}$  (ASAXS). The uncertainty of  $I(q_i)$  was estimated using  $\Delta I(q_i) \approx \sigma_i N_i^{-1/2}$ , where  $\sigma_i$  is the standard deviation among the  $N_i$  pixels in each bin. This uncertainty was propagated through subsequent averaging and buffer subtraction operations. The data were checked for radiation damage effects by comparing consecutive exposures before averaging. Finally, the scattering intensity was placed on an absolute scale using liquid water as a calibrant, as described in the Supporting Material, Pabit et al. (25), and Orthaber et al. (35). The scattering data were regularized and extrapolated to  $q = 0$  using a Bayesian indirect Fourier transform (36) implemented in MATLAB. At the low [DNA] employed, we do not anticipate significant modifications to the  $q = 0$  value resulting from interparticle interference (37,38). The signal would change by  $<3\%$  under these experimental conditions (Fig. S1). Our results are robust when the scattering profiles are truncated below  $q = 0.05 \text{ \AA}^{-1}$  (Fig. S3), further allaying concerns about the effect of interparticle interactions.

## RESULTS

### X-ray scattering from DNA with group-I cations

Duplex DNA (25 basepair, formal charge of  $-48 e$ ) was equilibrated with neutral pH buffers containing 100 mM chloride salts with group-I monovalent cations NaCl, KCl, RbCl, and CsCl. The activity coefficients of these salts differ by  $<4\%$  at 100 mM (39). SAXS profiles were acquired from solution with a low DNA concentration to minimize interparticle interference effects (37,38). The scattering profiles for each sample were normalized by the DNA concentration and placed on an absolute, calibrated intensity scale of electrons<sup>2</sup> per molecule. Repeated concentration measurements gave a normalization precision of  $\pm 3\%$  (SD). Further details for sample preparation and data analysis are provided in Materials and Methods and in the Supporting Material.

SAXS profiles for the cation series (Fig. 1) show enhancement at low scattering angles that increases with cation atomic number, suggesting the presence of excess cations associated with the DNA molecules. The simplest model for this ion-species-dependent effect is that of a two-phase system consisting of 1) a cloud of  $N_I$  excess ions, and 2) the hydrated macromolecule. Then, the forward scattering intensity  $I(q \rightarrow 0)$  is expected to vary with the ion contrast,  $\delta_I$ , and macromolecular contrast,  $\delta_M$ , according to

$$I(q \rightarrow 0) = (\delta_M + N_I \delta_I)^2. \quad (1)$$

If  $\delta_M$  and  $N_I$  are the same in all salt solutions, then a plot of  $I(q \rightarrow 0)^{1/2}$  versus  $\delta_I$  should yield a straight line. Assuming that the hydration of each ion (i.e., its interaction with water) is unaffected by the DNA,  $\delta_I$  can be estimated from the available data on dilute salt solutions. Ions exclude and attract water molecules via steric effects and electrostriction. At low angles, the hydrated ion scatters as though it

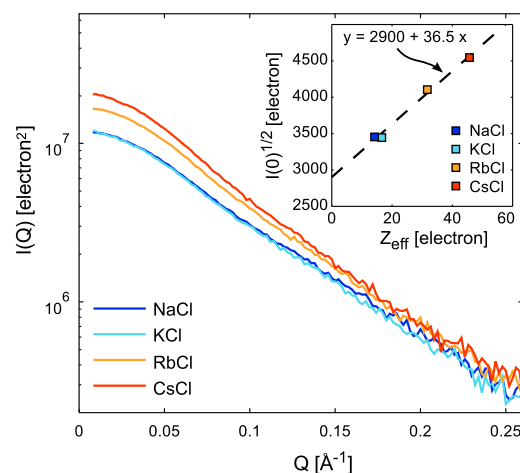


FIGURE 1 Small-angle x-ray scattering from 25 basepair DNA in 100 mM chloride salts. The background-subtracted scattering intensity,  $I(q)$ , versus momentum transfer,  $q$ , is shown in semilog scale with absolute units (electrons<sup>2</sup> per DNA molecule). The intensity at low angles depends on the cation type, indicated in the legend. (Inset) The extrapolated forward-scattering amplitude,  $I(q \rightarrow 0)^{1/2}$ , varies linearly with the effective ion contrast,  $Z_{\text{eff}}$  (see Eqs. 1 and 2).

had an effective number of electrons,  $Z_{\text{eff}}$ , that is different from the actual number,  $Z_I$ , according to

$$Z_{\text{eff}} = Z_I - \bar{\rho}_e V_I, \quad (2)$$

where  $\bar{\rho}_e$  is the average electron density of the solvent (for water at room temperature,  $\bar{\rho}_e = 0.334 \text{ \AA}^{-3}$ ) and  $V_I$  is the change in solution volume due to the addition of a single ion. The latter quantity,  $V_I$ , is related to the absolute limiting partial molar volume of the ion, which has been measured for common salts (40). Relevant values for  $Z_{\text{eff}}$ ,  $Z_I$ , and  $V_I$  are provided in Table 1.

A plot of  $I(q \rightarrow 0)^{1/2}$  versus  $Z_{\text{eff}}$  (Fig. 1, inset) is well fit by a straight line, conforming to expectations for the two-phase model,  $I(q \rightarrow 0)^{1/2} = \delta_M + N_I Z_{\text{eff}}$ . From the slope, a number of excess ions,  $N_I = 36.5 \pm 2.4$ , is obtained. Since the DNA charge is  $-48e$  at neutral pH, the measured value of  $N_I$  indicates that 75 – 80% of the charge is neutralized by monovalent counterions, in accord with previous equilibrium dialysis measurements that found  $80 \pm 4\%$  neutralization for a DNA duplex of similar length and composition in 100 mM NaCl (14). The remaining DNA

TABLE 1 X-ray scattering contrasts and sizes of ions in water

	$Z_I$	$V_I (\text{\AA}^3)^a$	$Z_{\text{eff}}^b$	$R_{M-O} (\text{\AA})^c$
$\text{Cl}^-$	18	40.2	4.6	–
$\text{Na}^+$	10	–12.6	14.2	2.3
$\text{K}^+$	18	4.3	16.6	2.8
$\text{Rb}^+$	36	12.8	31.7	2.9
$\text{Cs}^+$	54	24.7	45.8	3.1

<sup>a</sup>The absolute limiting partial volume per ion ( $\pm 1 \text{ \AA}^3$ ), from Marcus (40).

<sup>b</sup>Calculated from Eq. 2 ( $\pm 0.4e^-$ )

<sup>c</sup>Distance between the ion and the first hydration shell, from Marcus (41).

charge is neutralized by coion exclusion (14,42). Note that the low x-ray contrast of hydrated  $\text{Cl}^-$  (Table 1) makes the heavy-atom measurement relatively insensitive to the coions: the exclusion of  $\sim 11$  Cl ions has the expected effect of decreasing  $I(0)^{1/2}$  by  $\sim 50$  electrons, which is within the uncertainty of the measurement.

By linear extrapolation of  $I(q \rightarrow 0)^{1/2}$  to the point where the net ion contrast is zero, we obtain the DNA contrast according to Eq. 1,  $\delta_M = 2900 \pm 50$ . This contrast depends on the number of water molecules attracted to the DNA surface (the hydration waters,  $N_H$ ), and the number excluded ( $N_E$ ). Since each water molecule contributes 10 electrons,

$$\delta_M = Z_{\text{DNA}} + 10(N_H - N_E), \quad (3)$$

where  $Z_{\text{DNA}}$  is the number of electrons bound to the DNA, known from the chemical formula (our DNA construct has 7950 electrons assuming a charge of  $-48$ , or 8018 after applying the anomalous corrections at 10.5 keV). Thus, from the measured value of  $\delta_M$ , we obtain  $N_H - N_E$  directly. Furthermore, if either  $N_H$  or  $N_E$  can be estimated (for example, by geometrical arguments), the other is constrained by the measurement. We exploit this fact in the next section when modeling DNA hydration.

### Mathematical separation of the partial scattering intensities

The full scattering profile corresponding to the two-phase system of Eq. 1 is

$$I(q) = \delta_M^2 P_M(q) + 2\delta_M(\delta_I N_I) P_{MI}(q) + (\delta_I N_I)^2 P_I(q), \quad (4)$$

where  $P_M(q)$  and  $P_I(q)$  are the partial scattering form factors of the macromolecule and associated counterions, and  $P_{MI}(q)$  is the cross-term ( $P_M(0) = P_I(0) = P_{MI}(0) = 1$ , as described in the Supporting Material). From the  $I(q)$  profiles for the monovalent ion series (Fig. 1) and the known values of  $\delta_M$ ,  $\delta_I$ , and  $N_I$ , Eq. 4 was solved for  $P_M(q)$ ,  $P_I(q)$ , and  $P_{MI}(q)$  as described in the Supporting Material. The results are shown in Figs. 2 and 3.

The first term in Eq. 4 is the partial scattering intensity of the DNA,  $I_M(q) = \delta_M^2 P_M(q)$ , conceptually equal to the total scattering if the ion atmosphere were completely invisible ( $\delta_I = 0$ ). Theoretically,  $I_M(q)$  depends only on the scattering amplitudes of the macromolecule, its hydration layer, and the solvent excluded by the DNA:

$$I_M(q) = \langle |A_M(q)|^2 \rangle_{\Omega, \Pi} = \langle |A_{\text{DNA}}(q) + A_H(q) - \langle A_E(q) \rangle_{\Pi}|^2 \rangle_{\Omega, \Pi}. \quad (5)$$

The complex scattering amplitudes  $A_{\text{DNA}}(q)$ ,  $A_H(q)$ , and  $A_E(q)$  are equal to Fourier transforms of the electron density of the DNA, the hydration waters, and the excluded solvent, respectively (illustrated in Fig. 2 b). Brackets denote the average over DNA orientation ( $\Omega$ ) and solvent degrees of

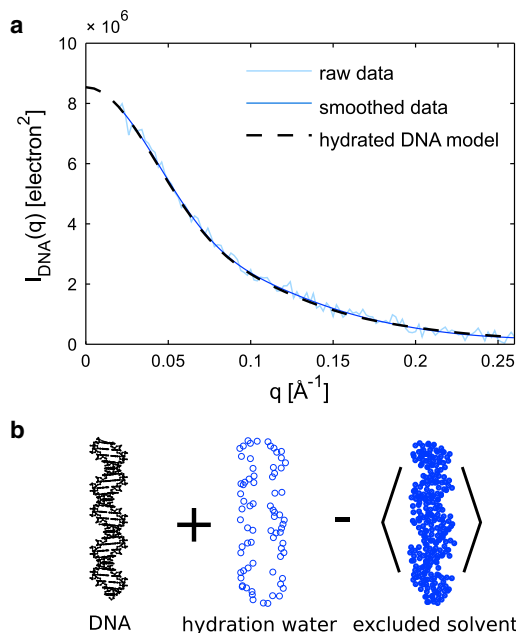


FIGURE 2 (a) The ion-independent (DNA-only) component of the scattering was found from decomposition of the scattering profiles in Fig. 1 (raw data) and from decomposition of the same profiles with prior smoothing by regularization (smoothed data). (b) The theoretical scattering intensity for hydrated DNA shown in (a) is the absolute square of the amplitude components illustrated here. Details of this calculation are provided in the text. For clarity, only water molecules within a 5 Å slab centered on the DNA are shown. To see this figure in color, go online.

freedom ( $\Pi$ ). To predict  $I_M(q)$ , we modeled the DNA as an ideal B-form duplex, added a uniform layer of hydration waters on the surface of the DNA, and chose the excluded solvent volume so that the number of excluded waters,  $N_E$ , reproduced the experimental value of  $\delta_M$  according to Eq. 3. Finally, the scattering profiles were calculated from the models using Eq. 5. Details of the modeling and calculations are included in the Supporting Material. We point out that commonly used software such as CRY SOL (43) fits hydration parameters to the shape of the experimental scattering curve on an arbitrary intensity scale, whereas here, the intensity is measured on an absolute scale, so fitting is not required. In Fig. 2 a, the experimentally determined DNA component is compared with the theoretical prediction for hydrated DNA. Visually, the agreement is excellent. For a quantitative comparison, we compute a reduced  $\chi^2$  for experiment and prediction (Eq. S42) and find that the two agree within the uncertainty ( $\chi_{\text{rel}}^2 = 1.14$ ). Thus, the hydrated DNA component obtained by mathematical decomposition of the scattering data agrees with a simple, geometric hydration model.

### Validation of partial scattering intensities by ASAXS

The three-term decomposition of the SAXS data to  $P_M(q)$ ,  $P_I(q)$ , and  $P_{MI}(q)$  assumes that the ion distributions are the

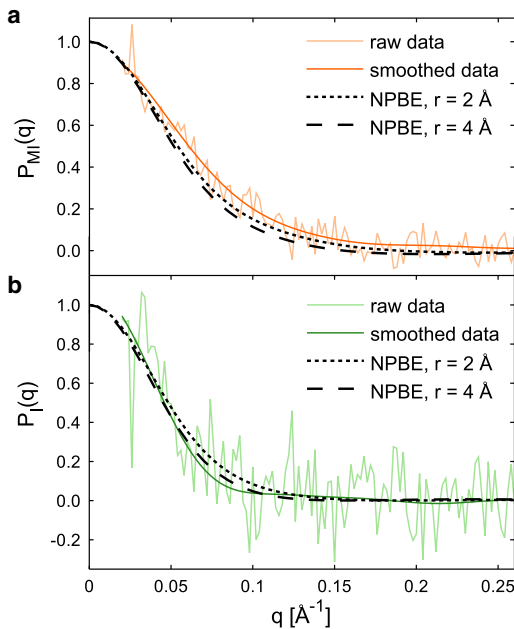


FIGURE 3 Ion-dependent components of the scattering from decomposition of the isomorphous replacement data (Fig. 1). The DNA-ion cross term,  $P_{MI}(q)$  (a), and the ion-only term,  $P_I(q)$  (b), were obtained in two ways: by direct decomposition (raw data) and by decomposition with prior smoothing by regularization (36) (smoothed data). Also shown are predictions derived from the NPBE with Stern layer thicknesses,  $r$ , indicated in the legend. To see this figure in color, go online.

same for all salts employed. To validate this assumption, we performed anomalous (or resonant) scattering measurements (ASAXS) of the DNA sample in 100 mM RbCl (15,23). In ASAXS, the strength of scattering from the Rb atoms is varied by tuning the x-ray energy near its absorption edge; this has an effect similar to that of isomorphous replacement but only one sample is required (see Materials and Methods). The number of Rb<sup>+</sup> ions was obtained from the ASAXS data as described in Pabit et al. (25). We find  $N_I = 39 \pm 3$ , in good agreement with the value of  $N_I$  from isomorphous replacement.

For the two-phase system, the anomalous difference signal,  $I_{lo}(q) - I_{on}(q)$ , is a mixture of the cross-term and the ion-only terms (Eq. S28):

$$I_{lo}(q) - I_{on}(q) = 2\delta_M N_I (f'_{lo} - f'_{on}) (P_{MI}(q) + \gamma P_I(q)), \quad (6)$$

where  $\gamma = N_I(Z_{\text{eff}} + 0.5(f'_{lo} + f'_{on}))\delta_M^{-1}$  and  $f'_{lo}$  and  $f'_{on}$  are the real parts of the anomalous scattering factor for Rb<sup>+</sup> at  $E_{lo}$  and  $E_{on}$  (for numerical values of  $f'$ , see Fig. 4 a). To provide a cross-check for the basis functions determined by the isomorphous replacement method, we computed this difference using the experimentally determined partial form factors  $P_{MI}(q)$  and  $P_I(q)$  shown in Fig. 3 for comparison with the ASAXS data. The ASAXS and isomorphous replacement methods agree (Fig. 4,  $\chi_{\text{rel}}^2 = 1.39$ ), giving further support for our interpretation of the isomorphous ion replacement data.

## Prediction of ion scattering terms by PB theory and comparison with experiment

Information about the shape of the ion atmosphere is contained within the partial form factors  $P_{MI}(q)$  and  $P_I(q)$  (Fig. 3, a and b, respectively). We solved the NPBE around the DNA to model the ion atmosphere and calculated  $P_{MI}(q)$  and  $P_I(q)$  as described in the Supporting Material. The only undetermined parameter in this electrostatics model is the Stern layer thickness,  $r$ , nominally equal to the radius of the hydrated counterion. Radii for hydrated ions considered here range from 3.8 to 4.5 Å ( $R_{M-O}$  plus 1.4 Å for the water molecule radius; see Table 1 and Bai et al. (14)). However, previous DNA studies using the NPBE have assumed  $r = 2$  Å (16). Therefore, we performed calculations using both  $r = 2$  Å and  $r = 4$  Å.

As expected, the shape of the predicted cross-term,  $P_{MI}(q)$  (Fig. 3 a), depends on the value of  $r$ . Comparing data and calculations,  $r = 2$  Å shows a slight improvement over  $r = 4$  Å ( $\chi_{\text{rel}}^2 = 1.28$  for  $r = 2$  Å vs.  $\chi_{\text{rel}}^2 = 1.92$  for  $r = 4$  Å). Stern-layer-dependent differences are also observed in the calculated ion-only term,  $P_I(q)$  (Fig. 3 b), but the experimental  $P_I(q)$  is too noisy to discriminate between models ( $\chi_{\text{rel}}^2 = 1.01$  for  $r = 2$  Å and  $\chi_{\text{rel}}^2 = 1.02$  for  $r = 4$  Å). Finally, from the predictions for  $P_I(q)$  and  $P_{MI}(q)$ , we calculated the anomalous difference signal using Eq. 6. Comparing the ASAXS data and the NPBE-based prediction (Fig. 4 c), we find that the agreement is poor for  $r = 4$  Å ( $\chi_{\text{rel}}^2 = 15.4$ ) and improved for  $r = 2$  Å ( $\chi_{\text{rel}}^2 = 1.65$ ).

Theoretically, the cross-term  $P_{MI}(q)$  is the Fourier transform of a distance histogram between electron density belonging to the macromolecule phase and that belonging to the ion atmosphere phase. Its width in reciprocal space is inversely related to the distances between DNA and counterions. Regardless of which value of the Stern radius is assumed,  $P_{MI}(q)$  computed from the NPBE ion distribution decays more rapidly than the experimental curves. Therefore, compared to experimental data, the ions as predicted by NPBE appear to be farther away from the DNA on average. A similar conclusion can be drawn from the ASAXS data, where the predicted anomalous difference signal (a linear combination of  $P_{MI}(q)$  and  $P_I(q)$ ) decays more rapidly than the experimental data. These observations point to a single conclusion: ion locations according to the NPBE are too far away from the DNA molecule on average.

## SAXS data fit using an explicit ion/PB hybrid model

One hypothesis, consistent with observations above, is that PB theory underestimates the number of counterions present at the DNA surface. To further explore this idea, we tested whether alternate ion atmosphere models could provide improved agreement with the data. For this purpose, we adopt a hybrid explicit ion/PB modeling approach motivated

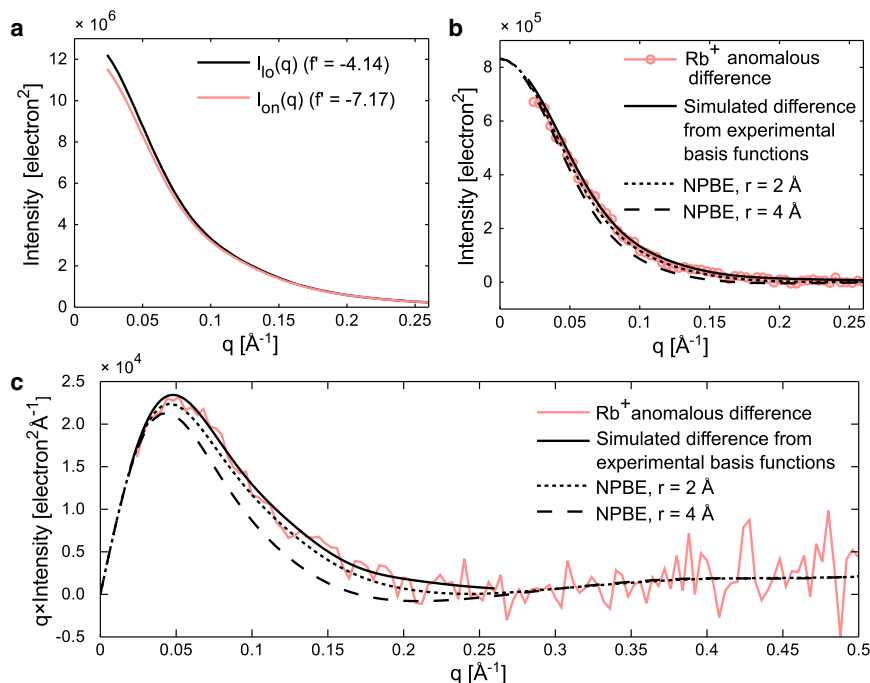


FIGURE 4 Anomalous x-ray scattering (ASAXS) data for DNA with Rb<sup>+</sup> counterions. (a) Scattering profiles were measured on an absolute scale at two energies below the absorption edge of Rb. The real part of the anomalous scattering factor of Rb<sup>+</sup> is given in the legend at each energy (labeled  $I_0$  and  $I_{on}$ ). (b) The anomalous difference between curves in (a),  $I_0(q) - I_{on}(q)$ , is compared with a curve synthesized from the experimental basis functions in Fig. 3 using Eq. 6, and with predictions from PB theory. (c) The profiles in (b) are plotted on  $q \times I(q)$  versus  $q$  axes to facilitate visual comparison. To see this figure in color, go online.

by the tightly bound ion (TBI) model (44). In the TBI model, ions at the DNA surface are modeled explicitly using statistical mechanics, whereas those farther away are modeled using the NPBE. The number of localized ions in TBI depends on an interaction parameter. Since our goal is not to test TBI, but to generate a variety of physically motivated ion distributions, we modified the TBI approach in two important ways. First, we neglect the energetics of the tight-ion-DNA interaction and instead treat the number of localized ions as a fitting parameter. Second, we model a single low-energy distribution for each number of bound ions rather than the full partition function for all binding modes. Starting with a PB model, the procedure for placing counterions was as follows. First, an ion was placed at the site with the greatest magnitude of the electrostatic potential. Then, the NPBE was re-solved with this ion present, and the new potential guided the placement of a second ion, etc. The explicit ion distributions that result from applying this method show a preference for the minor groove (Fig. 5 a). As ions are added to the surface, the DNA charge is gradually neutralized, and the number of ions modeled by the NPBE,  $N_1^{(D)}$ , decreases. Here, D stands for diffuse ions and S for surface-localized ions. Fig. 5 b shows the cumulative radial ion distribution corresponding to each model in Fig. 5 a. The total number of excess ions,  $N_1^{(S+D)}$ , depends weakly on the number placed at the surface,  $N_1^{(S)}$ . Within 50 Å from the central axis of the DNA helix,  $N_1^{(S+D)}$  ranges from 36.7 ( $N_1^{(S)} = 0$ ) to 39.0 ( $N_1^{(S)} = 20$ ). This small difference is within the experimental uncertainty of available ion counting methods, including equilibrium dialysis (14) and ASAXS (25). However, because the shape of the ion atmosphere depends on

$N_1^{(S)}$  (Fig. 5 b), SAXS may be expected to discriminate between scattering profiles predicted for each model.

We calculated SAXS profiles for each ion atmosphere model including explicit and diffuse (NPBE) ions (Fig. 5), as described in the Supporting Material, and compared them with the experimental SAXS profiles. Relative to the partial form factors, experimental SAXS curves have a greater signal/noise ratio and therefore represent a more challenging test of the model. The predictions from the hydrated DNA model with an NPBE-derived ion atmosphere with  $r = 2$  Å (Fig. 6 a,  $N_1^{(S)} = 0$ ) do not agree well with the experimental data. The disagreement between model and data is most obvious for the electron-rich counterions Cs<sup>+</sup> and Rb<sup>+</sup>, where the ion atmosphere represents a more significant part of the total scattering. From visual comparison of hybrid explicit/PB models and experiment for DNA in CsCl and RbCl salts (Fig. 6 a), the agreement improves and then worsens with increasing  $N_1^{(S)}$ , showing a best-fit value between 10 and 15. More quantitatively, the total  $\chi^2$  between model and experiment has a minimum in this range of  $N_1^{(S)}$  (Fig. 6 b).

The distinction between explicitly placed and diffuse ions in our model should not be overinterpreted. Given the limited resolution of our SAXS measurements, we expect them to be sensitive not to the precise ion configuration, but rather to the low-order moments of the ion distribution. In other words, the  $N_1^{(S)} = 10$  model in Fig. 5 a should not be viewed as a unique solution. To underscore this point, we repeated the model calculations with different parameter values and fit these to the data. Since the Stern radius is effectively an ad hoc parameter—within a physically meaningful range, it controls the surface potential and ion accessibility—we repeated the

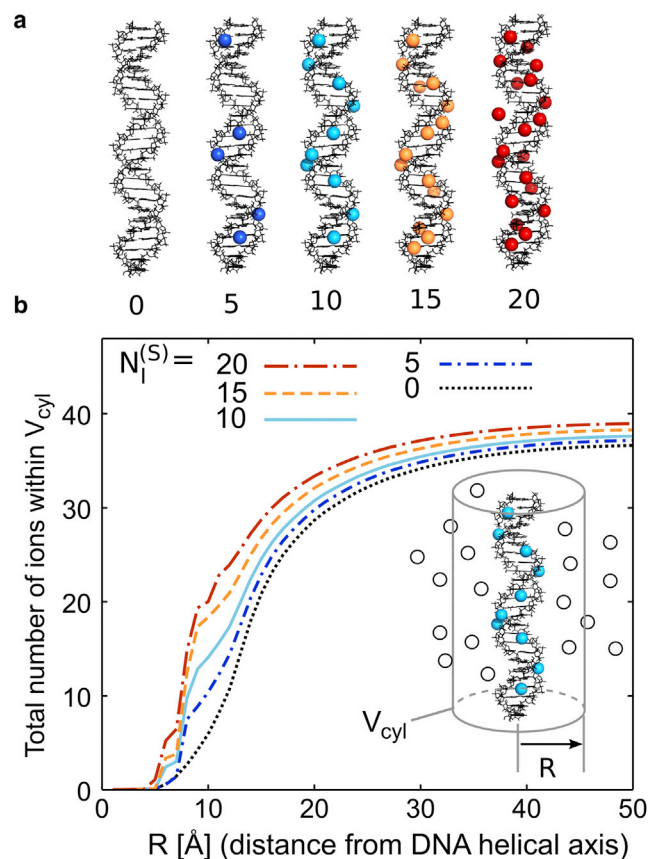


FIGURE 5 Modeling of surface-localized ions and their effect on the diffuse ion atmosphere. (a) Ions were placed sequentially at sites near the surface of the DNA molecule with the greatest magnitude of the potential, and the NPBE was solved again after each ion addition (see text). DNA models with ions represented as  $r = 2$  Å spheres are shown for  $N_1^{(S)} = 0, 5, 10, 15,$  and  $20$  (rendered using Pymol version 1.2r1, DeLano Scientific, Portland, OR). (b) For each model, the total number of ions within a cylindrical volume,  $V_{cyl}$ , with radius  $R$  centered on the helical axis of the DNA (inset) was computed. The number of excess ions includes the explicit ions shown in (a) and the excess ion density from the NPBE solution (open circles). To see this figure in color, go online.

calculations with  $r = 2 - 4$  Å (Figs. S2 and S3). The best-fit range of  $N_1^{(S)}$  changed from  $\sim 10 - 15$  ( $r = 2$  Å) to  $\sim 20 - 25$  ( $r = 3$  Å) and  $\sim 25 - 35$  ( $r = 4$  Å). However, the radius of gyration (second moment) of the ion atmosphere about the  $z$  axis of the DNA ( $R_{GZ}$ ) in each of these best-fit ranges is the same (Fig. S3, inset). Thus,  $R_{GZ}$  is a more salient aspect of the ion atmosphere than  $N_1^{(S)}$  from the perspective of the SAXS data. At best fit, the value of  $R_{GZ} \approx 16$  Å is significantly smaller than the value obtained from PB theory (for  $r = 2$  Å,  $R_{GZ} = 18$  Å), and this conclusion is robust to the details of the calculation. This smaller value underscores our main point that ions must be closer to the surface than PB predicts.

## DISCUSSION

We developed methods for collecting and analyzing SAXS data using the contrast variation technique of isomorphous

replacement. These methods built on previous experience with ASAXS by adding absolute-intensity calibration to the isomorphous replacement method. We found that the scattering strength of monovalent ions around DNA was well-described by an effective scattering factor calculated from their partial molar volumes, and we obtained excess ion numbers consistent with previous measurements. In addition, the number of water molecules displaced by DNA was measured by extrapolation to zero ion contrast. Isomorphous replacement data were decomposed into contributions from the ions and those from DNA alone. The DNA-only profile agreed with predictions from a B-form duplex model with hydration parameters set according to the experimentally measured number of water molecules displaced. However, when the NPBE was used to model the counterion atmosphere, we found significant disagreement between predicted x-ray scattering profiles and three related measurements: 1) the ion-DNA cross-term determined by heavy-atom substitution; 2) the ASAXS anomalous difference profiles; and 3) SAXS profiles for DNA with electron-dense counterions. All are sensitive to the spatial distribution of ions around DNA and suggest that the counterions are more closely localized to the DNA surface than predicted.

From a theoretical perspective, the failure of PB in this case is not surprising. However, previous x-ray scattering studies of DNA tend to confirm predictions from PB (15,21,24,45). There are several important differences between this study and those preceding it. First, we calculate the ion distribution using an all-atom representation of the DNA molecule, a clear improvement over previous studies that modeled DNA as a charged cylinder (21,24). In addition, we model hydration and excluded volume in real space instead of relying on the effective scattering factor approach (1,15). Earlier studies that used ASAXS to probe the ion atmosphere and all-atom NPBE for modeling reported qualitative (45) and quantitative (15) agreement with the data. The latter study also applied the isomorphous replacement method. However, the ionic strength,  $I$ , of the solution was 400 mM, as compared to 100 mM used here. Because the ionic strength affects the length scale of the diffuse ion atmosphere through the Debye screening length, proportional to  $I^{-1/2}$ , we speculate that the SAXS measurements of Das et al. (15) were unable to distinguish between surface-localized and diffuse counterions because the diffuse ion atmosphere was too compact (screening length  $\sim 5$  Å) relative to the resolution of SAXS. In the low ionic strength measured here (screening length  $\sim 10$  Å), the spatial separation between diffuse and surface-localized ions is more pronounced, and therefore our measurement is more sensitive to the ion distribution.

Although x-ray scattering data for DNA have not previously shown disagreement with PB theory, theoretical and experimental studies have revealed limitations. Dialysis experiments show that PB theory does not account for the

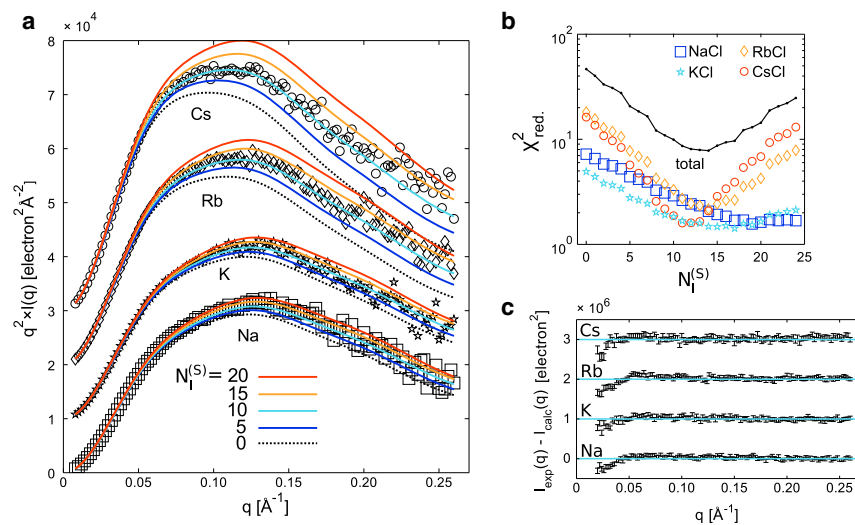


FIGURE 6 Adding surface-localized ions to DNA models improves the fit with isomorphous replacement SAXS profiles. (a) Theoretical scattering profiles were computed from the hydrated DNA model (Fig. 2 a) plus an ion atmosphere modeled using the NPBE (dotted line) with varying numbers of explicit ions,  $N_I^{(S)}$ , at the DNA surface (models shown in Fig. 5 a). Profiles were computed for each ion type (Na, K, Rb, and Cs), plotted as  $I(q) \times q^2$  to emphasize the high-angle features, and offset by  $1 \times 10^4$  for clarity. Experimental intensities (from Fig. 1) were rescaled slightly by a factor  $\alpha$  to minimize the  $\chi^2$  for each predicted curve ( $\alpha$  deviated from 1 by  $<3\%$ ). In (a),  $\alpha$  has the value corresponding to the  $N_I^{(S)} = 10$  theoretical profiles. (b) The agreement between experimental data and each theoretical profile was assessed using a reduced  $\chi^2$ , shown on a logarithmic scale for each experimental curve (NaCl, KCl, RbCl, and CsCl) and for the whole data set (total). (c) The residual for  $N_I^{(S)} = 10$  is plotted on a standard intensity scale with experimental error bars. Each curve is offset by  $1 \times 10^6$  for clarity.

competition of  $\text{Mg}^{2+}$  against a background of 20 mM  $\text{Na}^+$  (14), even with modifications to account for the unequal ion sizes. In explicit-solvent MD simulations of RNA duplexes with monovalent cations, PB theory does not account for the ion distributions (17). Of the several simplifications inherent in PB theory, which ones, if corrected, would improve ion atmosphere models for DNA? PB theory neglects the possibility of direct interactions between ions and DNA involving full or partial dehydration of the ion (46). We find that the accuracy of the PB-derived SAXS profile increases as the Stern layer thickness decreases, which would seem to support the idea that the ions closest to the DNA are dehydrated. However, dehydration is unlikely to be important for the monovalent ions studied here, as NMR measurements show that Group I cations interact with DNA in a fully hydrated state (47). The mean-field PB theory also neglects ion-ion correlations, which may affect the free energy of ions occupying the highly concentrated region near the molecular surface. Chu et al. argue for the importance of ion correlations when there is a discrepancy between counterion competition data and size-modified NPBE calculations (16). In addition, PB theory neglects solvent degrees of freedom, treating water as a uniform dielectric medium. However, the water near molecular surfaces is highly perturbed, and it is thus unlikely to mediate electrostatic interactions in the same way as would bulk water. Interestingly, Kirmizialtin and co-workers have found that discrepancies between NPBE and all-atom MD for  $\text{Na}^+$  and  $\text{Sr}^{2+}$  ions around RNA could be largely corrected by artificially lowering the solvent dielectric constant near the surface (17). The physical justification for this correction is unclear, but it points to a central role of water in mediating ion-DNA interactions.

We posit that under physiological conditions, the ion atmospheres around DNA are strongly influenced by the elec-

trostatic screening properties of the solvent near the molecular surface. These screened electrostatic fields are important determinants of the overall free energy for conformation changes and ligand interactions. SAXS-based measurements of the spatial distribution of ions and water around highly charged molecules such as DNA provides a valuable experimental constraint on theories that seek to account for such complex solvent interactions. The recent application of integral equation theory to macromolecular systems promises to provide a realistic treatment of all solvent degrees of freedom without the computational cost of explicit-solvent MD (19,48). The x-ray scattering methods presented here will be essential for testing these and other emerging first-principles computational models.

## SUPPORTING MATERIAL

Supporting Theory, Supporting Materials and Methods, and three figures are available at [http://www.biophysj.org/biophysj/supplemental/S0006-3495\(15\)00490-7](http://www.biophysj.org/biophysj/supplemental/S0006-3495(15)00490-7).

## AUTHOR CONTRIBUTIONS

L.P. designed the research, S.P.M. and S.A.P. performed the research, and S.P.M. analyzed the data. All authors wrote the article.

## ACKNOWLEDGMENTS

We thank Cornell High Energy Synchrotron Source (CHESS) scientists Arthur Woll and Ken Finkelstein and CHESS staff for assistance with the experiments; members of the Pollack Lab including Julie Sutton, Andrea Katz, Joshua Tokuda, Yujie Chen, Joshua Blose, and Huimin Chen for help with x-ray data collection; and David Case, Hung Nguyen, Nathan Baker, Serdal Kirmizialtin, and Ron Elber for insightful discussions.

This work was supported by the National Institutes of Health (NIH) (R01-GM085062). CHESS is supported by the National Science Foundation



(NSF) and by the National Institutes of Health/National Institute of General Medical Sciences (NIH/NIGMS) via NSF award DMR-0936384, and the MacCHESS resource is supported by NIGMS award GM-103485.

## SUPPORTING CITATIONS

References (49–59) appear in the Supporting Material.

## REFERENCES

1. Qiu, X., L. W. Kwok, ..., L. Pollack. 2006. Measuring inter-DNA potentials in solution. *Phys. Rev. Lett.* 96:138101.
2. Qiu, X., K. Andresen, ..., L. Pollack. 2008. Abrupt transition from a free, repulsive to a condensed, attractive DNA phase, induced by multivalent polyamine cations. *Phys. Rev. Lett.* 101:228101.
3. Qiu, X., V. A. Parsegian, and D. C. Rau. 2010. Divalent counterion-induced condensation of triple-strand DNA. *Proc. Natl. Acad. Sci. USA.* 107:21482–21486.
4. Bloomfield, V. A. 1997. DNA condensation by multivalent cations. *Biopolymers.* 44:269–282.
5. Abouzar, M. H., A. Poghosian, ..., M. J. Schöning. 2012. Label-free electrical detection of DNA by means of field-effect nanoplate capacitors: experiments and modeling. *Physica Status Solidi A Appl. Mater. Sci.* 209:925–934.
6. Poghosian, A., A. Cherstvy, ..., M. Schöning. 2005. Possibilities and limitations of label-free detection of DNA hybridization with field-effect-based devices. *Sens. Actuators B Chem.* 111–112:470–480.
7. Stern, E., R. Wagner, ..., M. A. Reed. 2007. Importance of the Debye screening length on nanowire field effect transistor sensors. *Nano Lett.* 7:3405–3409.
8. Star, A., E. Tu, ..., C. Valcke. 2006. Label-free detection of DNA hybridization using carbon nanotube network field-effect transistors. *Proc. Natl. Acad. Sci. USA.* 103:921–926.
9. Wong, I. Y., and N. A. Melosh. 2009. Directed hybridization and melting of DNA linkers using counterion-screened electric fields. *Nano Lett.* 9:3521–3526.
10. Grochowski, P., and J. Trylska. 2008. Continuum molecular electrostatics, salt effects, and counterion binding—a review of the Poisson-Boltzmann theory and its modifications. *Biopolymers.* 89:93–113.
11. Rocchia, W., E. Alexov, and B. Honig. 2001. Extending the applicability of the nonlinear Poisson-Boltzmann equation: multiple dielectric constants and multivalent ions. *J. Phys. Chem. B.* 105:6507–6514.
12. 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.
13. Holst, M., N. Baker, and F. Wang. 2000. Adaptive multilevel finite element solution of the Poisson-Boltzmann equation I. Algorithms and examples. *J. Comput. Chem.* 21:1319–1342.
14. 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.
15. Das, R., T. T. Mills, ..., L. Pollack. 2003. Counterion distribution around DNA probed by solution x-ray scattering. *Phys. Rev. Lett.* 90:188103.
16. 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.
17. Kirmizialtin, S., A. R. Silalahi, ..., M. O. Fenley. 2012. The ionic atmosphere around A-RNA: Poisson-Boltzmann and molecular dynamics simulations. *Biophys. J.* 102:829–838.
18. Tan, Z.-J., and S.-J. Chen. 2009. Predicting electrostatic forces in RNA folding. *Methods Enzymol.* 469:465–487.
19. Giambaşu, G. M., T. Luchko, ..., D. A. Case. 2014. Ion counting from explicit-solvent simulations and 3D-RISM. *Biophys. J.* 106:883–894.
20. 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.
21. Chang, S. L., S. H. Chen, ..., J. Lin. 1990. Measurements of monovalent and divalent counterion distributions around persistence length DNA fragments in solution. *J. Phys. Chem.* 94:8025–8028.
22. Morfin, I., F. Horkay, ..., E. Geissler. 2004. Adsorption of divalent cations on DNA. *Biophys. J.* 87:2897–2904.
23. Pabit, S. A., K. D. Finkelstein, and L. Pollack. 2009. Using anomalous small angle x-ray scattering to probe the ion atmosphere around nucleic acids. *Methods Enzymol.* 469:391–410.
24. Andresen, K., R. Das, ..., L. Pollack. 2004. Spatial distribution of competing ions around DNA in solution. *Phys. Rev. Lett.* 93:248103.
25. 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.
26. Greenfeld, M., and D. Herschlag. 2009. Probing nucleic acid-ion interactions with buffer exchange-atomic emission spectroscopy. *Methods Enzymol.* 469:375–389.
27. Strauss, U. P., C. Helfgott, and H. Pink. 1967. Interactions of polyelectrolytes with simple electrolytes. II. Donnan equilibria obtained with DNA in solutions of 1-1 electrolytes. *J. Phys. Chem.* 71:2550–2556.
28. Grilley, D., A. M. Soto, and D. E. Draper. 2009. Direct quantitation of Mg<sup>2+</sup>-RNA interactions by use of a fluorescent dye. *Methods Enzymol.* 455:71–94.
29. Kuwamoto, S., S. Akiyama, and T. Fujisawa. 2004. Radiation damage to a protein solution, detected by synchrotron x-ray small-angle scattering: dose-related considerations and suppression by cryoprotectants. *J. Synchrotron Radiat.* 11:462–468.
30. Luzzati, V., F. Masson, ..., P. Saludjian. 1967. X-ray scattering study of rigid polyelectrolyte solution. lithium, sodium, and cesium salts of DNA. *Biopolymers.* 5:491–508.
31. Meisburger, S. P., J. L. Sutton, ..., L. Pollack. 2013. Polyelectrolyte properties of single stranded DNA measured using SAXS and single-molecule FRET: beyond the wormlike chain model. *Biopolymers.* 99:1032–1045.
32. Bardhan, J. P. 2012. Biomolecular electrostatics—I want your solvation (model). *Comput. Sci. Discov.* 5:013001.
33. Nielsen, S. S., M. Møller, and R. E. Gillilan. 2012. High-throughput biological small-angle x-ray scattering with a robotically loaded capillary cell. *J. Appl. Cryst.* 45:213–223.
34. Huang, T. C., H. Toraya, ..., Y. Wu. 1993. X-ray powder diffraction analysis of silver behenate, a possible low-angle diffraction standard. *J. Appl. Crystallogr.* 26:180–184.
35. Orthaber, D., A. Bergmann, and O. Glatter. 2000. SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. *J. Appl. Crystallogr.* 33:218–225.
36. Hansen, S. 2000. Bayesian estimation of hyperparameters for indirect Fourier transformation in small-angle scattering. *J. Appl. Crystallogr.* 33:1415–1421.
37. Pollack, L. 2011. SAXS studies of ion-nucleic acid interactions. *Annu Rev Biophys.* 40:225–242.
38. Qiu, X., K. Andresen, ..., L. Pollack. 2007. Inter-DNA attraction mediated by divalent counterions. *Phys. Rev. Lett.* 99:038104.
39. Haynes, W. M. 2014. CRC Handbook of Chemistry and Physics, 95th ed. CRC Press, Boca Raton, FL.
40. Marcus, Y. 1985. Ion Solvation. Wiley, Hoboken, NJ.
41. Marcus, Y. 1988. Ionic radii in aqueous solutions. *Chem. Rev.* 88:1475–1498.
42. 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.
43. Svergun, D., C. Barberato, and M. H. J. Koch. 1995. CRYSOLE—a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* 28:768–773.

44. Tan, Z.-J., and S.-J. Chen. 2005. Electrostatic correlations and fluctuations for ion binding to a finite length polyelectrolyte. *J. Chem. Phys.* 122:44903.
45. 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.
46. Draper, D. E. 2004. A guide to ions and RNA structure. *RNA*. 10: 335–343.
47. Braunlin, W. H. 1996. NMR studies of cation-binding environments on nucleic acids. In *Advances in Biophysical Chemistry, Vol. 5*. Allen Bush, editor. Newnes, Burlington, MA.
48. Nguyen, H. T., S. A. Pabit, ..., D. A. Case. 2014. Accurate small and wide angle x-ray scattering profiles from atomic models of proteins and nucleic acids. *J. Chem. Phys.* 141:22D508.
49. Andresen, K., X. Qiu, ..., L. Pollack. 2008. Mono- and trivalent ions around DNA: a small-angle scattering study of competition and interactions. *Biophys. J.* 95:287–295.
50. Li, L., S. A. Pabit, ..., L. Pollack. 2011. Double-stranded RNA resists condensation. *Phys. Rev. Lett.* 106:108101.
51. Blöse, J. M., S. A. Pabit, ..., L. Pollack. 2011. Effects of a protecting osmolyte on the ion atmosphere surrounding DNA duplexes. *Biochemistry*. 50:8540–8547.
52. Cantor, C. R., M. M. Warshaw, and H. Shapiro. 1970. Oligonucleotide interactions. 3. Circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers*. 9:1059–1077.
53. Tataurov, A. V., Y. You, and R. Owczarzy. 2008. Predicting ultraviolet spectrum of single stranded and double stranded deoxyribonucleic acids. *Biophys. Chem.* 133:66–70.
54. Park, S., J. P. Bardhan, ..., L. Makowski. 2009. Simulated x-ray scattering of protein solutions using explicit-solvent models. *J. Chem. Phys.* 130:134114.
55. Durchschlag, H., and P. Zipper. 2003. Modeling the hydration of proteins: prediction of structural and hydrodynamic parameters from x-ray diffraction and scattering data. *Eur. Biophys. J.* 32:487–502.
56. Yang, S., M. Parisien, ..., B. Roux. 2010. RNA structure determination using SAXS data. *J. Phys. Chem. B*. 114:10039–10048.
57. Merzel, F., and J. C. Smith. 2002. SASSIM: a method for calculating small-angle x-ray and neutron scattering and the associated molecular envelope from explicit-atom models of solvated proteins. *Acta Crystallogr. D Biol. Crystallogr.* 58:242–249.
58. Dolinsky, T. J., J. E. Nielsen, ..., N. A. Baker. 2004. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* 32:W665–W667.
59. Kell, G. S. 1970. Isothermal compressibility of liquid water at 1 atm. *J. Chem. Eng. Data*. 15:119–122.

# Supporting Material for: “Determining the locations of ions and water around DNA from X-ray scattering measurements”

Steve P. Meisburger<sup>†</sup>, Suzette A. Pabit<sup>†</sup>, and Lois Pollack<sup>†\*</sup>

<sup>†</sup>School of Applied and Engineering Physics, Cornell University, Ithaca, New York

## SUPPORTING THEORY

The following sections provide derivations of equations 1-6 in the Main Text, with supporting references and discussion.

### Fundamental equations of X-ray scattering from a macromolecule in solution

The fundamental equation that describes a scattering experiment for a dilute solution of macromolecules is (1)

$$I_{sub}(q)/N = \left\langle \left| A_M(\mathbf{q}) + A_{solv}(\mathbf{q}) - \overline{B_{solv}(\mathbf{q})} \right|^2 \right\rangle_{\Omega, \Pi} - \left\langle \left| B_{solv}(\mathbf{q}) - \overline{B_{solv}(\mathbf{q})} \right|^2 \right\rangle_{\Omega, \Pi} \quad (S1)$$

where  $\mathbf{q}$  is the momentum transfer,  $I_{sub}(q)/N$  is the scattering per macromolecule (in excess of the bulk solvent background),  $A_M(\mathbf{q})$  is the scattering amplitude of the macromolecule (in vacuum),  $A_{solv}(\mathbf{q})$  is amplitude of the solvent molecules in a volume  $V$  that encloses the macromolecule and its region of influence, and  $B_{solv}(\mathbf{q})$  is the amplitude of bulk-like solvent occupying a volume identical in shape to  $V$  but with no macromolecule present. The angle brackets with subscripts  $\Omega$  and  $\Pi$  signify the averages over orientation and configuration respectively. The over-bar indicates a configuration-average with the positions of the macromolecule and solvent particles underlying  $A_M(\mathbf{q})$  and  $A_{solv}(\mathbf{q})$  held constant.

In Equation S1, the term  $\left\langle \left| B_{solv}(\mathbf{q}) - \overline{B_{solv}(\mathbf{q})} \right|^2 \right\rangle_{\Omega, \Pi}$  is very small at low angles, and its calculation is only necessary for modeling WAXS profiles ( $q$  of order  $1 \text{ \AA}^{-1}$ ). Here, we set it equal to zero and approximate the total scattering intensity as

$$I_{sub}(q)/N \approx \left\langle \left| A_M(\mathbf{q}) + A_{solv}(\mathbf{q}) - \overline{B_{solv}(\mathbf{q})} \right|^2 \right\rangle_{\Omega, \Pi} \quad (S2)$$

Each amplitude term can be expressed as a sum over  $K$  different types of spherically symmetric particles,

$$A(\mathbf{q}) = \sum_{k=1}^K f_k(q) \sum_{j=1}^{N_k} e^{-i\mathbf{q}\cdot\mathbf{r}_j} \quad (S3)$$

where  $f_k(q)$  is the scattering factor of the  $k^{\text{th}}$  type of particle, and  $r_j$  is the position of the  $j^{\text{th}}$  particle of type  $k$ . Because hydrogen atoms have weak scattering, for computational efficiency their contributions are folded into an effective scattering factor for their chemical group, as in the CRY SOL program (2). Similarly, water molecules are modeled by a spherically-symmetric, effective scattering factor  $f_w(q)$  that combines the contributions from the oxygen the hydrogen atoms (3). Otherwise,  $f_k(q)$  are the atomic form factors (4).

### Excess scattering amplitude of a charged macromolecule and its ion atmosphere

We consider a macromolecule with fixed atomic coordinates whose solvent atmosphere consists of water and one ion species. This case applies polyions (such as DNA) in low concentration monovalent salts, and is easily generalized to more complex

solvents. First, we derive a straightforward expression for the excess solvent amplitude,  $A_{solv}(\mathbf{q}) - \overline{B_{solv}(\mathbf{q})}$ . The solvent around the macromolecule, consisting of  $N_w$  water molecules and  $N_I$  ions, has a scattering amplitude given by Equation S3,

$$A_{solv}(\mathbf{q}) = f_w(q) \sum_{j=1}^{N_w} e^{-i\mathbf{q}\cdot\mathbf{r}_j} + f_I(q) \sum_{k=1}^{N_I} e^{-i\mathbf{q}\cdot\mathbf{r}_k} \quad (\text{S4})$$

The configurational average of the excluded solvent amplitude can be expressed as a Fourier transform of the average electron density in the excluded volume  $V$ ,

$$\overline{B_{solv}(\mathbf{q})} = \overline{\rho_{solv}} \int_V e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} \quad (\text{S5})$$

where the average bulk solvent electron density is  $\overline{\rho_{solv}}$ . The exact volume  $V$  remains to be defined. A minimal choice includes the immediate region around the macromolecule (labeled by the superscript  $M$ ) plus the region around each ion (the  $k^{\text{th}}$  ion is labeled by the superscript  $I_k$ ). If the  $N_w$  water molecules are partitioned into these sub-volumes ( $N_w^{(M)}$  waters in the volume  $V^{(M)}$ , and  $N_w^{(I_k)}$  waters in the volume  $V^{(I_k)}$ ), the excess solvent amplitude is

$$\begin{aligned} A_{solv}(\mathbf{q}) - \overline{B_{solv}(\mathbf{q})} &= f_w(q) \sum_{j=1}^{N_w^{(M)}} e^{-i\mathbf{q}\cdot\mathbf{r}_j} - \overline{\rho_{solv}} \int_{V^{(M)}} e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} \\ &+ \sum_{k=1}^{N_I} e^{-i\mathbf{q}\cdot\mathbf{r}_k} \left( f_I(q) + f_w(q) \sum_{l=1}^{N_w^{(I_k)}} e^{-i\mathbf{q}\cdot\mathbf{r}'_l} - \overline{\rho_{solv}} \int_{V^{(I_k)}} e^{-i\mathbf{q}\cdot\mathbf{r}'} d\mathbf{r}' \right) \end{aligned} \quad (\text{S6})$$

The primed coordinates are relative to each ion's position (e.g.  $\mathbf{r}'_l = \mathbf{r}_l - \mathbf{r}_k$ ).

If an ion's hydration shell is independent of its position around the macromolecule (i.e. ions do not dehydrate or otherwise "bind" the macromolecule), the term in parenthesis is approximately the same for all ions (independent of  $k$ ), and can be replaced by an averaged effective scattering factor. The simplest choice of effective scattering factor is a constant,  $Z_{\text{eff}}$ , obtained by setting  $q = 0$  inside the parentheses:

$$Z_{\text{eff}} = f_I(0) + f_w(0) N_w^{(I)} - \overline{\rho_{solv}} V^{(I)} \quad (\text{S7})$$

This effective scattering factor is evidently valid only at low scattering angles. In the case where the bulk solvent electron density is close to that of pure water,

$$\overline{\rho_{solv}} \approx f_w(0) \overline{n_w} \quad (\text{S8})$$

where  $\overline{n_w}$  is the number density of water molecules. Then the approximate scattering factor of the hydrated ions is

$$Z_{\text{eff}} \approx f_I(0) - f_w(0) \overline{n_w} V_I \quad (\text{S9})$$

where  $V_I$  is the volume change due to the presence of the ion

$$V_I = V^{(I)} - N_w^{(I)} / \overline{n_w} \quad (\text{S10})$$

equal to the absolute limiting partial molar volume (5) divided by Avogadro's number. Equation S9 is the same as Equation 2 in the Main Text, with the notational change  $Z_I = f_I(0)$  and  $\overline{\rho}_e = f_w(0) \overline{n_w}$ .

Finally, substituting  $Z_{\text{eff}}$  into Equation S6, we obtain an expression for the excess solvent scattering amplitude

$$A_{solv}(\mathbf{q}) - \overline{B_{solv}(\mathbf{q})} \approx f_w(q) \sum_{j=1}^{N_w^{(M)}} e^{-i\mathbf{q}\cdot\mathbf{r}_j} - \overline{\rho_{solv}} \int_{V^{(M)}} e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} + Z_{\text{eff}} \sum_{k=1}^{N_I} e^{-i\mathbf{q}\cdot\mathbf{r}_k} \quad (\text{S11})$$

After substitution of Equation S11 into Equation S2, the intensity is

$$I_{sub}(q)/N \approx \left\langle \left| A_M(\mathbf{q}) + f_w(q) \sum_{j=1}^{N_w^{(M)}} e^{-i\mathbf{q}\cdot\mathbf{r}_j} - \overline{\rho_{solv}} \int_{V^{(M)}} e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} + Z_{\text{eff}} \sum_{k=1}^{N_I} e^{-i\mathbf{q}\cdot\mathbf{r}_k} \right|^2 \right\rangle_{\Omega, \Pi} \quad (\text{S12})$$

## Total scattering intensity as a two-phase system

Central to the decomposition method is a treatment of the scattering as a two-phase system: the macromolecule and associated water constitute the first phase (M), hydrated ions are the second phase (I). The intensity for this two-phase system is

$$I_{sub}(q)/N \approx \left\langle |F_M(q) + F_I(q)|^2 \right\rangle_{\Omega, \Pi} \quad (S13)$$

where  $F_M(q)$  and  $F_I(q)$  are the amplitudes of each phase. With reference to Equation S12, the amplitudes are

$$F_M(q) = A_M(\mathbf{q}) + f_w(q) \sum_{j=1}^{N_w^{(M)}} e^{-i\mathbf{q}\cdot\mathbf{r}_j} - f_w(0) \overline{n_w} \int_{V^{(M)}} e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r} \quad (S14)$$

and

$$F_I(q) = Z_{\text{eff}} \sum_{k=1}^{N_I} e^{-i\mathbf{q}\cdot\mathbf{r}_k} \quad (S15)$$

Next, we expand Equation S13

$$I_{sub}(q)/N \approx \left\langle |F_M(q)|^2 \right\rangle_{\Omega, \Pi} + 2 \langle \text{Re}(F_M(q) F_I^*(q)) \rangle_{\Omega, \Pi} + \left\langle |F_I(q)|^2 \right\rangle_{\Omega, \Pi} \quad (S16)$$

The magnitude of each term at  $q = 0$  can be separated from the  $q$ -dependent part by defining real functions  $P(q)$  with  $P(0) = 1$ ,

$$I_{sub}(q)/N \approx |F_M(0)|^2 P_M(q) + 2 \text{Re}(F_M(0) \langle F_I^*(0) \rangle_{\Pi}) P_{MI}(q) + \left\langle |F_I(0)|^2 \right\rangle_{\Pi} P_I(q) \quad (S17)$$

To simplify the notation, we define the partial amplitudes

$$\delta_M = F_M(0) \quad (S18)$$

$$\delta_I N_I = F_I(0) \quad (S19)$$

If the X-ray energy is far from the absorption edges for the atoms, the scattering factors at  $q = 0$  are real numbers, and  $\delta_M$  and  $\delta_I$  are real. Furthermore, the configurational averages may be dropped from Equation S17 provided the number fluctuations for ions and water are small (i.e.  $\langle N^2 \rangle \sim \langle N \rangle^2$ ), leaving

$$I_{sub}(q)/N \approx \delta_M^2 P_M(q) + 2\delta_M N_I \delta_I P_{MI}(q) + (N_I \delta_I)^2 P_I(q) \quad (S20)$$

Equation S20 is identical to Equation 4 in the Main Text. From Equations S14, S15, S18 and S19, the partial amplitudes  $\delta$  are equal to

$$\delta_M = A_M(0) + f_w(0) \left( N_w^{(M)} - \overline{n_w} V^{(M)} \right) \quad (S21)$$

$$\delta_I = Z_{\text{eff}} \quad (S22)$$

Equation 3 in the Main Text,  $\delta_M = Z_{DNA} + 10(N_H - N_E)$ , can be obtained from Equation S21, above, by the change of notation,  $A_M(0) = Z_{DNA}$ ,  $f_w(0) = 10$ ,  $N_w^{(M)} = N_H$ , and  $\overline{n_w} V^{(M)} = N_E$ .

## Two-phase model for anomalous scattering

In an anomalous scattering experiment, the X-ray energy is tuned near the absorption edge for an atom of interest. If the ion contrast is varied in this way, its scattering factor is

$$f_I(q, E) = f_I(q) + f'(E) + i f''(E) \quad (S23)$$

Thus, the effective scattering amplitude of the ion is modified by

$$\delta_I(E) = Z_{\text{eff}} + f'(E) + i f''(E) \quad (S24)$$

The anomalous scattering experiment measures  $I_{sub}(q)$  at two energies,  $E_1$  and  $E_2$ . The anomalous difference is

$$I_{anom}(q) = I_{sub}(q, E_2) - I_{sub}(q, E_1) \quad (S25)$$

According to Equations S17, S18, S19 and S24,

$$I_{anom}(q)/N = 2 \langle N_I \rangle \delta_M (Z_{eff} + f'(E_2)) P_{MI}(q) + \langle N_I^2 \rangle \left\{ (Z_{eff} + f'(E_2))^2 + f''(E_2)^2 \right\} P_I(q) \\ - 2 \langle N_I \rangle \delta_M (Z_{eff} + f'(E_1)) P_{MI}(q) - \langle N_I^2 \rangle \left\{ (Z_{eff} + f'(E_1))^2 + f''(E_1)^2 \right\} P_I(q) \quad (S26)$$

This simplifies to

$$I_{anom}(q)/N = 2 \langle N_I \rangle \delta_M (f'(E_2) - f'(E_1)) P_{MI}(q) \\ + \langle N_I^2 \rangle \left\{ 2 Z_{eff} (f'(E_2) - f'(E_1)) + (f'(E_2) - f'(E_1)) (f'(E_2) + f'(E_1)) + f''(E_2)^2 - f''(E_1)^2 \right\} P_I(q) \quad (S27)$$

In the ASAXS experiment, we chose the energies  $E_1$  and  $E_2$  to be below the absorption edge, so that  $f''(E_2) \approx f''(E_1)$ . Thus, Equation S27 simplifies further:

$$I_{anom}(q)/N = 2 \langle N_I \rangle \delta_M (f'(E_2) - f'(E_1)) (P_{MI}(q) + \gamma P_I(q)) \quad (S28)$$

where

$$\gamma = \frac{\langle N_I^2 \rangle}{\langle N_I \rangle \delta_M} \left( Z_{eff} + \frac{1}{2} (f'(E_2) + f'(E_1)) \right) \quad (S29)$$

Again, assuming the number density fluctuations are small, so that  $\langle N_I^2 \rangle \sim \langle N_I \rangle^2$ ,

$$\gamma \approx \frac{\langle N_I \rangle}{\delta_M} \left( Z_{eff} + \frac{1}{2} (f'(E_2) + f'(E_1)) \right) \quad (S30)$$

Equations S28 and S30 are the same as Equation 6 in the Main Text.

## SUPPORTING MATERIALS AND METHODS

### Sample preparation

The 25 base-pair DNA duplex was composed of two single stranded oligonucleotides with sequence GCATCTGGGC-TATAAAGGGCGTCG and its complement. The sequence has been used in many previous small-angle X-ray scattering (SAXS) studies (6–14). Oligonucleotides were synthesized and HPLC purified by Integrated DNA Technologies (Coralville, IA) and delivered as lyophilized powders. Each strand was rehydrated in aqueous buffer containing 10 mM TRIS, 50 mM NaCl, 1 mM EDTA, pH 8.0. The concentration of each strand was calculated from the UV absorption at 260 nm using extinction coefficients derived from the nearest-neighbor model (15);  $\epsilon_{S1} = 244,400 \text{ L mol}^{-1} \text{ cm}^{-1}$ , and  $\epsilon_{S2} = 228,500 \text{ L mol}^{-1} \text{ cm}^{-1}$ . Strands were mixed in an equimolar ratio, annealed at 94 C for 4 minutes, and allowed to cool on the bench.

Buffered salt solutions were prepared using 1 mM Na-MOPS pH 7.0, 100 mM of 1:1 salt (NaCl, KCl, RbCl, CsCl) and NANOpure water (Barnstead, Dubuque, IA). All reagents were purchased from Sigma-Aldrich, unless specified. DNA in each buffered salt solution was prepared by spin dialysis using an Amicon Ultra-0.5 mL with a 10 kDa cutoff (Millipore, Billerica, MA), repeated to ensure complete exchange. The DNA concentration during spin dialysis did not exceed 0.1 mM. The final flow-through was retained for SAXS background subtraction. The UV absorbance of each duplex DNA solution at 260 nm was converted to concentration using the hypochromicity-corrected extinction coefficient (16),  $\epsilon_{S1+S2} = 397,600 \text{ L mol}^{-1} \text{ cm}^{-1}$ .

In previous studies, 50  $\mu\text{M}$  DNA was shown to agree with the infinite dilution limit within the measurement noise (6). However, considering improvements in signal strength and extended  $q$ -range of the measurements, some residual interparticle interference is possible at this concentration, if difficult to confirm experimentally (e.g. by collecting data at even lower concentration). Therefore, we modeled the DNA structure factor theoretically using parameters determined previously for this system at high DNA concentrations (6). At 50  $\mu\text{M}$ , the structure factor contribution is predicted to be less than 3% of the signal at  $q = 0$ , and it decays rapidly to zero by  $q \sim 0.05 \text{ \AA}^{-1}$  (Figure S1). When the fits shown in Figure 6 in the Main Text are repeated using data truncated at  $q = 0.05 \text{ \AA}^{-1}$ , we find little effect on the minimum value of  $N_I^{(S)}$  (Figure S3).

### Absolute intensity calibration

For the ASAXS and heavy atom experiments, the scattering intensity was placed on an absolute scale using liquid water as a calibrant (17). The forward X-ray scattering of a liquid (i.e. the macroscopic scattering cross-section,  $d\Sigma/d\Omega$ ) is proportional to the isothermal compressibility:

$$\frac{d\Sigma}{d\Omega} = n r_0^2 Z^2 k_B T \chi_T \quad (\text{S31})$$

where  $r_0$  is the classical electron radius,  $n$  is the molecular number density,  $Z$  is the number of electrons per molecule (10 for water), and  $\chi_T$  is the (temperature-dependent) isothermal compressibility, which has been determined accurately using speed of sound measurements (18). For liquid water at 23C,  $\chi_T = 4.55 \times 10^{-10} \text{Pa}^{-1}$  and  $d\Sigma/d\Omega = 0.0164 \text{cm}^{-1}$ .

For calibrating X-ray data, the scattering curve for pure water was measured in the sample cell, and the scattering from the empty cell was measured and subtracted. The scattering curves for each macromolecular sample,  $I(q)_{\text{sample}}$ , were scaled by a constant:

$$\left. \frac{d\Sigma(q)}{d\Omega} \right|_{\text{sample}} = \frac{I(q)_{\text{sample}}}{I(q \rightarrow 0)_{\text{water}}} \left. \frac{d\Sigma}{d\Omega} \right|_{\text{water}} \quad (\text{S32})$$

The absolute scattering cross section was converted to units of electron<sup>2</sup> per molecule by dividing by the concentration of molecules (number density  $n$ ) and the square of the classical electron radius  $r_0$ :

$$I(q) [\text{electron}^2] = \frac{1}{n r_0^2} \frac{d\Sigma(q)}{d\Omega} \quad (\text{S33})$$

Equations S32 and S33 can be combined in a convenient form

$$I(q)_{\text{sample}} [\text{electron}^2] = \frac{1000 \text{cm}^3/\text{L}}{N_A c_{\text{sample}}} \cdot \frac{1}{r_0^2} \cdot \frac{I(q)_{\text{sample}}}{I(q \rightarrow 0)_{\text{water}}} \left. \frac{d\Sigma}{d\Omega} \right|_{\text{water}} \quad (\text{S34})$$

where  $c_{\text{sample}}$  is the sample concentration in units of mole/L (as measured from standard techniques like UV-VIS),  $r_0$  and  $d\Sigma/d\Omega$  are in CGS units, and  $N_A$  is Avogadro's number.

### Robust matrix method for decomposition of isomorphous replacement SAXS data

Equation S20 can be written in matrix form as  $I = P \cdot C$ , where the columns of  $I$  contain the scattering intensities, the columns of  $P$  contain the basis functions  $P_M(q)$ ,  $P_{MI}(q)$ , and  $P_I(q)$ , and  $C$  is a matrix of coefficients. Written explicitly,

$$\begin{bmatrix} I^{(1)}(q_1) & I^{(2)}(q_1) & \cdots \\ I^{(1)}(q_2) & I^{(2)}(q_2) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} P_M(q_1) & P_{MI}(q_1) & P_I(q_1) \\ P_M(q_2) & P_{MI}(q_2) & P_I(q_2) \\ \vdots & \vdots & \vdots \end{bmatrix} \cdot \begin{bmatrix} \delta_M^2 & \delta_M^2 & \cdots \\ 2\delta_M N_I \delta_I^{(1)} & 2\delta_M N_I \delta_I^{(2)} & \cdots \\ N_I^2 (\delta_I^{(1)})^2 & N_I^2 (\delta_I^{(2)})^2 & \cdots \end{bmatrix} \quad (\text{S35})$$

where superscripts enumerate the different salt solutions. In the case where the matrix  $I$  contains experimental data, Equation S35 is

$$I_{\text{exp}} = P C + E \quad (\text{S36})$$

where  $E$  is a matrix of errors due to measurement noise. The solution  $P$  that minimizes the least-squares error is

$$P = I_{\text{exp}} C^+ \quad (\text{S37})$$

where  $C^+$  is the Moore-Penrose inverse of  $C$ . However, finding  $P$  in this way will only work with perfectly-calibrated data. The solution can be found robustly if both sides of Equation S35 are normalized by the forward scattering:

$$I_{\text{exp}} I(0)_{\text{exp}}^{-1} = P (C I(0)^{-1}) \quad (\text{S38})$$

where  $I(0)_{\text{exp}}$  is a diagonal matrix containing the zero-angle scattering (determined from the experimental data, e.g. using the indirect Fourier transform),

$$I(0)_{\text{exp}} = \begin{bmatrix} I^{(1)}(0) & 0 & \cdots \\ 0 & I^{(2)}(0) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad (\text{S39})$$

and  $I(0)$  is the prediction for  $I(0)_{exp}$ ,

$$I(0) = \begin{bmatrix} (\delta_M + N_I \delta_I^{(1)})^2 & 0 & \cdots \\ 0 & (\delta_M + N_I \delta_I^{(2)})^2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad (\text{S40})$$

The robust least-squares solution for  $P$  in Equation S38 is

$$P = I_{exp} I(0)_{exp}^{-1} (C I(0)^{-1})^+ \quad (\text{S41})$$

where  $(C I(0)^{-1})^+$  is the Moore-Penrose inverse of  $(C I(0)^{-1})$ , calculated using the PINV function in MATLAB.

## DNA hydration model and X-ray scattering simulation

The spherical harmonic expansion method enables efficient calculation of the orientationally-averaged intensity from a system composed of discrete, spherically-symmetric particles (2). To take advantage of this method, the electron densities of the DNA, its associated solvent molecules, and the excluded solvent were modeled using an explicit particle representation.

The atomic coordinates of a 25 base-pair, B-form DNA duplex were generated by Nucleic Acid Builder (19). A geometric hydration model described previously (20) was implemented in MATLAB: water-like particles were placed at random but non-overlapping positions  $>1.50 \text{ \AA}$  from the DNA's Van-der-Waals surface. A total of 384 (8 per phosphate group) were used. To model the excluded solvent, a pre-computed water box (21) was superimposed on the DNA model, and water molecules outside a cutoff distance from the DNA surface were discarded. This cutoff was chosen to match the experimentally determined  $\delta_M$  (see Results in the Main Text). For averaging over solvent configurations, the excluded volume calculation was repeated for 8 different positions of the DNA within the water box and the particles were assigned spherical Gaussian scattering factors to better approximate uniform density (2, 22).

The ion atmosphere was modeled by solving the nonlinear Poisson-Boltzmann equation (NPBE) numerically on a grid using the program APBS (23). First, the DNA's atomic charges and radii were assigned by PDB2PQR with Amber parameters (24). The APBS program was configured with a  $100 \text{ \AA} \times 100 \text{ \AA} \times 168 \text{ \AA}$  box size,  $129 \times 129 \times 193$  mesh elements, and a zero-potential boundary condition. The excess ion density was computed from the NPBE solution, and discretely sampled using Monte Carlo with 10,000 points (the probability for a point to occupy a cell was proportional to the excess ion density times the cell volume, and within each cell the coordinates were assigned at random with uniform probability).

## Reduced $\chi^2$ statistic for comparing scattering profiles

To provide a quantitative comparison between experimentally-determined terms  $P_{exp}(q)$  and those predicted by models  $P_{calc}(q)$ , we compute the following chi-squared statistic:

$$\chi_{rel}^2 = \frac{1}{\beta^2 N} \sum_{i=1}^N \left( \frac{P_{exp}(q_i) - P_{calc}(q_i)}{\sigma_{exp}(q_i)} \right)^2 \quad (\text{S42})$$

where  $\sigma_{exp}$  is an estimate of the experimental uncertainty and  $\beta$  is a scale factor so that  $\chi_{rel}^2 = 1$  when  $P_{calc}(q)$  is replaced by the regularized version of  $P_{exp}(q)$  (smoothed data in Figures 2 and 3 in the Main Text). Although in the ideal case  $\beta = 1$ , we find  $\beta \sim 0.94$  when  $\sigma_{exp}$  is determined by propagation of uncertainty through Equation S41.

## SUPPORTING REFERENCES

1. Park, S., J. P. Bardhan, B. Roux, and L. Makowski, 2009. Simulated X-ray scattering of protein solutions using explicit-solvent models. *J. Chem. Phys.* 130:134114.
2. Svergun, D., C. Barberato, and M. H. J. Koch, 1995. CRY SOL – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr.* 28:768–773.
3. Wang, J., A. N. Tripathi, and V. H. Smith Jr, 1994. Chemical binding and electron correlation effects in x-ray and high energy electron scattering. *The Journal of chemical physics* 101:4842–4854.
4. Cromer, D. T., and J. B. Mann, 1968. X-ray scattering factors computed from numerical Hartree-Fock wave functions. *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography* 24:321–324.



5. Marcus, Y., 1985. Ion Solvation. Wiley.
6. Qiu, X., L. W. Kwok, H. Y. Park, J. S. Lamb, K. Andresen, and L. Pollack, 2006. Measuring Inter-DNA Potentials in Solution. Phys. Rev. Lett. 96:138101+.
7. Pabit, S. A., S. P. Meisburger, L. Li, J. M. Blose, C. D. Jones, and L. Pollack, 2010. Counting ions around DNA with anomalous small-angle X-ray scattering. J. Am. Chem. Soc. 132:16334–16336.
8. Qiu, X., K. Andresen, J. S. Lamb, L. W. Kwok, and L. Pollack, 2008. Abrupt transition from a free, repulsive to a condensed, attractive DNA phase, induced by multivalent polyamine cations. Phys. Rev. Lett. 101:228101.
9. Andresen, K., R. Das, H. Y. Park, H. Smith, L. W. Kwok, J. S. Lamb, E. Kirkland, D. Herschlag, K. Finkelstein, and L. Pollack, 2004. Spatial distribution of competing ions around DNA in solution. Phys. Rev. Lett. 93:248103.
10. Andresen, K., X. Qiu, S. A. Pabit, J. S. Lamb, H. Y. Park, L. W. Kwok, and L. Pollack, 2008. Mono- and trivalent ions around DNA: a small-angle scattering study of competition and interactions. Biophys. J. 95:287–295.
11. Pabit, S. A., X. Qiu, J. S. Lamb, L. Li, S. P. Meisburger, and 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.
12. Li, L., S. A. Pabit, S. P. Meisburger, and L. Pollack, 2011. Double-stranded RNA resists condensation. Phys. Rev. Lett. 106:108101.
13. Blose, J. M., S. A. Pabit, S. P. Meisburger, L. Li, C. D. Jones, and L. Pollack, 2011. Effects of a Protecting Osmolyte on the Ion Atmosphere Surrounding DNA Duplexes. Biochemistry 50:8540–8547.
14. Das, R., T. Mills, L. Kwok, G. Maskel, I. Millett, S. Doniach, K. Finkelstein, D. Herschlag, and L. Pollack, 2003. Counterion distribution around DNA probed by solution X-ray scattering. Phys. Rev. Lett. 90:188103.
15. Cantor, C. R., M. M. Warshaw, and H. Shapiro, 1970. Oligonucleotide interactions. III. Circular dichroism studies of the conformation of deoxyoligonucleotides. Biopolymers 9:1059–1077.
16. Tataurov, A. V., Y. You, and R. Owczarzy, 2008. Predicting ultraviolet spectrum of single stranded and double stranded deoxyribonucleic acids. Biophys. Chem. 133:66–70.
17. Orthaber, D., A. Bergmann, and O. Glatter, 2000. SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. J. Appl. Crystallogr. 33:218–225.
18. Kell, G. S., 1970. Isothermal compressibility of liquid water at 1 atm. J. Chem. Eng. Data 15:119–122.
19. Macke, T. J., and D. A. Case, 1998. Modeling Unusual Nucleic Acid Structures. In N. B. Leontis, and J. SantaLucia, editors, Molecular Modeling of Nucleic Acids, American Chemical Society, Washington, DC, 379–393.
20. Durchschlag, H., and P. Zipper, 2003. Modeling the hydration of proteins: prediction of structural and hydrodynamic parameters from X-ray diffraction and scattering data. Eur. Biophys. J. 32:487–502.
21. Yang, S., M. Parisien, F. Major, and B. Roux, 2010. RNA structure determination using SAXS data. J. Phys. Chem. B 114:10039–10048.
22. Merzel, F., and J. C. Smith, 2002. SASSIM: a method for calculating small-angle X-ray and neutron scattering and the associated molecular envelope from explicit-atom models of solvated proteins. Acta Crystallogr. D Biol. Crystallogr. 58:242–249.
23. Holst, M., N. Baker, and F. Wang, 2000. Adaptive multilevel finite element solution of the Poisson–Boltzmann equation I. Algorithms and examples. J. Comput. Chem. 21:1319–1342.
24. Dolinsky, T. J., J. E. Nielsen, J. A. McCammon, and N. A. Baker, 2004. PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. Nucleic Acids Res. 32:W665–W667.

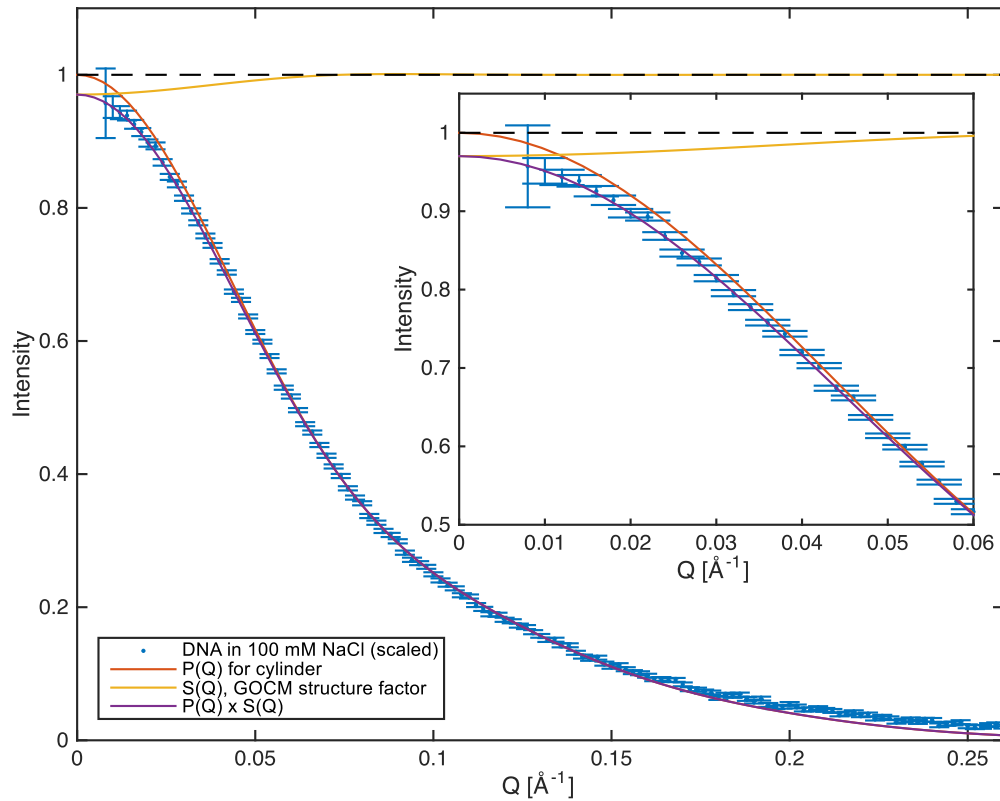


Figure S1: Modeling interparticle interference. The structure factor for a DNA duplex in 100 mM NaCl was calculated using the generalized one component macroion theory (GOCM) and decoupling approximation with a cylindrical form factor (6). Model parameters were determined from SAXS measurements of DNA solutions at high concentration, published previously (6):  $\sigma = 55.0 \text{ \AA}$  (effective particle diameter),  $z_m = 9.1$  (effective charge), and  $I = 100 \text{ mM}$  (ionic strength). The predicted concentration-dependent scattering profile is  $I(q) = P_{CYL}(q, R, H) \times S(q, c, R, H)$ , where  $P(q) = P_{CYL}(q, R, H)$  is the form factor for a cylinder with radius  $R$  and height  $H$ , and  $S(q)$  is the structure factor (note that  $S(q)$  depends on  $R$  and  $H$  through the decoupling approximation). To check for consistency between the structure factor model and the observed data at low concentration, we fit the model to the SAXS data (50  $\mu\text{M}$  DNA in 100 mM NaCl, Figure 1 in the Main Text) on an arbitrary intensity scale over the  $q$ -range  $0.008 < q < 0.1 \text{ \AA}^{-1}$ . The model with parameters  $R = 11.8 \text{ \AA}$  and  $H = 82 \text{ \AA}$  is shown along with the corresponding form factor  $P(q)$  and structure factor  $S(q)$ . According to the model, the data differs from the dilute limit,  $P(q)$ , in the low- $q$  region only (inset), and the magnitude of the interparticle interference effect is at most 3%.

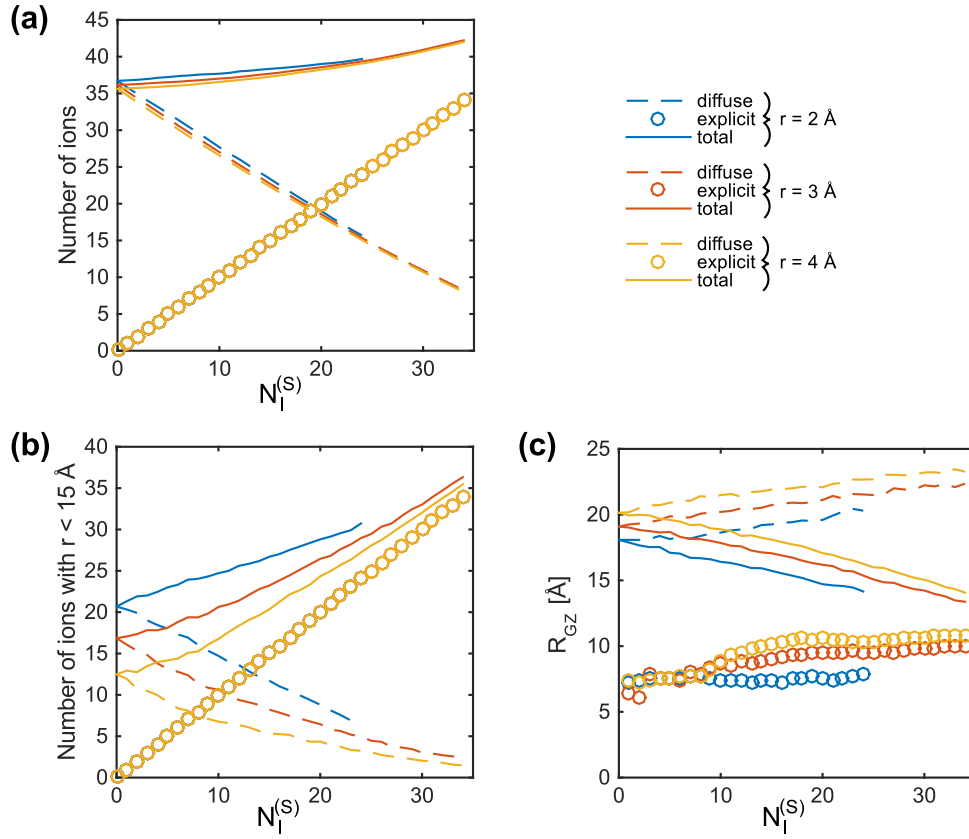


Figure S2: Statistics from ion atmosphere models generated using the explicit / nonlinear Poisson Boltzmann equation (NPBE) approach described in the Main Text. The ion atmosphere has two classes of ions, explicit ions fixed in place near the DNA surface, and diffuse ions modeled by a continuous distribution from the NPBE solution. The total number of ions is the sum of the explicit (“surface”) ions,  $N_I^{(S)}$ , and the number of diffuse ions  $N_I^{(D)}$  (equal to the integral of the diffuse density,  $N_I^{(D)} = \int_V \rho(r) dV$ ). The calculation was repeated for three different values of the Stern layer thickness,  $r = 2, 3, \text{ or } 4$  Å and for  $N_I^{(S)}$  between 0 and 34. (a) As  $N_I^{(S)}$  increases, the  $N_I^{(D)}$  decreases, while the total number  $N_I^{(S)} + N_I^{(D)}$  depends only weakly on  $N_I^{(S)}$ . (b) Although the total number of ions is relatively constant, their distribution is not. To illustrate, we counted the number of diffuse and explicit ions within a cylinder of radius  $r = 15$  Å centered on the DNA molecule. The overall effect of adding explicit ions is to increase the total number of ions near the DNA surface, at the expense of the diffuse component further away. The number of diffuse ions near the surface depends strongly on the choice of Stern radius. (c) The size of the ion distribution for the different models was compared by calculating the radius of gyration of the ions about the Z-axis of the DNA,  $R_{GZ}$ , defined as  $R_{GZ,S}^2 = \left(N_I^{(S)}\right)^{-1} \sum_{i=1}^{N_S} (x_i^2 + y_i^2)$  for the explicit ions, and  $R_{GZ,D}^2 = \left(N_I^{(D)}\right)^{-1} \int_V \rho(x, y, z) (x^2 + y^2) dV$  for the diffuse ions. The value of  $R_{GZ}$  for the total ion atmosphere is  $R_{GZ}^2 = \left(N_I^{(S)} R_{GZ,S}^2 + N_I^{(D)} R_{GZ,D}^2\right) / \left(N_I^{(S)} + N_I^{(D)}\right)$ .

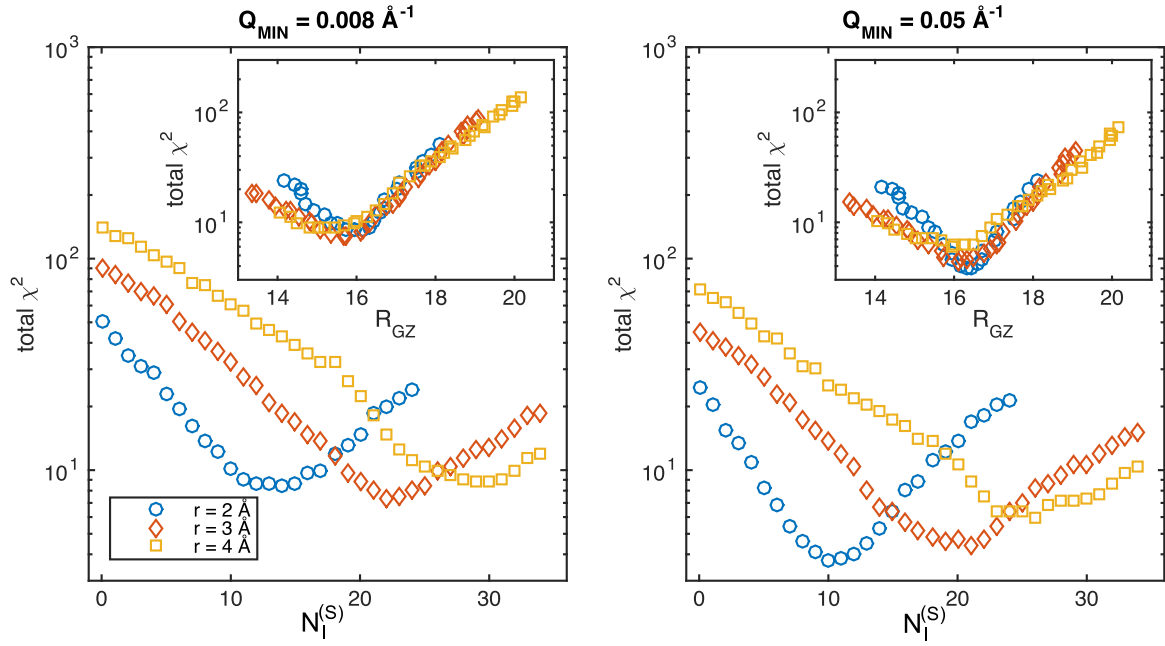


Figure S3: Ad-hoc parameters of the ion atmosphere model and fitting procedure do not affect overall conclusions. (left) Results of fitting the explicit ion / nonlinear Poisson Boltzmann equation (NPBE) hybrid models to the entire heavy atom data set. Results for three Stern layer thickness indicated in the legend are compared (points for  $r = 2 \text{ \AA}$  are identical to Figure 6b in the Main Text). (right) The first were repeated using a truncated  $q$ -range, chosen to minimize the potential effects of residual inter-particle interference (see Figure S1). Truncation improves the overall quality of the fits, but does not alter the interpretation. In both cases, the minimum  $\chi^2$  occurs for  $N_I^{(S)}$  between 0 and 34. The location of the minimum depends on the choice of Stern layer thickness. As discussed in Figure S2, the Stern layer thickness affects the spatial extent of the ion atmosphere at  $N_I^{(S)} = 0$ , and as  $N_I^{(S)}$  increases the average ion atmosphere size is reduced. Since SAXS data is low resolution, it is reasonable to guess that the primary feature of the ion atmosphere determining  $\chi^2$  is the rotational moment about the Z-axis ( $R_{GZ}$  in Figure S2c), and not the fine details of the ion placement. Indeed, when  $\chi^2$  is plotted as a function of  $R_{GZ}$  rather than  $N_I^{(S)}$  (insets), models with different Stern radii collapse to a single curve with a minimum around  $R_{GZ} \sim 16 \text{ \AA}$  (full  $q$ -range) or  $R_{GZ} \sim 16.4 \text{ \AA}$  (truncated  $q$ -range). The best-fit value is significantly more compact than the Poisson-Boltzmann (PB) solution ( $R_{GZ} \sim 18 \text{ \AA}$  with  $r = 2 \text{ \AA}$ ). Thus, our conclusion that PB overestimates the spatial extent of the ion atmosphere does not depend the details of the calculation, i.e. the Stern radius value or the exclusion of low-angle data.