Invited Review

Polyelectrolyte Properties of Single Stranded DNA Measured Using SAXS and Single-Molecule FRET: Beyond the Wormlike Chain Model

Steve P. Meisburger,[1] Julie L. Sutton,[1] Huimin Chen,[1] Suzette A. Pabit,[1] Serdal Kirmizialtin,[2,3] Ron Elber,[2,3] Lois Pollack[1]

[1] *School of Applied and Engineering Physics, Cornell University, Ithaca, NY*

[2] *Department of Chemistry and Biochemistry, University of Texas at Austin, Austin, TX*

[3] *Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, TX*

Received 14 February 2013; accepted 11 April 2013
Published online 20 April 2013 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/bip.22265

ABSTRACT:

Nucleic acids are highly charged polyelectrolytes that interact strongly with salt ions. Rigid, base-paired regions are successfully described with wormlike chain models, but nonbase-paired single stranded regions have fundamentally different polymer properties because of their greater flexibility. Recently, attention has turned to single stranded nucleic acids due to the growing recognition of their biological importance, as well as the availability of sophisticated experimental techniques sensitive to the conformation of individual molecules. We investigate polyelectrolyte properties of poly(dT), an important and widely studied model system for flexible single stranded nucleic acids, in physiologically important mixed mono- and divalent salt. We report measurements of the form factor and interparticle interactions using SAXS, end-to-end distances using smFRET, and number of excess ions using ASAXS. We present a coarse-grained model that accounts for flexibility, excluded volume, and electrostatic interactions in these systems. Predictions of the model are validated against experiment. We also discuss the state of all-atom, explicit solvent molecular dynamics simulations of poly(dT), the next step in understanding the complexities of ion interactions with these highly charged and flexible polymers. © 2013 Wiley Periodicals, Inc. Biopolymers 99: 1032–1045, 2013.

Keywords: *single stranded DNA; ion–DNA interactions; small angle X-ray scattering*

Correspondence to: Lois Pollack; e-mail: lp26@cornell.edu
Contract grant sponsor: NIH
Contract grant number: R01-GM085062
Contract grant sponsor: NSF
Contract grant number: DMR-0936384
Contract grant sponsor: National Institute of General Medical Services (NIGMS)
Contract grant number: GM-103485

This article was originally published online as an accepted preprint. The "Published Online" date corresponds to the preprint version. You can request a copy of the preprint by emailing the Biopolymers editorial office at biopolymers@wiley.com

INTRODUCTION

The growing appreciation for the roles that nucleic acids play in biology calls for a thorough description of these biopolymers, including an understanding of how their mechanical properties couple to their biological function. Much of the effort thus far has focused on double stranded structures, which are well described by wormlike chain (WLC) models with ionic strength dependent persistence lengths that exceed 100 base pairs.[1] However, experience with other biopolymers, like proteins, demonstrates that although rigid structures are most amenable to experimental characterization, the flexible regions often impart biological function.[2] The most flexible regions of nucleic acids are nonbase-paired and include single stranded DNA (ssDNA) and RNA (ssRNA) regions that

are involved in crucial biological processes. For example, polymerases unwind dsDNA, yielding stretches of ssDNA whose genetic information is transcribed into messenger ssRNA. The nonbase-paired regions of ssRNA may be recognized by proteins involved in gene regulation or transport. The mechanical properties of ssRNA are exploited by riboswitches, where single stranded regions serve as actuators.[3] Finally, ssDNA is a tool in bioengineering, used for example as a tunable ligand for building nanoparticle superlattices.[4]

Although the WLC model (and associated polyelectrolyte theory) has been successful in describing dsDNA, biophysical studies of single stranded nucleic acids in the last decade have found varying degrees of success applying WLC models. Estimates of the persistence lengths and contour lengths in ssDNA and ssRNA vary widely among different experimental techniques, which have included fluorescence-based measurements,[5–8] single-molecule force extension,[9–11] and small angle x-ray scattering (SAXS).[5,12] The polyelectrolyte theory describing electrostatic effects on polymer flexibility predicts an electrostatic component of the persistence length that has a power law dependence on the Debye screening length, where the exponent is different depending on assumptions about the intrinsic flexibility.[13] Two limiting theories, that of Odijk, Skolnick, and Fixman (OSF)[14,15] and Barrat and Joanny (BJ),[16] predict exponents of 2 and 1, respectively. Experiments on dsDNA agree with OSF,[1] but there is no consensus on whether the many experiments testing ssRNA and ssDNA fit either theory (reviewed in Ref. 12).

To this end, we recently performed both SAXS and single-molecule Förster resonance energy transfer (smFRET) measurements of homopolymeric deoxythymidylate (poly(dT)) and uridylate (poly(rU)) molecules in solution and constrained a WLC model to simultaneously fit both the end-to-end distance measured by FRET and the entire scattering profile measured by SAXS.[5] SmFRET measurements over a wide range of monovalent and divalent salt concentrations were interpreted in this context. Surprisingly, we found that the power-law dependence of persistence length predicted by polyelectrolyte theory did not apply over the entire range of salt concentration. Furthermore, the power law exponents fall between the OSF and BJ values and vary depending on the sugar moiety (ribose vs. deoxyribose) and the identity of the counterion (Mg^{2+} vs Na^{+}). However, the smFRET data also hinted at a possible reason for the discrepancy: divalent ions show an anomalously strong effect on structure, which suggests that ion condensation plays a major role in the conformations of ssDNA. The theories of OSF and BJ, which are based on the Debye–Hückel approximation for electrostatics, do not account for this phenomenon.

Another shortcoming of the WLC model is its neglect of excluded volume interactions. For dsDNA, the diameter (~20 Å) is much smaller than the persistence length (~500 Å), but for single stranded nucleic acids they are comparable (~5–10 Å). Thus, excluded volume might be safely neglected for dsDNA, but for ssDNA it is potentially important, even for short chains. Indeed, excluded volume effects have been observed in single-molecule diffusion[6] and SAXS[12] studies of 8–128 nucleotide poly(dT), and in single-molecule force spectroscopy measurements[10] of comparatively long (~10 kb) ssDNA. The picture of ssDNA as an electrostatically swollen coil, rather than a WLC, more strongly resembles recent models for highly charged intrinsically disordered proteins (IDPs) that invoke Flory theory,[17] where excluded volume and solvent character play a dominant role. In one study, FRET measurements of IDP dimensions in an ionic denaturant were modeled analytically with chain monomers possessing a salt-dependent effective volume.[18] These results were consistent with a computational study using implicit solvent[19] that predicted expansion behavior of IDPs in accordance with hydrodynamic measurements. The success these approaches for charged IDPs suggests that similar models might be used for ssDNA.

To develop and test such a model, we carried out SAXS, single-molecule FRET, and anomalous SAXS (ASAXS) measurements of 30 nucleotide poly(dT) (dT_{30}) in physiologically important mixed-salt solutions, where mono- and divalent ions compete. We begin with a coarse-grained representation of ssDNA as a freely rotating chain of virtual bonds between the backbone C4' and P atoms,[20,21] and add hard-core excluded volume interactions between nonbonded monomers and mean-field electrostatics. The precise nature of ssDNA–ion interactions is unknown, and thus our model includes a single parameter for charge renormalization by strongly condensed ions. We determine this parameter for each salt condition by measuring the intermolecular pair potential of dT_{30} molecules in solution using SAXS.[22] With our model constrained by the renormalized charge, we compute ensemble properties such as the radius of gyration, form factor, end-to-end distances, and ion distributions. These predictions are compared with corresponding measurements.

Like the WLC model and polyelectrolyte theory, our model involves coarse-graining and mean-field approximations. Ideally, single stranded nucleic acids and associated salt ions could be modeled without such compromises using explicit solvent molecular dynamics (MD). Recent MD predictions of the complex ion atmosphere around dsRNA have been directly compared with SAXS and ASAXS data.[23,24] However, the large number of polymer degrees of freedom for single stranded nucleic acids poses a computational challenge that currently limits the technique to relatively short chains. Therefore, we also report explicit solvent MD simulations of dT_{10} with monovalent ions, and discuss the future applications of

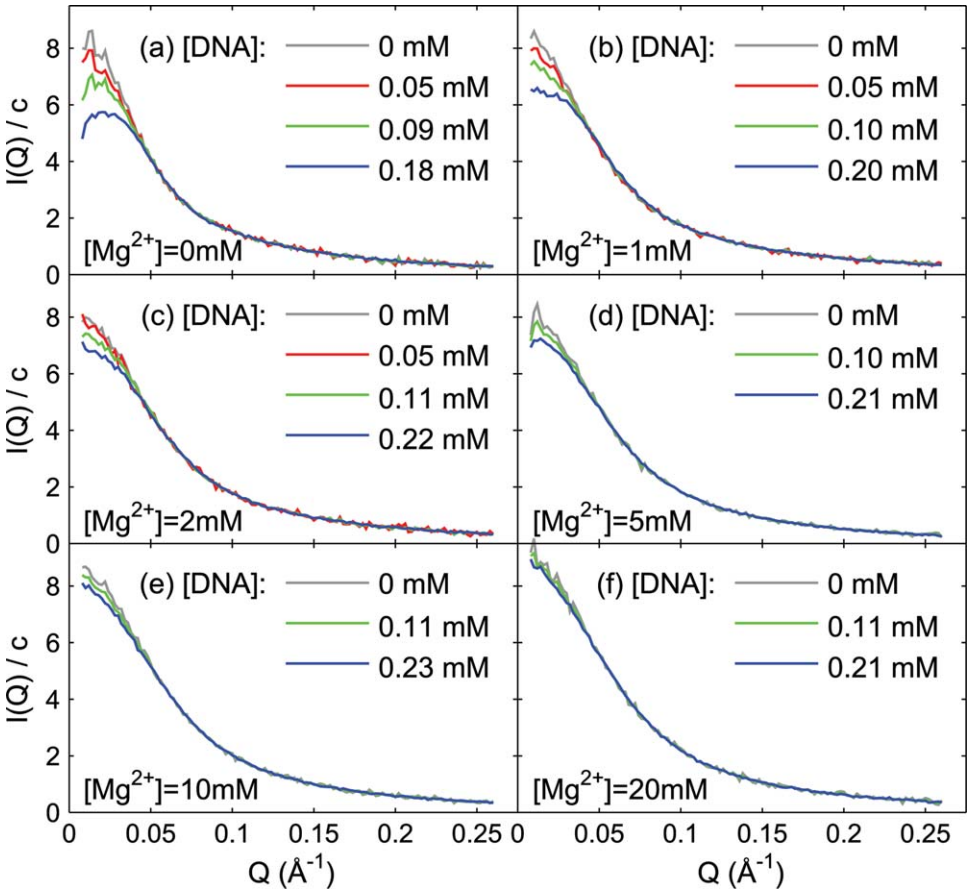

FIGURE 1 Concentration-normalized SAXS profiles of dT_{30} in 20 mM NaCl with added $MgCl_2$ and extrapolated form factors ([DNA]: 0 mM) in gray.

molecular dynamics in conjunction with ion-selective experimental techniques (such as ASAXS) for obtaining a more complete understanding of the interactions between ions and single stranded nucleic acids. Progress in this area will enable accurate modeling of folding and dynamics of these biologically important macromolecules.

RESULTS

SAXS and smFRET Measurements of dT_{30} Compaction and Charge Screening in Mixed Na^+ and Mg^{2+} Salt

SAXS measurements were performed on dT_{30} in 20 mM NaCl and $MgCl_2$ concentrations of 0–20 mM, as described in Methods. For each solution condition, a series of DNA concentrations was measured to investigate the interparticle interference effect, where variation in the shape of the SAXS curve with concentration arises because of interparticle interactions. Measurement of the interparticle interference serves a dual purpose: first for finding the form factor of dT_{30} by extrapolation to infinite dilution, and second for quantifying the strength of interactions between molecules.[25] Concentration-normalized SAXS profiles are plotted in Figure 1. In 20 mM NaCl (Figure 1a), the data show strong variation with DNA concentration for $q < 0.05$ Å^{-1}, consistent with repulsive interactions between molecules.[26] As Mg^{2+} content increases (Figures 1b–1f), the net interparticle interaction weakens significantly but remains repulsive up to 20 mM $MgCl_2$.

The interparticle interference effect was quantified using an analysis of the SAXS data in terms of pair-wise interactions. In dilute solutions where the interference effect is small, the modulation of the scattering profile is linear in concentration and proportional to the second virial coefficient, B_2. This model was fit to the data to obtain B_2 for each solution condition, as described in Methods. The results are plotted in Figure 2. A sharp decrease is observed between 0 and 2 mM $MgCl_2$, followed by a steady decline between 2 and 20 mM $MgCl_2$.

Next, the extrapolated form factors were compared to investigate the effect of Mg^{2+} on the chain conformations. The changes appear subtle in Figure 1, but can be seen more clearly in a real-space analysis. The radius of gyration R_G and maximum dimension D_{MAX} were computed using a Bayesian

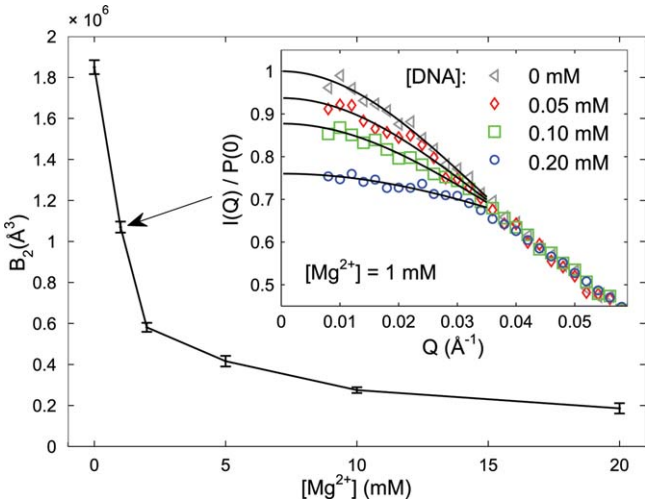

FIGURE 2 The second virial coefficient, B_2, was obtained by double-extrapolation of concentration-normalized SAXS profiles in Figure 1 to $q = 0$ and $c = 0$, as described in Methods. The inset shows this extrapolation for dT_{30} in 20 mM NaCl and 1 mM $MgCl_2$.

indirect Fourier transform[27] implemented in MATLAB, and are given in Table I. As can be seen from plots of R_G and D_{MAX} against Mg^{2+} concentration in Figure 3a, the overall chain dimensions decrease as Mg^{2+} is added.

As a complementary probe of chain dimensions, smFRET measurements were performed on end-labeled dT_{30} in 20 mM NaCl with added $MgCl_2$, and the FRET-averaged end-to-end distance, $<R>_{FRET}$, was determined from the FRET efficiency and the Förster radius as described in Methods. The results are given in Table I, and plotted as a function of Mg^{2+} concentration in Figure 3b. An increase in Mg^{2+} concentration is accompanied by a decrease in $<R>_{FRET}$. As was observed with SAXS, Mg^{2+} has the effect of decreasing overall chain dimensions.

The Number of Excess Monovalent Ions Around dT_{30}

The number of monovalent ions around dT_{30} in excess of the bulk concentration was measured using two different SAXS techniques: heavy ion replacement[28] and ASAXS.[29] While these techniques have been applied to double stranded RNA and DNA, ion populations for single stranded nucleic acids have not been previously reported. Both techniques obtain information about the ion atmosphere by varying the scattering contrast of the ions, either by comparing high-Z and low-Z atomic species (heavy ion method), or by exploiting the energy-dependent scattering factor for the ion of interest near an X-ray absorption edge (ASAXS). This information can include the number of ions per macromolecule when the scattering patterns are placed on an absolute scale. For the heavy ion method, we compare the scattering from two identically prepared solutions of 0.1 mM dT_{30} in 100 mM RbCl or 100

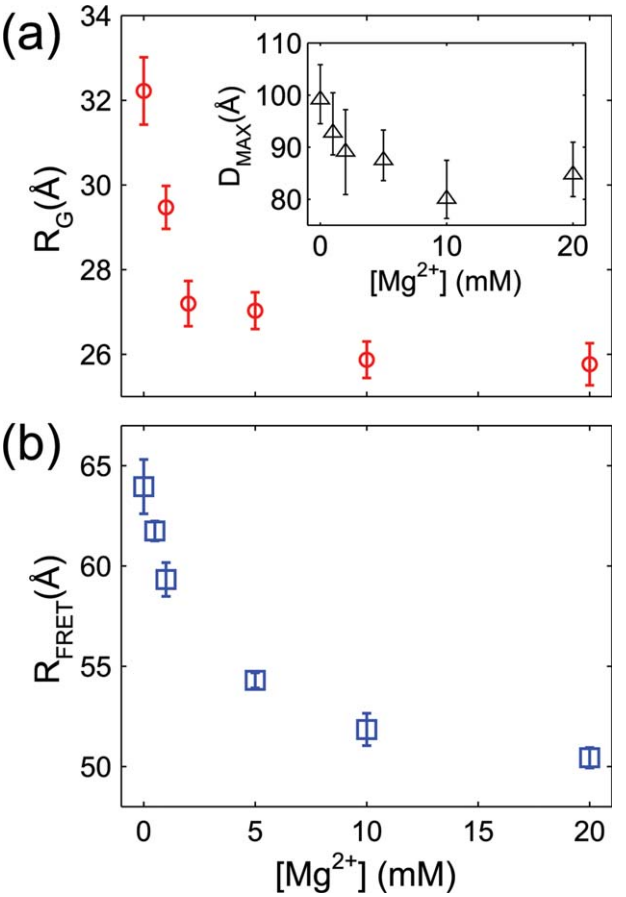

FIGURE 3 Chain dimensions of dT_{30} measured by SAXS and smFRET in 20 mM NaCl with added $MgCl_2$. (a) The radius of gyration R_g and the maximum dimension D_{max} were determined from SAXS data by Bayesian indirect Fourier transform of the extrapolated form factors shown in Figure 1. Error bars correspond to the uncertainty of the fit. (b) The FRET-averaged end-to-end distance, R_{FRET}, was measured with smFRET of freely diffusing, fluorescently labeled dT_{30} in buffer containing 20 mM NaCl and 0, 0.5, 1, 5, 10, and 20 mM $MgCl_2$. Error bars show the standard deviation of at least four repeated measurements.

Table I Experimental Data in Figs. 2 and 3, for dT_{30} with 20 mM NaCl and Varying $MgCl_2$

$[Mg^{2+}]$ (mM)	R_{FRET} (Å)	R_G (Å)	D_{MAX} (Å)	$B_2(\times 10^6$ Å)
0.0	64.0 ± 1.4	32.2 ± 0.8	99 ± 6	1.850 ± 0.034
0.5	61.7 ± 0.5	n.d.	n.d.	n.d.
1.0	59.3 ± 0.8	29.5 ± 0.5	93 ± 6	1.070 ± 0.027
2.0	n.d.	27.2 ± 0.5	89 ± 8	0.581 ± 0.022
5.0	54.3 ± 0.4	27.0 ± 0.4	88 ± 5	0.416 ± 0.026
10.0	51.8 ± 0.8	25.9 ± 0.4	80 ± 6	0.275 ± 0.014
20.0	50.4 ± 0.5	25.8 ± 0.5	85 ± 5	0.186 ± 0.025

n.d. = not determined.

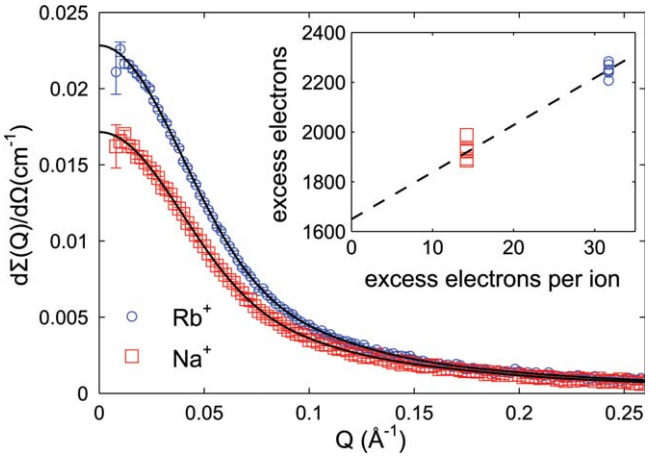

FIGURE 4 Quantitative heavy atom replacement measurement of the ion atmosphere around dT_{30}. The macroscopic scattering cross section of dT_{30} increases when electron-rich Rb^+ ions comprise the ion atmosphere, as opposed to Na^+. The number of excess monovalent ions was determined as shown in the inset, from a plot of total scattering contrast vs. the contrast per ion.

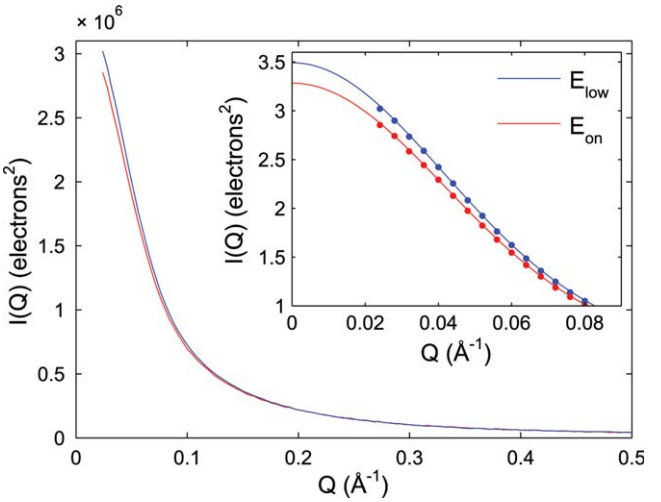

FIGURE 5 Quantitative ASAXS measurements to probe the monovalent ion atmosphere around dT_{30}. The scattering contrast of Rb^+ ions around dT_{30} was varied by tuning the X-ray energy between E_{on} and E_{lo}, defined in the text. The number of ions was determined by the change in the forward scattering with extrapolation to $q = 0$, as shown in the inset.

mM NaCl. ASAXS measurements were performed on dT_{30} with 100 RbCl at two energies near the K-edge of Rb, as described in methods.

For the heavy ion experiments, the macroscopic scattering cross-section, $d\Sigma(q)/d\Omega$, depends on the contrast of the ions at $q = 0$ as,

$$\frac{d\sum(0)}{d\Omega} = nr_0^2\left(\Delta N^e_{\text{total}}\right)^2 = nr_0^2\left(\Delta N^e_{\text{DNA}} + N_{\text{Ions}} \times \Delta N^e_{\text{Ion}}\right)^2 \quad (1)$$

where r_0 is the classical electron radius, n is the concentration (N/V) of DNA molecules, and $\Delta N^e_{\text{total}}$ is the total excess electron density contrast, with contributions from DNA and ion components: ΔN^e_{DNA} is the number of excess electrons for DNA, and ΔN^e_{Ion} is the number of excess electrons per ion, which we calculate using

$$\Delta N^e_{\text{Ion}} = N^e_{\text{Ion}} - \rho_w V$$

where ρ_w is the electron density of water and V is the absolute limiting partial molar volume of the ion ($V_{Na} = -12.6$ Å^3 and $V_{Rb} = 12.8$ Å^3)[30]. The number of excess ions around dT_{30} was determined from a plot of $\Delta N^e_{\text{total}}$ vs. ΔN^e_{Ion}, shown in Figure 4. The slope yields $N_{\text{Ions}} = 18.9 \pm 1.1$.

The ASAXS technique for measuring ion numbers takes advantage of the anomalous scattering properties of Rb ions.[29] The real part of the scattering factor of Rb, f'_{Rb}, changes by several electrons as the X-ray energy is varied below the K-edge at 15.2 keV. Calibrated SAXS profiles from a solution of dT_{30} in 100 mM RbCl are plotted in Figure 5, and show a decrease in forward scattering near the Rb edge due to the lower contrast of Rb at that energy. The number of Rb ions was calculated as described in Methods. With ASAXS, we find 19.1 ± 0.9 excess Rb ions around dT_{30}.

A Charged, Freely Rotating Chain Model for ssDNA

The DNA backbone has six bonds per nucleotide that adopt many orientations. For efficient computational sampling of conformations, it is desirable to reduce the number of degrees of freedom by coarse-graining. The WLC model is one way of achieving this, but its continuously deformable nature makes it a more useful analytic tool than a computational one. WLCs are members of a family of inextensible statistical chains that include freely rotating and freely jointed chains (FRC and FJC, respectively) that are statistically equivalent in the limit of large number of segments (or with contour length much longer than persistence length). Therefore, an FRC model for ssDNA is developed in analogy to the virtual bond description of polynucleotide backbones, which reduces the number of bonds per nucleotide to two. For simplicity, the virtual bond segments are assigned equal length l_0 and valence angle θ, while the dihedral angles are free to rotate. The synthetic poly(dT) molecules modeled here lack 5′ phosphate groups, so a chain with N bases has N-1 phosphates and $2(N-1)$ virtual bonds. Allowable conformations are restricted by a hard-core excluded volume interaction, such that nonbonded atoms have a minimum allowable separation, d.

Polyelectrolytes such as ssDNA swell in response to ionic environments. To model electrostatic effects, charges (with

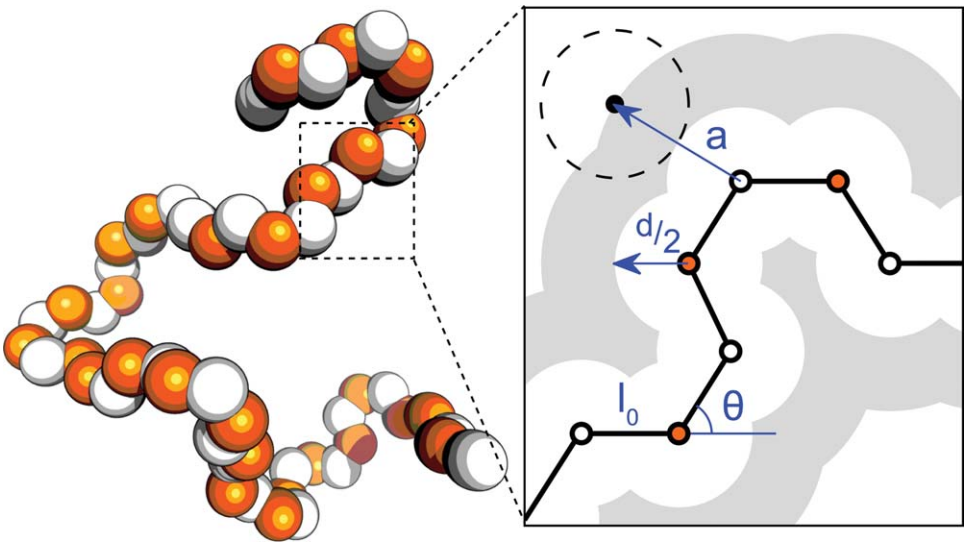

FIGURE 6 Illustration of geometric parameters for the charged, freely rotating chain model of ssDNA. Virtual bonds of length l_0 span backbone C4' and P atoms (open and orange filled circles, respectively). The valence angle between bonds is θ and d is the chain diameter. For electrostatics calculations, each P position is assigned a negative charge and the distance a defines ion accessibility. A representative ion is drawn as a dashed circle, and the ion-inaccessible region is shaded in gray. The cartoon drawing on the left shows a virtual bond model for dT_{30} rendered using Pymol version 1.2r1 (DeLano Scientific LLC).

valence $Z=-1$) are placed at the P-coordinates along the chain. The parameter a is the closest approach for a hydrated ion and (virtual) chain atom, defining the boundary between the macromolecule and the electrolyte. The geometric parameters l_0, θ, a, and d are illustrated in Figure 6.

Electrostatic effects are modeled using the Poisson–Boltzmann (PB) equation as a starting point. Outside the macromolecule, where the solvent has a uniform dielectric constant $\varepsilon=78.5$ and an electrolyte with m species of ion with valence z_i and bulk density n_i^{∞}, the electrostatic potential satisfies the PB equation,

$$\nabla^2\Phi(\vec{r})=-\frac{1}{\varepsilon\varepsilon_0}\sum_{i=1}^{m}e_0 z_i n_i^{\infty}\exp\left(-z_i e_0\Phi(\vec{r})/k_BT\right) \qquad (2)$$

where ε_0 is the permittivity of vacuum, and e_0 is the elementary charge. The linearized form of PB, also called the Debye–Hückel (DH) equation, is

$$\nabla^2\Phi(\vec{r})=\kappa^2\Phi(\vec{r}) \qquad (3)$$

where $\kappa=(8\pi l_B A_0 I)^{1/2}$ is the inverse Debye length with $A_0=6.022\times10^{-4}$ M^{-1} Å^{-3}, $l_B=e_0^2/(4\pi\varepsilon_0\varepsilon k_BT)$ is the Bjerrum length, equal to 7.14 Å for water at room temperature, and I is the ionic strength.

The DH equation is an accurate approximation of PB when the potential is weak ($|\Phi|<k_BT/e_0$), far from the macromolecule. However, the magnitude of the solution to the DH equation can be inaccurate because it depends on the application of Gauss' law at the molecular surface, where the potential is in general much larger than k_BT/e_0. A common way to "rescue" the DH solution is to renormalize the charge of the macromolecule in the DH solution so that it approximates the true potential at large distances.[31] The charge renormalization factor, $f=Z_{eff}/Z$ can be derived by solving the PB and DH equations for a model system (e.g., Ref. 32) or left as an experimentally determined parameter (e.g., Ref. 22). Here, we choose the latter option.

Using the linearity of the DH equation, the electrostatic free energy for the chain is the sum over all pair-wise electrostatic interactions:

$$W_{\text{chain}}=f^2\frac{1}{2}\sum_i\sum_{j\neq i}W_2\left(\left|\left(\vec{r}_i-\vec{r}_j\right)\right|\right) \qquad (4)$$

where $W_2(r)$ is the free energy for a pair of phosphates at a distance r apart, and the summations run over all P coordinates in the chain. We approximate $W_2(r)$ using the repulsive DLVO potential[33] for charged spheres with diameter d:

$$\frac{W_{\text{DLVO}}(r)}{k_BT}=l_B\left(\frac{Z}{1-\kappa d/2}\right)^2\frac{\exp\left(-\kappa(r-d)\right)}{r} \qquad (5)$$

Table II Poly(dT) Models from the Protein Data Bank (PDB) Analyzed here

PPB ID	Chain	Residues	Resolution	Structure Context
3VDY	F	1–10	2.8 A	B. Subtilis, ssDNA-binding protein B[34]
2VW9	C	18–26	2.3 A	H. Pylori, ssDNA-binding protein[35]
4GOP	K	1–25	3.1 A	Eukaryotic Replication protein A[36]
1XHZ	G	1–5	2.7 A	Phi29 DNA polymerase[37]

Insight Into Model Parameters From Crystal Structures of Poly(dT)

To choose physically realistic geometric parameters for the model, we examined X-ray crystal structures deposited in the protein data bank (PDB) that contain single stranded poly(dT) bound to proteins. Continuous strands of poly(dT) resolved in four crystal structures (Table II) were analyzed using the virtual bond representation. Histograms for the virtual bond length and valence angle are shown in Figures 7a and 7b. Bond lengths l_0 ranged from 2.9 to 4 Å, while the valence angle θ spanned 10°–110°. The average values $l_0 = 3.69$ and $\theta = 57.8°$ were chosen for the model.

The minimum distance between two chain atoms, d, should be comparable to the Van der Waals diameter of the sugar phosphate backbone, in the range of roughly 5.4–6.0 Å. The distance of closest approach for ions and DNA, a, is not straightforward to determine because of hydration effects. Recently, molecular dynamics simulations of ions around an RNA duplex with explicit solvent[23] revealed two binding modes for Na^+ characterized by ion-O2P distances of 2.3 and 4.5 Å, while hydrated Mg^{2+} approaches O2P at an intermediate distance of ~4 Å. Adding 1.5 Å for P-O2P bond, the distance a should fall in the range of 3.8–6.0 Å. For simplicity, ions and chain atoms were treated equivalently, with the choice of $a = d = 5.6$ Å.

To calculate accurate SAXS profiles, particularly at high-q, it is important to account for finite thickness of the chain. As in previous SAXS studies of poly(dT), the X-ray scattering pattern is calculated using

$$I(q) = P_{\text{chain}}(q) \times I_{\text{CS}}(q) \tag{6}$$

where $I_{CS}(q)$ accounts for the finite thickness. In earlier measurements of poly(dT),[5] a cylinder model was used for $I_{CS}(q)$

$$I_{\text{CYL}}(q, R) = (2J_1(qR)/(qR))^2 \tag{7}$$

with the radius R as a fitting parameter. Here, we improve on this method by calculating an average $I_{CS}(q)$ directly from the crystal structures listed in Table II. The chains were divided into 4-nucleotide segments, and the scattering patterns were computed using CRYSOL[38] with default hydration parameters. Then, $I_{CS}(q)$ was found by dividing the scattering profile by $P_{\text{chain}}(q)$ calculated from the virtual bond representation of the same structure, as described in Methods. The profiles and their average are shown in Figure 7c. The average $I_{CS}(q)$ resembles $I_{CYL}(q, R)$ used previously, however the fit can be made essentially exact for $q < 0.3$ Å^{-1} by adding a constant: $I_{CS}(q) \approx (1 - c) \times I_{CYL}(q, R) + c$ with $R = 6.812$ Å and $c = 0.2838$.

To verify that the above choice of model parameters gives a reasonable approximation of the intrinsic flexibility and excluded volume of poly(dT), the mean squared radius of

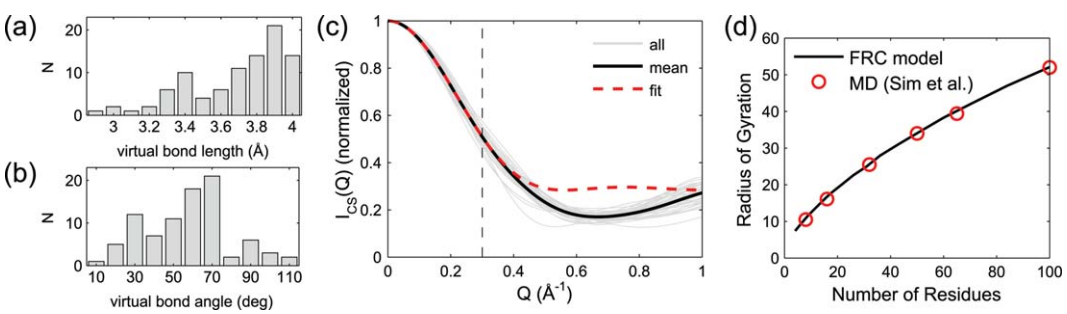

FIGURE 7 Virtual bond analysis of atomic models of poly(dT). (a,b) Histograms of bond lengths, l_0 and angles θ were generated from crystallographic models described in Table II. The average virtual bond parameters are $l_0 = 3.69$ Å and $\theta = 57.8°$. (c) Effective scattering of the cross-section $I_{CS}(q)$ was calculated from four-nucleotide subsets of the atomic models as described in the text. The solid line shows the average, and the dashed line shows a convenient fit in the experimental q range ($0 < q < 0.3$ Å^{-1}) (see text). (d) The radius of gyration was computed for virtual bond chains with varying numbers of residues using the average parameters from (a) and (b), and a chain diameter $d = 5.6$ Å. The results (solid line) are plotted with "sterics only" molecular dynamics simulation results (open circles) reported in Ref. 12.

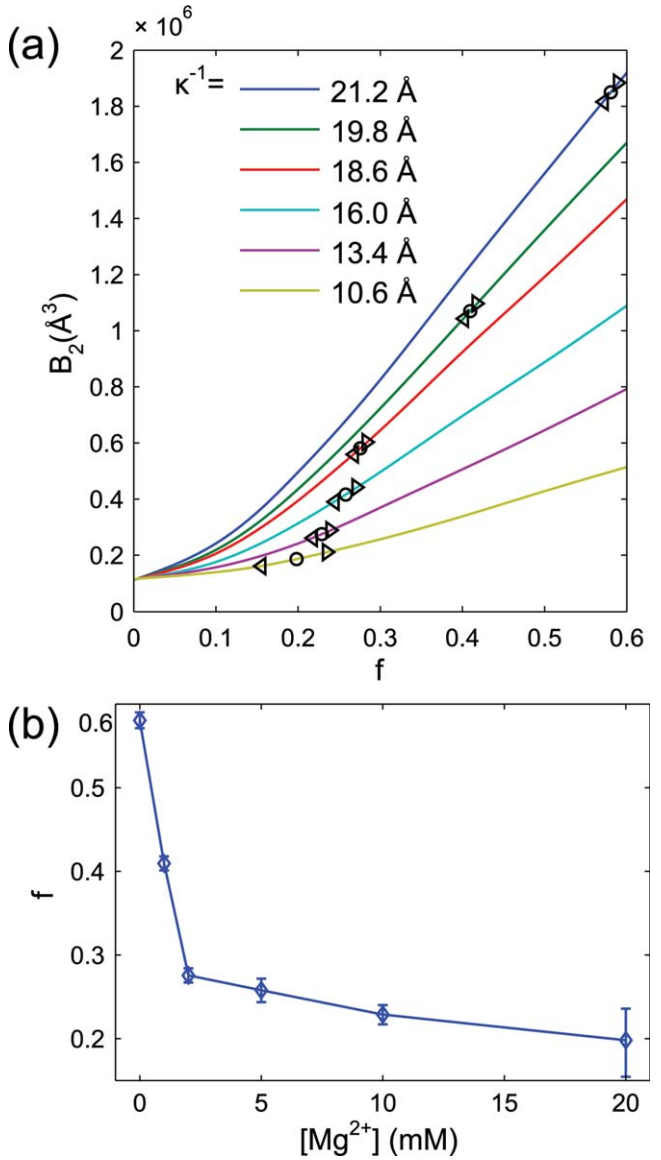

FIGURE 8 Relationship between the second virial coefficient B_2 and the charge renormalization parameter f for the FRC model of dT_{30}. (a) B_2 vs. f was calculated for values of the Debye screening length κ^{-1} corresponding to 20 mM NaCl plus 0, 1, 2, 5, 10, or 20 mM $MgCl_2$. Experimental values of B_2 (circles) ± standard errors (triangles) were mapped onto f at each salt condition. (b) Plot of f vs. Mg^{2+} concentration obtained from (a).

gyration for ensembles of chains was compared with "sterics only" molecular dynamics calculations recently reported for poly(dT), as the number of nucleotides was varied from $N = 8$ to 100,[12] shown in Figure 7d. Agreement between the coarse-grained and all-atom representations is excellent. Furthermore, we find that the radius of gyration follows a polymer scaling law,[39] $R_g \propto N^{\nu}$ with $\nu = 0.62$, which matches the reported value from an earlier MD calculation of poly(dT) for $N = 8$–128 nucleotides.[6]

Determination of the Charge Renormalization Parameter From Measurements of B_2

With the geometric parameters for the poly(dT) model fully specified, we turn to the electrostatic parameters. The Debye screening length κ^{-1} is determined solely by the ionic strength of the solution, but f must be fixed experimentally for each ionic condition by matching B_2. Therefore, for each experimental value of κ, and for a range of f values, B_2 was computed as described in Materials and Methods. Briefly, a set of chains, $\Omega_{\kappa,f}$, was generated using the Metropolis algorithm[40] with the energy function W_{chain} for every (κ, f) pair of parameters. Next, pairs of chains were drawn from $\Omega_{\kappa,f}$ and their centers of mass were displaced from each other by a random vector $\boldsymbol{R}$, chosen with uniform probability density within a sphere of radius $R_{\max}$. If the chains clashed, the interaction potential was assigned $W_m = \infty$. Otherwise, W_m was found in an analogous way to the single chain energy in Eqs. (4) and (5).

After M iterations, the second virial coefficient was computed,

$$B_2 = \frac{1}{2} V_{\text{excl}} = \frac{V_{\max}}{2M} \sum_{m=1}^{M} (1 - \exp(-W_m)) \tag{8}$$

where $V_{\max} = (4/3)\pi R_{\max}^3$. The value of $R_{\max}$ must be large enough that chains have negligible interactions at that distance. Therefore, B_2 was calculated with increasing values of $R_{\max}$ until no significant change was observed.

The result of the calculation is a set of curves for B_2 vs. f, shown in Figure 8a. For each salt condition, experimental values of B_2 (± standard errors) were mapped onto the parameter f. Results are plotted vs. Mg concentration in Figure 8b and Table III.

Prediction of Chain Conformations

With the value of f constrained by the measurement of B_2 at each salt condition, it is now possible to predict other observables using the model, such as the radius of gyration,

Table III Model Parameters for dT_{30} and Predicted Ensemble Average Dimensions

$[Mg^{2+}]$ (mM)	κ^{-1} (Å)	f	$\langle R_{G0}^2 \rangle^{1/2}$ (Å)	$\langle R_G^2 \rangle^{1/2}$ (Å)	$\langle R^2 \rangle^{1/2}$ (Å)	$\langle R \rangle_{\text{FRET}}$ (Å)
0	21.2	0.581	30.0	30.4	83.4	71.2
1	19.8	0.409	27.4	27.8	74.3	64.7
2	18.6	0.276	25.6	26.1	68.1	60.6
5	16.0	0.258	25.3	25.8	66.8	59.8
10	13.4	0.229	24.8	25.4	65.3	58.8
20	10.6	0.198	24.4	25.0	64.1	58.0

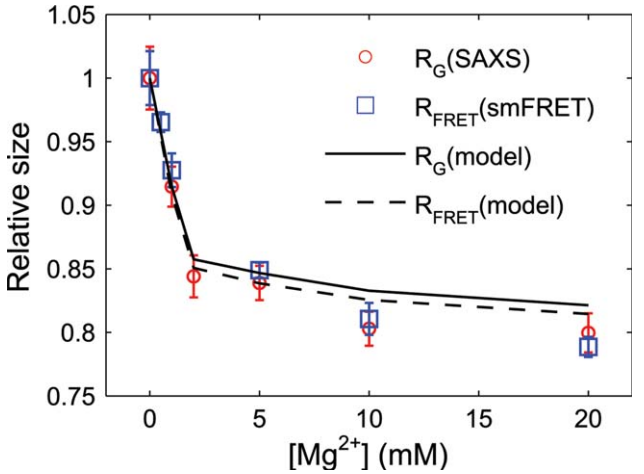

FIGURE 9 Chain compaction measurements by SAXS and smFRET are compared with model predictions. The ratio between R_g (or R_{FRET}) in 20 mM NaCl with added $MgCl_2$ and the value in 20 mM NaCl alone is plotted as a function of added Mg^{2+}.

the end-to-end distance distribution, the scattering profile, and the interparticle interference function. These predictions are compared directly with experimental data.

At each condition, a set of chain conformations was generated by Monte Carlo sampling. The end-to-end distance histogram was converted to $<R>_{FRET}$ as described previously.[5] The average pair distance distribution function was used to calculate the mean square radius of gyration of the chain,

$$\langle R_{g0}^2 \rangle = \frac{1}{2}\left(\sum_i \bar{h}_i\right)^{-1} \sum_i \bar{h}_i r_i^2 \qquad (9)$$

and the form factor,

$$P_{chain}(q) = \left(\sum_i \bar{h}_i\right)^{-1} \sum_i \bar{h}_i \frac{\sin(qr_i)}{qr_i} \qquad (10)$$

To account for the finite thickness of the ssDNA in calculation of the scattering profiles, the form factor was multiplied by the effective scattering of the cross-section derived from crystal structures (see Figure 7): $I(q) = P_{chain}(q) \times I_{CS}(q)$. The effect of multiplication by $I_{CS}(q)$ is that the SAXS radius of gyration is slightly larger than that of the skeletal chain, $<R_g^2> = <R_{g0}^2> + (5.04\ \text{Å})^2$, where 5.04 Å is the effective radius of gyration according to a Guinier fit of $I_{CS}(q)$ at low-q. Calculation of the interparticle interference function follows the method of Hubbard and Doniach[41] generalized for nonidentical particles, and is described in Methods. Numerical results are given in Table III.

Chain Compaction: Model vs. Experiment

Comparison of modeled and experimental values for R_g and R_{FRET} in Tables I and III shows that the SAXS predictions are about 4% lower than observation, while FRET predictions are about 10% higher than observed. However, on a relative scale, the model captures the extent of chain compaction. In Figure 9, R_g and R_{FRET} from the model and experiment were normalized by their value at 20 mM NaCl. The predictions reproduce the magnitude of chain compaction and overall shape of the curve at all Mg concentrations. Furthermore, the model suggests that R_{FRET} and R_g are nearly proportional, and correspondingly the SAXS and smFRET measurements of the chain size are proportional within error.

Polymer Scaling Properties of Poly(dT): Model vs. Experiment

At high q, the polymer scaling behavior becomes important, and this is a valuable way to test whether the data are consistent with a swollen polymer. The inset of Figure 10 compares the calculated form factor of the chain multiplied by q, at each salt condition, where the arrow indicates the direction of increasing Mg^{2+}. In the main part of Figure 10, the SAXS data are plotted as I(q)*q vs q, along with model predictions. The

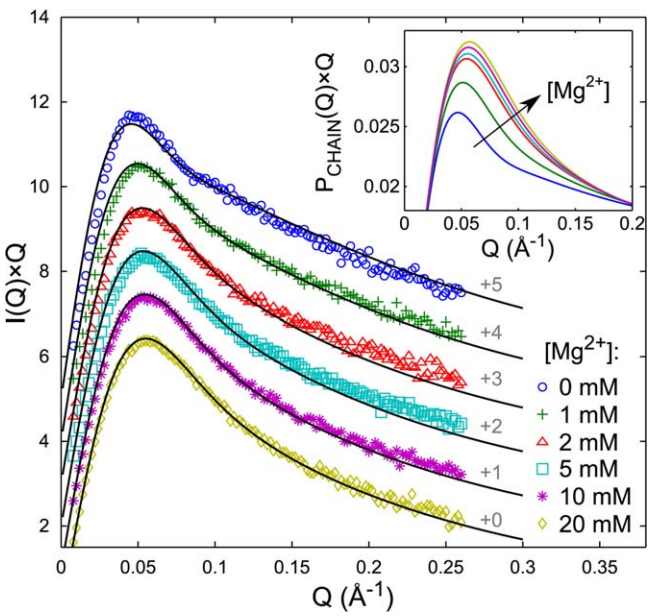

FIGURE 10 At high q, the form factor is sensitive to polymer scaling properties. Model form factors at each Mg concentration indicated in the legend and 20 mM NaCl are multiplied by q to emphasize the high q region, and plotted in the inset. To compare with experiment, form factors were multiplied by the effective cross section and scaled to match experimental curves at $q = 0.1$ Å^{-1}. The curves are offset vertically (for clarity) as indicated next to each curve.

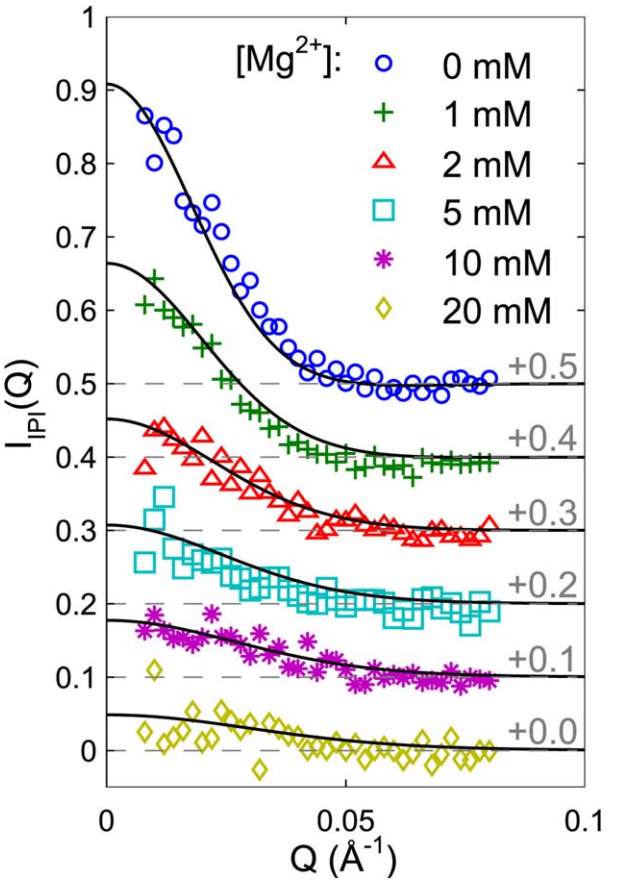

FIGURE 11 Interparticle interference functions $I_{IPI}(q)$ for model (solid lines) and experiment (points) at the highest DNA concentration (nominally 0.2 mM) are plotted for each Mg^{2+} concentration, given in the legend, and offset vertically by the amount indicated.

high concentration (0.2 mM) DNA data are shown, rather than the extrapolated form factor, because they have much higher signal-to-noise: interparticle interference is present at low-q, but I(q) for $q > 0.05$ Å^{-1} can be compared with the predictions. The model captures the change in shape of the scattering profile as Mg is added, indicating that the interpretation of poly(dT) as a swollen coil is essentially correct.

The predicted and experimental interparticle interference functions plotted in Figure 11 show a change in shape consistent with repulsive interactions between chains becoming shorter-ranged with increasing Mg. Note that because B_2 was determined by experiment, the magnitude of the interparticle interference function must agree; however, the shape of the interference function is an independent prediction.

Ion Atmosphere Around dT_{30}

The charge renormalization parameter can be interpreted physically in terms of the ion atmosphere using PB theory, where $(1 - f)$ corresponds to a fraction of the macromolecule's charge that is compensated by strongly condensed ions.[31] The physical interpretation of f becomes more complicated in mixed salt, so here we concentrate on the case of 20 mM Na with no added Mg. Then, the number of strongly condensed sodium ions is

$$N_+^{\text{condensed}} = (1-f)N_{\text{phos}}$$

For dT_{30}, $(1 - 0.581) \times 29 \sim 12.1$ ions. The number of strongly condensed ions is not easily measured experimentally; rather, ASAXS and heavy-ion report the total number of excess counterions,

$$N_+^{\text{excess}} = N_+^{\text{condensed}} + N_+^{\text{diffuse}}$$

To predict $N_+{}^{\text{excess}}$, one needs to calculate the number of diffusively bound ions, $N_+{}^{\text{diffuse}}$. Unfortunately, there is no unique way to distinguish between a strongly condensed and diffusively bound ion within PB theory on the basis of distance from the macromolecule or electrostatic potential, except for simple geometries.[31] Here, we follow previous work (e.g., Ref. 42) and define the diffuse ions as those bound with a potential less than k_BT/e_0. According to PB theory, the excess counterion density is

$$\Delta n_+(\vec{r}) = n_+^{\infty}(\exp(-e_0\Phi(\vec{r})/k_BT) - 1) \qquad (11)$$

where $n_+{}^{\infty}$ is the bulk density. The potential $\Phi(\vec{r})$ is given by the renormalized DH solution for charged spheres with radius a:

$$\frac{e_0\Phi(\vec{r})}{k_BT} = f\sum_{i=1}^{N_{\text{phos}}}\left(\frac{Zl_B}{1-\kappa a}\right)\frac{\exp(-\kappa(|\vec{r}-\vec{r}_i|-a))}{|\vec{r}-\vec{r}_i|} \qquad (12)$$

The total number of diffusively bound ions, $N_+{}^{\text{diffuse}}$ is the integral of $\Delta n_+(\vec{r})$ over the electrolyte volume where $|\Phi(\vec{r})| < k_BT/e_0$. The calculation of $N_+{}^{\text{diffuse}}$ for poly(dT) is slightly more complicated: the integral also must be performed over the distribution of chain conformations. To accomplish this efficiently, Monte Carlo integration was performed as described in Methods.

The calculated number of excess ions within a radius R of the center of mass of the chain is shown in Figure 12. The distribution of condensed ions follows the phosphate distribution, while the diffuse ion cloud extends outside the chain. The number of excess Na^+ approaches a value of ~19 far from the chain. This calculation agrees with the measured values within experimental error.

Explicit Solvent Molecular Dynamics Simulations of dT_{10}

The number of excess counterions was also computed by all atom molecular dynamics simulations (MD). MD accounts for

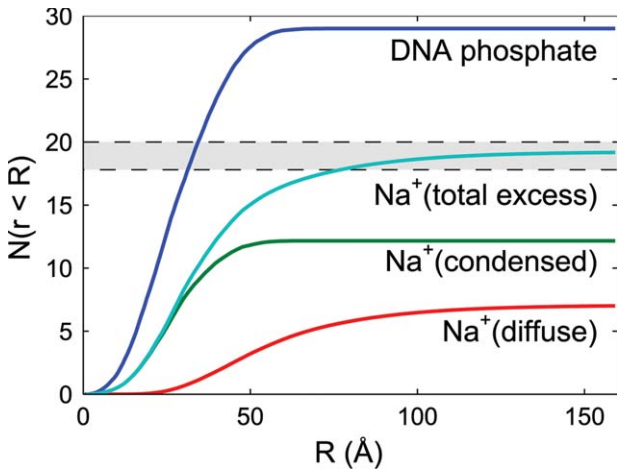

FIGURE 12 Number of charged phosphate groups and excess counterions within a radius R of the chain center of mass computed using the model for dT_{30} in 20 mM NaCl. The total number of excess sodium ions is the sum of the ions in diffuse and condensed regions, defined in the text. The experimental values for the number of excess monovalent ions are represented as a gray box bounded by dashed lines showing the $\pm 1\sigma$ confidence interval.

ions and water as well as the atoms of the nucleic acid explicitly, in contrast to PB approaches that model the solvent and ions as a continuous medium.[31] Here we report simulation results for dT_{10} in 50 mM Na^+, which yield the number of excess Na^+ directly (see Methods). The simulations find 5 ± 0.3 ions around dT_{10}, corresponding to 56% charge compensation (polymer charge of 9) due to excess ions. The results are lower than those obtained from experiments described above, where 66% (19 ions and polymer charge of 29) of the charge is compensated by excess Na+ ions around dT_{30}. The variance between experiment and MD simulation is discussed below.

DISCUSSION

We applied a simple polymer model for poly(dT), in which ionic strength dependent interactions between chains are fixed by experimental measures of the second virial coefficient and observables are predicted without additional input from experiment. This represents an improvement over previous applications of WLC, where the ionic strength dependent persistence length was a fitting parameter. As a simplifying assumption, the bond lengths and angles are fixed, and the potentials determine the distribution of chain conformations. While it would seem that this is fundamentally different from the WLC model (with no explicit potential, and a salt-dependent persistence length), it is merely a different way of accounting for bond correlations. The main difference is that for excluded volume, the bond correlations may be much longer-ranged than in the WLC, where by definition they decay exponentially.

A critical assumption of the proposed model is that the same driving force for interchain repulsion (e.g., screened electrostatics) is responsible for changes in the chain dimensions. Measurements by SAXS and smFRET allowed us to test this assumption. The model predictions reproduce the salt-dependent changes in chain dimensions, as well as the number of ions the chains attract.

While model predictions for the radius of gyration and FRET distance agree on a relative basis, they fall outside the experimental error on an absolute basis. Because for this study we focused on constraining the electrostatic aspects of poly(dT) rather than the backbone properties, it is possible that the geometrical parameters could be tuned for better agreement with experiment. However, we also neglected the scattering of the ion atmosphere when computing R_g, as well as the possible influence of dye labels and chain dynamics on FRET.[43] Future studies that rigorously test MD using FRET and SAXS data will account for these effects.

Interestingly, the MD simulations do not reproduce the same fraction of charge neutralization as measured experimentally. Although shorter chains and lower ambient salt concentrations are used for MD than for experiment (dT_{10} as opposed to dT_{30}, and 50 mM Na as opposed to 100 mM Na), we expect the number of excess ions to be similar across these different conditions.[44] The deviation may be statistical (even hundreds of nanoseconds may be insufficient to sample the relevant range of conformations of dT10). Moreover the small number of ions in the simulation box (see Methods) implies other statistical inaccuracies. While the differences can also be a result of force field tuning, we comment that all atom simulations quantitatively captured ion distribution near the A form of dsRNA.[24]

SUMMARY AND OUTLOOK

The work described herein focuses on poly(dT) measured over a relatively small range of ionic conditions, less than, but approaching physiological ionic strength. Future work will extend these measurements to provide a broader test of the model. At lower ionic strength, repulsive forces increase dramatically enabling more robust comparison with electrostatic parameters. At much higher ionic strength, the possibility of achieving theta solvent conditions[10] will allow an assessment of the balance between attractive and repulsive forces. Opportunities also exist to explore the distinctly different roles of mono and multi (e.g., di-) valent ions in providing charge compensation. Although we described measurements in mixed salt solutions, all of these important details were absorbed by the

charge renormalization parameter, *f*. Our simple model does not address the distinctly different contributions of mono- or divalent ions to screening. The surprisingly efficient charge compensation measured in the presence of Mg^{2+} in this simple system may be an ideal vehicle for probing effects such as fluctuations, correlations, or polarizability. The inclusion of models of more complex electrostatic environments,[31,45,46] might be readily testable in terms of the apparent charge fraction, *f*. Furthermore, a detailed atomic level understanding of the differing roles of monovalent and divalent ions may be gleaned from MD simulations, once discrepancies with experiment are resolved. MD simulations will also be able elucidate the presence of counterion-induced deformations in the chain, or "wrinkles," that have been observed in coarse-grained MD with explicit ions.[47]

Finally, and most significantly, we plan to extend these approaches to study chains of mixed sequence, to gain insight into the sequence preferences found in important biological molecules such as mRNA or riboswitches. Sequence and sugar (ribose vs. deoxyribose) effects are expected to change base stacking interactions, which adds another level of complexity, but in principle can also be modeled using virtual bonds with measured B_2. Tight collaboration with MD will be essential to explore the subtle details that lead to highly specialized biological function.

MATERIALS AND METHODS

Preparation of dT_{30} Samples for X-ray Scattering Experiments

DT_{30} was synthesized and HPLC purified by Integrated DNA Technologies. For SAXS studies of the conformation and electrostatics in mixed ion environments, dT_{30} was buffer exchanged with solutions containing 1 mM Na-MOPS, pH 7.0, 20 mM NaCl and 0, 1, 2, 5, 10, or 20 mM $MgCl_2$ using spin concentrators (Amicon Ultra-0.5, 3 kDa cutoff, EMD Millipore). For each ionic condition, a DNA concentration series was prepared by dilution with the matching buffer. For measurements of the ion atmosphere, dT_{30} was buffer exchanged with 1 mM Na MOPS pH 7.0 and 100 mM NaCl or 100 mM RbCl. The DNA concentration was determined by UV absorption at room temperature (Cary 50, Varian) assuming an extinction coefficient of $\varepsilon_{260nm} = 2.436 \times 10^5\ M^{-1}\ cm^{-1}$.

SAXS Data Collection and Analysis

SAXS data were collected at CHESS beamline G1 using an X-ray energy of 10.53 keV. Samples and matching buffers were loaded sequentially in a 2-mm diameter, 10-μm-thick quartz capillary. During each exposure, scattering patterns were collected on a photon-counting area detector (Pilatus 100K, Dectris) at a distance of 1.689 m from the sample and normalized by the beamstop PIN diode current. Oscillation of the sample within the capillary was used to reduce radiation damage.[48] Data processing was performed in MATLAB. X-ray images were azimuthally averaged, and repeated exposures of the same sample were compared to verify that no damage had occurred. The uncertainty of $I(q_i)$ at each bin of N_i pixels was estimated from the standard deviation σ_i as $\sigma_i N_i^{(-1/2)}$. This uncertainty was propagated through buffer subtraction and averaging of repeated exposures.

SAXS profiles at varying DNA concentration c were matched over the range $0.1 < q < 0.26$ Å^{-1}, and the form factor P(q) was obtained by linear extrapolation to c = 0 of I(q,c) at each value of *q*. To determine the second virial coefficient, B_2, this set of curves was fit using a model for the interparticle interference of dilute solutions,

$$I(q,c) = P(q) - 2P(0)B_2A_0cP_{IPI}(q) \quad (13)$$

where $P_{IPI}(q)$ is the q-dependent part of the interparticle interference term with $P_{IPI}(0) = 1$, and A_0 is defined so that *c* is in M and B_2 is in Å^3. To find P(0), the low-q portion of P(q) was fit with the Debye function for a random coil $P_D(q)$.[49] The interference terms $|I(q,c) - P(q)|$ were found empirically to have a Gaussian shape for $q < 0.035$ Å$^{(-1)}$. The Gaussian shape is also observed in theoretical calculations of the interference function (see Figure 11). Therefore, we chose $P_{IPI}(q) = \exp(-q^2d^2)$, where the length scale *d* is independent DNA concentration, but may depend on salt. For each salt condition, we obtained B_2 (and *d*) using a simultaneous nonlinear least squares fit of the interference terms to $2P(0)B_2A_0c\ P_{IPI}(q,d)$ for $q < 0.035$ Å^{-1}. To double-check the validity of the above assumptions, the low *q* profiles were reconstructed using $I(q,c) = P_D(q) - 2P(0)B_2A_0c\exp(-q^2d^2)$ and compared directly with the experimental data, as in the inset of Figure 2.

Heavy Ion and ASAXS Measurement of Excess Ions Around dT_{30}

For the heavy ion replacement method, solutions of 0.1 mM dT_{30} in Na and Rb solutions were prepared as described above, and placed on an absolute scale using water as a calibrant.[50] The largest experimental uncertainty for this technique was from the measurement of the DNA concentrations. Therefore, ΔN^e_{tot} was calculated separately for repeated concentration measurements, and the uncertainty in N_{ions} was estimated using linear regression.

ASAXS profiles from 0.3 mM dT_{30} with Rb were acquired at beamline C1 at CHESS, as described in Ref. 29. Two X-ray energies were chosen below the Rb absorption edge, $E_{lo} = 15.093$ and $E_{on} = 15.191$ keV, and the X-ray fluorescence spectrum from a dilute RbCl solution was measured and used to calculate f' at E_{lo} and E_{on}.[51] The number of Rb ions was obtained using $N_{Rb} = (I_{low}(0)^{1/2} - I_{on}(0)^{1/2})\ S(0)^{-1/2}\ \Delta f'^{-1}$, where $S(0) = 0.95$ corrects for the interparticle interference effect.[29] The extrapolations of $I_{low}(q)$ and $I_{on}(q)$ to $q = 0$ shown in Figure 5 were performed using the Bayesian indirect Fourier transform, as described above.

Single-Molecule FRET

For smFRET measurements of the end to end distance, dT_{30} with 3′ Cy3 and 5′ Cy5 labels was synthesized and HPLC purified by Integrated DNA Technologies. DNA was prepared in solutions containing 20 mM TRIS buffer, pH 8, 20 mM NaCl and 0, 0.5, 1, 5, 10, or 20 mM $MgCl_2$.

The experimental methods were similar to those used in a previous smFRET study involving poly(dT).[5] Briefly, smFRET

measurements of the freely diffusing DNA were performed on a custom confocal microscope using an Olympus UAPO 40× objective. Donor and acceptor fluorescence were separated by a 660LP dichroic, then additionally filtered by a 570/40 band-pass (donor channel) and 660 long pass (acceptor channel). Each fluorescence signal was collected through a 30-μm pinhole and detected by a single-photon counting PMT (Hamamatsu). The intensity was sampled at 60 MHz using a Flex03LQ-01 correlator card (correlator.com) in the photon counting mode.

Data analysis was also similar to previous work,[5] one exception being that a threshold was only placed on the sum of donor and acceptor channels, rather than on each channel individually. Histograms of events passing the threshold criterion were fit to two Gaussians: the first low-FRET peak representing molecules with an inactive acceptor, and the second representing the signal of interest. The peak FRET efficiency E_{FRET} was converted to the FRET-averaged end-to-end distance $\langle R\rangle_{\mathrm{FRET}} = R_0\,(1/E_{\mathrm{FRET}} - 1)^{1/6}$ using the experimentally derived Förster radius for the dye pair conjugated to poly(dT): $R_0 = 56.4$ Å.[5]

Calculation of Scattering Profiles from the Virtual Bond Representation of Poly(dT)

For a chain with contour length L and a uniform density along the contour, the form factor is

$$P_{\mathrm{chain}}(q) = \int_0^1 \int_0^1 ds ds' \frac{\sin\left(q|\vec{r}(s) - \vec{r}(s')|\right)}{q|\vec{r}(s) - \vec{r}(s')|} \tag{14}$$

where s is the distance along the chain divided by the contour length. To efficiently calculate $P_{\mathrm{chain}}(q)$, a histogram of distances between points was generated and $P_{\mathrm{chain}}(q)$ was found using Eq. (10). For comparison between chain form factors and CRYSOL calculations (Figure 7), s was discretized with $L\Delta s \approx 0.1$Å, and the histogram was generated for all pairs of points. When computing form factors for a distribution of chain conformations (Figure 10), the histogram was generated from randomly sampled chains and values of s on the interval [0,1].

Calculation of the Interparticle Interference Function

Calculation of the interparticle interference function uses a Monte Carlo method,[41] and begins as described above for the calculation of B_2 by generating pairs of chains drawn from $\Omega_{\kappa,\mathrm{f}}$ and center of mass displacement vectors with uniform probability density in a volume $V_{\max}$. For each pair of chains labeled by the index m, an interparticle distance histogram $p_i(\mathrm{m})$ with bins centered at r_i is calculated for distance vectors that span the two chains, normalized so that $\sum_i pi^{(m)} = 1$. The interaction energy between chains W_{m} was calculated as described above. After M iterations, the interparticle interference function was computed using

$$I_{\mathrm{IPI}}(q) = A_0 c \frac{V_{\max}}{M} \sum_{i,m} p_i^{(m)} \left(1 - \exp\left(-W_m\right)\right) \frac{\sin\left(qr_i\right)}{qr_i} \tag{15}$$

Monte Carlo Integration of the Ion Atmosphere

At each step in the Monte Carlo integration, a chain was chosen from the distribution $\Omega_{\kappa,\mathrm{f}}$ and a point $\vec{r}$ was chosen uniformly within the integration volume (a sphere of radius $r_{\max} = 160$ Å where the origin is the chain's center of mass). The point was classified as being within one of three phases according to the electrostatic potential $\Phi(\vec{r})$ defined in Eq. (12), and the distance between the point at $\vec{r}$ and the nearest chain atom: $d_{\min}$,

$$\text{phase at } \vec{r} = \begin{cases} \text{chain}, & d_{\min} < a \\ \text{condensed ion}, & d_{\min} \geq a \text{ and } |\Phi(\vec{r})| \geq \Phi_{\mathrm{cutoff}} \\ \text{diffuse ion}, & d_{\min} \geq a \text{ and } |\Phi(\vec{r})| < \Phi_{\mathrm{cutoff}} \end{cases} \tag{16}$$

The chain and condensed ion phases were assigned densities of 1 to compute their volumes, and the density for the diffuse ion phase was calculated from the potential through Eq. (11). A radial density histogram for each phase was incremented accordingly. After 10^6 iterations, the number of diffuse ions within a radius r of the chain center of mass was determined from a cumulative sum of the density histogram. Similarly, the volumes of the chain and condensed ion phases within r, V (r), were given by the sum of the respective density histograms. The number within r was calculated assuming uniform density of particles within these phases: $N(r) = N_{\mathrm{tot}} V(r)/V(r_{\max})$.

Explicit Solvent MD Simulations of dT_{10}

About 150 ns of molecular dynamics simulations at 300 K was performed with MOIL suite of programs[52,53] with explicit account of water, ions, and DNA. To model the nucleic acid and its environment we used TIP3P[54] for water together with recent AMBER parameters for monovalent ions[55] and DNA.[56,57] The simulated system contained dT_{10}, 5,062 water molecules, and 10 Na^+, and 1 Cl^- ions. The center of mass of the DNA was constrained by harmonic springs to the center of the simulation box. The size of the simulation box was (55.75 Å)^3, and periodic boundary conditions were applied in all directions. The long-range component of the electrostatics interaction was calculated by particle mesh Ewald method[58] with a grid spacing of 32 × 32 × 32 Å^3. The cutoff for the real space part of the electrostatics and van der Waals interactions was set to 8.5 Å with a nonbonded list update for every eight steps. We used matrix version[59,60] of the SHAKE algorithm[61] to constrain the water bond lengths and angles; the bond lengths of the DNA, fixed by SHAKE, allowed us to use 1.5-fs time step. We also used RESPA[62] for dual time stepping. Here the reciprocal-space of the Ewald sum was calculated every four steps, while the rest of the forces were evaluated every step. We reported atom positions for every 3 ps for further analysis.

To compute the excess ions around the flexible chain (dT_{10}) we cut a slice of width d from the outer face of the cubic box in all directions. These regions are combined and used to compute the asymptotic concentration of Na^+ ions (c_{bulk}). Once we compute the c_{bulk} we find the excess ions simply as $N_{\mathrm{excess}} = N_{\mathrm{Na+}} - V_{\mathrm{box}}\, c_{\mathrm{bulk}}$ where $N_{\mathrm{Na+}}$ is the total number of Na^+ ions (10 here) and V_{box} is the total volume of the simulation box. The width of the cut d is an ad hoc parameter. We used $d = 3$, 4, 5, and 6 Å for comparison. The results for c_{bulk} with alternate cutoffs differ by no more than 10% suggesting convergence in estimating the bulk concentration. As a final result we reported the average of these results.

The authors thank Pollack group members and CHESS scientists Ken Finkelstein and Arthur Woll, and staff, for experimental assistance.

REFERENCES

1. Baumann, C. G.; Smith, S. B.; Bloomfield, V. A.; Bustamante, C. Proc Natl Acad Sci USA 1997, 94, 6185–6190.
2. Henzler-Wildman, K.; Kern, D. Nature 2007, 450, 964–972.
3. Eichhorn, C. D.; Feng, J.; Suddala, K. C.; Walter, N. G.; Brooks, C. L.; Al-Hashimi, H. M. Nucleic Acids Res 2012, 40, 1345–1355.
4. Campolongo, M. J.; Tan, S. J.; Smilgies, D. -M.; Zhao, M.; Chen, Y.; Xhangolli, I.; Cheng, W.; Luo, D. ACS Nano 2011, 5, 7978–7985.
5. Chen, H.; Meisburger, S. P.; Pabit, S. A.; Sutton, J. L.; Webb, W. W.; Pollack, L. Proc Natl Acad Sci USA 2012, 109, 799–804.
6. Doose, S.; Barsch, H.; Sauer, M. Biophys J 2007, 93, 1224–1234.
7. Laurence, T. A.; Kong, X.; Jäger, M.; Weiss, S. Proc Natl Acad Sci USA 2005, 102, 17348–17353.
8. Murphy, M. C.; Rasnik, I.; Cheng, W.; Lohman, T. M.; Ha, T. Biophys J 2004, 86, 2530–2537.
9. Dessinges, M. N.; Maier, B.; Zhang, Y.; Peliti, M.; Bensimon, D.; Croquette, V. Phys Rev Lett 2002, 89, 248102+.
10. Saleh, O. A.; McIntosh, D. B.; Pincus, P.; Ribeck, N. Phys Rev Lett 2009, 102, 068301.
11. Seol, Y.; Skinner, G. M.; Visscher, K. Phys Rev Lett 2004, 93, 118102+.
12. Sim, A. Y. L.; Lipfert, J.; Herschlag, D.; Doniach, S. Phys Rev E 2012, 86, 021901+.
13. Ha, B. Y.; Thirumalai, D. Macromolecules 1995, 28, 577–581.
14. Odijk, T. J. Polym Sci Polym Phys Ed 1977, 15, 477–483.
15. Skolnick, J.; Fixman, M. Macromolecules 1977, 10, 944–948.
16. Barrat, J. L.; Joanny, J. F. Europhys Lett 2007, 24, 333+.
17. Flory, P. J. Principles of Polymer Chemistry; Cornell University Press, Ithaca, NY, 1953.
18. Müller-Späth, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Rüegger, S.; Reymond, L.; Nettels, D.; Schuler, B. Proc Natl Acad Sci USA 2010, 107, 14609–14614.
19. Mao, A. H.; Crick, S. L.; Vitalis, A.; Chicoine, C. L.; Pappu, R. V. Proc Natl Acad Sci USA 2010, 107, 8183–8188.
20. Cao, S.; Chen, S. -J. RNA 2005, 11, 1884–1897.
21. Olson, W. K.; Flory, P. J. Biopolymers 1972, 11, 1–23.
22. Qiu, X.; Kwok, L. W.; Park, H. Y.; Lamb, J. S.; Andresen, K.; Pollack, L. Phys Rev Lett 2006, 96, 138101+.
23. Kirmizialtin, S.; Elber, R. J Phys Chem B 2010, 114, 8207–8220.
24. Kirmizialtin, S.; Pabit, S. A.; Meisburger, S. P.; Pollack, L.; Elber, R. Biophys J 2012, 102, 819–828.
25. Qiu, X.; Andresen, K.; Kwok, L. W.; Lamb, J. S.; Park, H. Y.; Pollack, L. Phys Rev Lett 2007, 99, 038104+.
26. Porod, G. In Small Angle X-ray Scattering; Glatter, O.; Kratky, O., Eds.; Academic Press: New York, 1982, 17–51.
27. Hansen, S. J Appl Crystallogr 2000, 33, 1415–1421.
28. Das, R.; Mills, T. T.; Kwok, L. W.; Maskel, G. S.; Millett, I. S.; Doniach, S.; Finkelstein, K. D.; Herschlag, D.; Pollack, L. Phys Rev Lett 2003, 90, 188103+.
29. Pabit, S. A.; Meisburger, S. P.; Li, L.; Blose, J. M.; Jones, C. D.; Pollack, L. J Am Chem Soc 2010, 132, 16334–16336.
30. Marcus, Y. Ion Solvation; Wiley: Chichester, NY, 1985.
31. Lamm, G.; Pack, G. R. Biopolymers 2010, 93, 619–639.
32. Zhang, Y.; Zhou, H.; Ou-Yang, Z. C. Biophys J 2001, 81, 1133–1143.
33. Verwey, E. J. W.; Overbeek, J. T. G. Theory of the Stability of Lyophobic Colloids; Elsevier: Amsterdam, 1948.
34. Yadav, T.; Carrasco, B. N.; Myers, A. R.; George, N. P.; Keck, J. L.; Alonso, J. C. Nucleic Acids Res 2012, 40, 5546–5559.
35. Chan, K.-W.; Lee, Y. -J.; Wang, C. -H.; Huang, H.; Sun, Y. -J. J Mol Biol 2009, 388, 508–519.
36. Fan, J.; Pavletich, N. P. Genes Dev 2012, 26, 2337–2347.
37. Kamtekar, S.; Berman, A. J.; Wang, J.; Lázaro, J. M.; de Vega, M.; Blanco, L.; Salas, M.; Steitz, T. A. Mol Cell 2004, 16, 609–618.
38. Svergun, D.; Barberato, C.; Koch, M. H. J. J Appl Crystallogr 1995, 28, 768–773.
39. De Gennes, P. -G. Scaling Concepts in Polymer Physics; Cornell University Press, Ithaca, NY, 1979.
40. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. J Chem Phys 1953, 21, 1087–1092.
41. Hubbard, S. R.; Doniach, S. J Appl Crystallogr 1988, 21, 953–959.
42. Alexandrowicz, Z.; Katchalsky, A. J Polym Sci A Gen Pap 1963, 1, 3231–3260.
43. Makarov, D. E.; Plaxco, K. W. J Chem Phys 2009, 131, 085105+.
44. Bai, Y.; Greenfeld, M.; Travers, K. J.; Chu, V. B.; Lipfert, J.; Doniach, S.; Herschlag, D. J Am Chem Soc 2007, 129, 14981–14988.
45. Kundagrami, A.; Muthukumar, M. J Chem Phys 2008, 128, 244901+.
46. Tan, Z. -J. J.; Chen, S. -J. J. Biophys J 2007, 92, 3615–3632.
47. Stevens, M. J.; McIntosh, D. B.; Saleh, O. A. Macromolecules 2012, 45, 5757–5765.
48. Nielsen, S. S.; Møller, M.; Gillilan, R. E. J Appl Crystallogr 2012, 45, 213–223.
49. Debye, P. J Phys Chem 1947, 51, 18–32.
50. Orthaber, D.; Bergmann, A.; Glatter, O. J Appl Crystallogr 2000, 33, 218–225.
51. Evans, G.; Pettifer, R. F. J Appl Crystallogr 2001, 34, 82–86.
52. Elber, R.; Roitberg, A.; Simmerling, C.; Goldstein, R.; Li, H.; Verkhivker, G.; Keasar, C.; Zhang, J.; Ulitsky, A. Comp Phys Commun 1995, 91, 159–189.
53. Ruymgaart, A. P.; Cardenas, A. E.; Elber, R. J Chem Theory Comput 2011, 7, 3072–3082.
54. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. J Chem Phys 1983, 79, 926–935.
55. Joung, I. S.; Cheatham, T. E. J Phys Chem B 2008, 112, 9020–9041.
56. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. J Am Chem Soc 1995, 117, 5179–5197.
57. Zgarbová, M.; Otyepka, M.; Sponer, J.; Mládek, A.; Banáš, P.; Cheatham, T. E.; Jurecka, P. J Chem Theor Comput 2011, 7, 2886–2902.
58. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. J Chem Phys 1995, 103, 8577–8593.
59. Elber, R.; Ruymgaart, A. P.; Hess, B. Eur Phys J Special Top 2011, 200, 211–223.
60. Weinbach, Y.; Elber, R. J Comput Phys 2005, 209, 193–206.
61. Ryckaert, J. -P.; Ciccotti, G.; Berendsen, H. J. C. J Comput Phys 1977, 23, 327–341.
62. Tuckerman, M.; Berne, B. J.; Martyna, G. J. J Chem Phys 1992, 97, 1990–2001.

Reviewing Editor: Nils G. Walter