SUPPORTING INFORMATION

# Visualizing disordered single-stranded RNA: connecting sequence, structure and electrostatics

Alex Plumridge[1,], Kurt Andresen[2] and Lois Pollack[1, †]

[1]School of Applied and Engineering Physics, Cornell University, Ithaca, NY 14853, USA

[2] Department of Physics, Gettysburg College, Gettysburg, PA 17325, USA

[†] To whom correspondence should be addressed.

# UV Methods

*UV Sample Preparation.*

High performance liquid chromatography (HPLC) purified RNA oligomers of $rU_{30}$ and $rA_{30}$ were purchased (deprotected) from Dharmacon (Lafayette, CO, USA). The lyophilized strands were re-suspended in sodium chloride-Tris-EDTA (STE) buffer (100mM NaCl, 10mM Tris, 1mM EDTA) and subsequently annealed at 90°C for 5 minutes, followed by slow cooling to room temperature. The annealed samples were dialyzed four times in 1mM Na MOPS (3-(N-morpholino)propanesulfonic acid) and 20µM EDTA, pH 7, with added salt as indicated for each sample, using Amicon Ultra-0.5 3kDa concentrators. Samples were spun at room temperature, at concentrations matching those used in the SAXS experiments (in the 100µM regime).

*UV Data acquisition.*

Absorption spectra were recorded over a wavelength range of 200 to 500 nm using a Cary 50 spectrophotometer. The optical density of each sample was monitored at 260 nm ($OD_{260}$) to report changes, such as precipitation or stabilization of structures. Uncertainties pertain to the dilution step required to measure nucleic acid concentrations on the Cary.

# CD Methods

*CD Sample Preparation.*

HPLC purified RNA oligomers of $rU_{30}$ and $rA_{30}$ were purchased (deprotected) from Dharmacon (Lafayette, CO, USA). The lyophilized strands were re-suspended in STE buffer and subsequently annealed at 90°C for 10 minutes, followed by slow cooling to room temperature. The annealed samples were dialyzed four times in 10mM Na MOPS and 20µM EDTA, pH 7, with added salt as indicated for each sample, using Amicon Ultra-0.5 3kDa concentrators. Samples were spun in a fridge at 4°C, and stored on ice until experiments.

*CD Data acquisition.*

Spectra were acquired using a BioLogic MOS 450 in CD mode. Eight scans were collected and averaged per spectrum, spanning wavelengths of 220–300 nm with a step size of 1 s/nm. The CD spectra were scaled to account for differing nucleic acid concentrations, which ranged from 40-60µM for $rU_{30}$, and 60-80µM for $rA_{30}$, and further adjusted to match at 300nm to allow comparison of the peak positions across solution conditions.

# SAXS Methods and Analysis

### *SAXS Sample Preparation.*

HPLC purified RNA oligomers of $rU_{30}$ and $rA_{30}$ were purchased (deprotected) from Dharmacon (Lafayette, CO, USA). The lyophilized strands were re-suspended in STE buffer, annealed at 90°C for 5 minutes, and slow cooled to room temperature. The annealed samples were dialyzed four times at room tempreature in solutions containing 1mM Na MOPS and 20µM EDTA, pH 7 and the quoted concentration of added salt, using Amicon Ultra-0.5 3kDa concentrators. RNA constructs were prepared on the same day as the experiments.

### *Raw SAXS curves and zero concentration extrapolations.*

Small angle X-ray scattering (SAXS) data were acquired at Cornell High Energy Synchrotron source (CHESS) beamline G1 using 11.18 keV X-rays. Scattering profiles were normalized by dividing by the beam intensity, measured after passage through a semi-transparent beamstop (250µm molybdenum foil). Samples were positioned in a 2 mm diameter glass capillary cell by a peristaltic pump. A syringe pump was used to oscillate the sample during beam exposure, to avoid radiation damage.

Three sample concentrations were measured for each experiment condition (salt concentration and construct) to control for the effects of interparticle interference in the SAXS data. Interparticle interference manifests as differences in the profiles in the low q regime, and can be removed by linearly extrapolating the low q portion of curves ($q<0.05Å^{-1}$) acquired at different sample concentrations to the zero concentration limit. Table S1 quotes the recovered sample concentration as measured by UV absorption, with error corresponding to the standard deviation of three independent dilutions. For each experimental condition, sample concentrations of approximately 150µM, 80µM and 40µM were used. At a given salt concentration, SAXS profiles were scaled to match for $q>0.05Å^{-1}$ to compensate for the differing sample concentrations. Interference-free SAXS profiles were generated by first linearly extrapolating the low q portion of these matched SAXS profiles to the zero concentration limit. Next, this extrapolated piece of the curve was stitched to the curve acquired at high sample concentration to yield the final structure factor free SAXS profile. The SAXS curves and zero-concentrated extrapolated curves are shown for each experimental condition in Figures S1-S4. In some cases, only two sample concentrations were measured due to limited sample quantity.

Due to a slight over-estimate of errors during the SAXS integration step, a rescaling of the uncertainties for the zero concentration curves was performed. The inverse Fourier transform (IFT) of the experimental data was calculated with GNOM [1], after which the uncertainties on the experimental curves were rescaled so that the chi-square for the IFT fits were equal to 1.

All SAXS data analysis was performed in MATLAB with in-house written code. The zero concentration extrapolated curves were used for all subsequent analysis ($R_g$ calculations, Kratky plots and fitting).

All of the final SAXS curves (zero concentration extrapolated) are deposited in the SASDBD under accession codes: SASDFA9, SASDFB9, SASDFC9, SASDFD9, SASDFE9, SASDFF9, SASDFG9, SASDFH9, SASDFJ9, SASDFK9, SASDFL9, SASDFM9, SASDFN9, SASDFP9, SASDFQP, SASDFR9, SASDFS9.

| Buffer Condition | $rU_{30}$ High (μM) | $rU_{30}$ Med (μM) | $rU_{30}$ Low (μM) | $rA_{30}$ High (μM) | $rA_{30}$ Med (μM) | $rA_{30}$ Low (μM) |
|---|---|---|---|---|---|---|
| 20mM NaCl | 145±12 | 69±4 | 34±10 | 123±9 | 65±10 | 27±3 |
| 100mM NaCl | 130±14 | 62±7 | 32±2 | 146±3 | 81±2 | 34±4 |
| 200mM NaCl | 134±6 | 68±7 | 29±3 | 126±11 | 66±5 | 40±10 |
| 400mM NaCl | 152±10 | 76±9 | 40±4 | 131±3 | 78±11 | 38±2 |
| 600mM NaCl | 124±3 | 62±7 | 32±4 | ND | 63±2 | 34±2 |
| 1mM $MgCl_2$ | 129±9 | 77±9 | 30±2 | 162±18 | 65±10 | 27±3 |
| 2mM $MgCl_2$ | 103±12 | 69±9 | 33±2 | 125±11 | 40±8 | 13±5 |
| 5mM $MgCl_2$ | 116±14 | 64±5 | 29±4 | ND | 86±11 | 15±2 |
| 10mM $MgCl_2$ | 116±2 | 72±3 | 37±4 | ND | ND | ND |

*Table S1. Recovered sample concentrations used for SAXS experiments.*

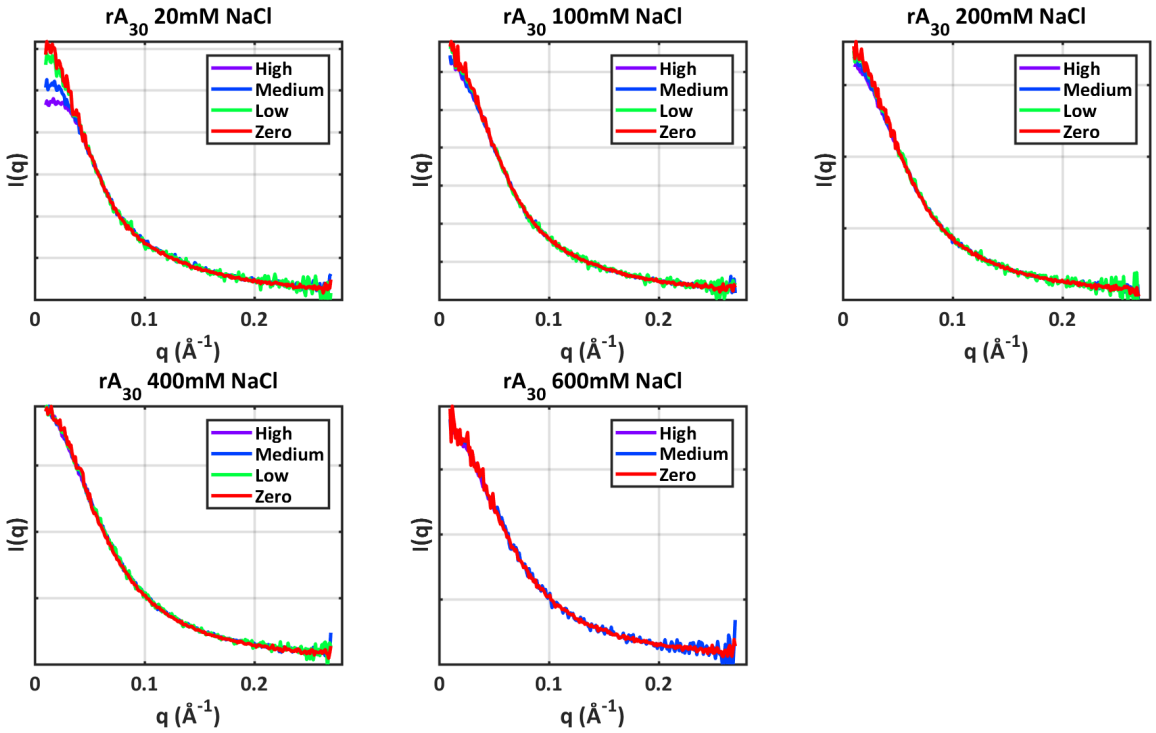



***Figure S1***. *Zero concentration extrapolation procedure for SAXS experiments on $rA_{30}$ as a function of [NaCl]. Multiple nucleic acid concentrations at a particular buffer composition were measured (high, medium and low, Table S1) and scaled to match at high-q. Next the low q portion of these curves was linearly extrapolated to the zero-concentration limit. The final structure factor free SAXS profiles (red) were constructed by stitching this low q extrapolated piece to the high q portion of the high sample concentration profile.*

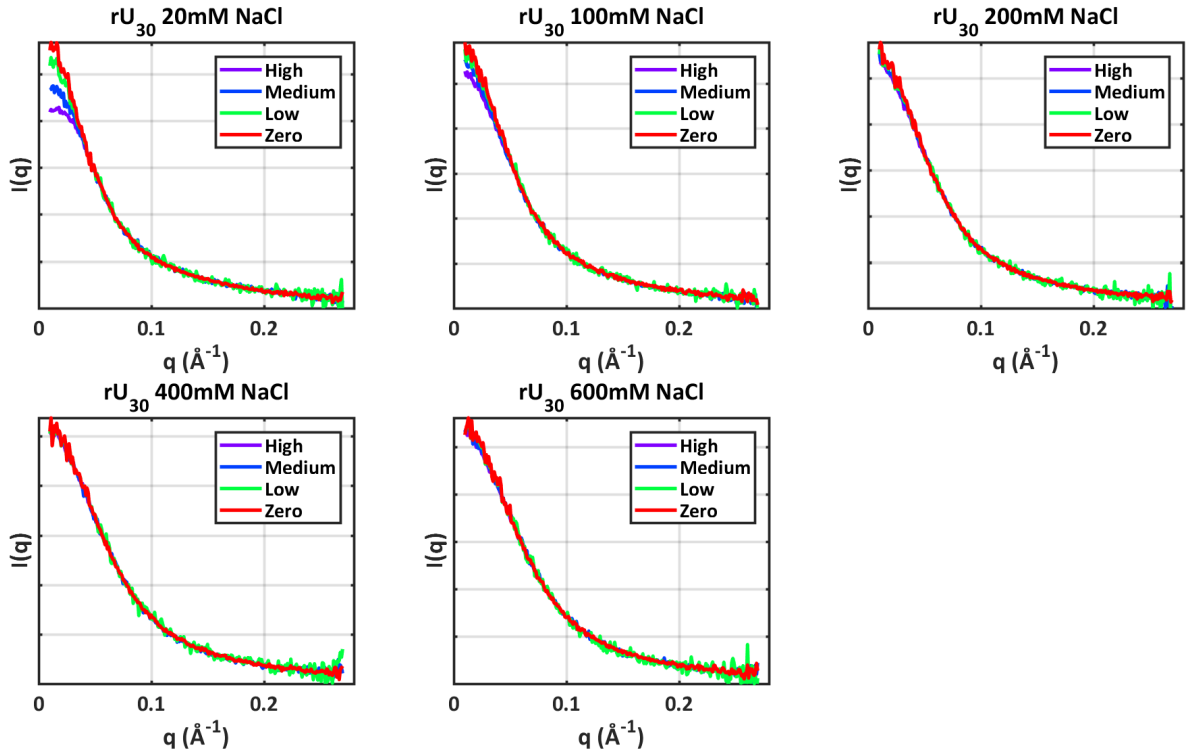



***Figure S2***. *Zero concentration extrapolation procedure for SAXS experiments on $rU_{30}$ as a function of [NaCl]. Multiple nucleic acid concentrations at a particular buffer composition were measured (high, medium and low, Table S1) and scaled to match at high-q. Next the low q portion of these curves was linearly extrapolated to the zero-concentration limit. The final structure factor free SAXS profiles (red) were constructed by stitching this low q extrapolated piece to the high q portion of the high sample concentration profile.*

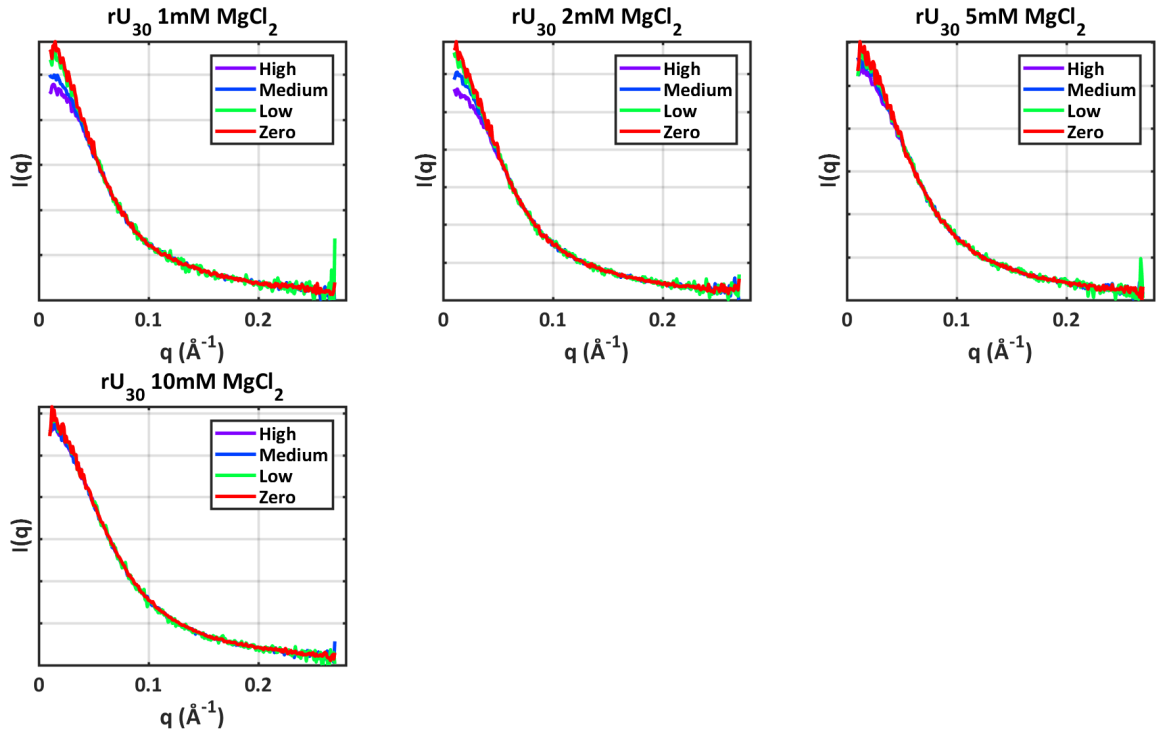



***Figure S3**. Zero concentration extrapolation procedure for SAXS experiments on $rU_{30}$ as a function of $[MgCl_2]$. Multiple nucleic acid concentrations at a particular buffer composition were measured (high, medium and low, Table S1) and scaled to match at high-q. Next the low q portion of these curves was linearly extrapolated to the zero-concentration limit. The final structure factor free SAXS profiles (red) were constructed by stitching this low q extrapolated piece to the high q portion of the high sample concentration profile.*

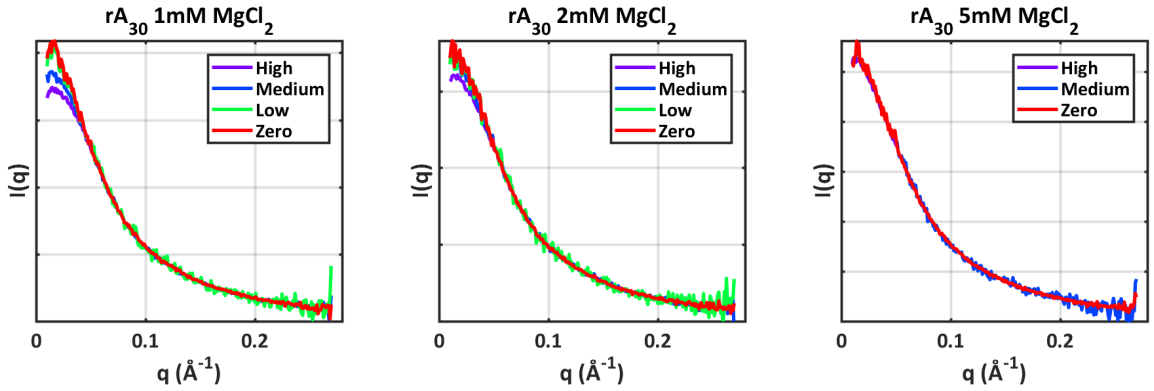



*Figure S4. Zero concentration extrapolation procedure for SAXS experiments on $rA_{30}$ as a function of [$MgCl_2$]. Multiple nucleic acid concentrations at a particular buffer composition were measured (high, medium and low, Table S1) and scaled to match at high-q. Next the low q portion of these curves was linearly extrapolated to the zero-concentration limit. The final structure factor free SAXS profiles (red) were constructed by stitching this low q extrapolated piece to the high q portion of the high sample concentration profile.*

*Basic SAXS Metrics - $R_g$*

The $R_g$ values provided in the Main text are tabulated in Table S2. The quoted values were determined by a Guinier fit in the appropriate region of the zero-concentration extrapolated SAXS curves: $qR_g < 1.3$. The quoted errors in $R_g$ are derived from the uncertainty in the straight line fit to the data.

| NaCl Conc. (mM) | $rU_{30}$ $R_g$ (Å) | $rA_{30}$ $R_g$ (Å) |
|---|---|---|
| 20 | 30.6 ± 0.8 | 27.2 ± 0.6 |
| 100 | 26.8 ± 0.6 | 24.5 ± 0.3 |
| 200 | 25.1 ± 0.3 | 25.1 ± 0.3 |
| 400 | 23.3 ± 0.3 | 23.8 ± 0.2 |
| 600 | 23.3 ± 0.3 | 22.2 ± 0.5 |

| $MgCl_2$ Conc. (mM) | $rU_{30}$ $R_g$ (Å) | $rA_{30}$ $R_g$ (Å) |
|---|---|---|
| 0 | 30.6 ± 0.8 | 27.2 ± 0.6 |
| 1 | 26.8 ± 0.7 | 24.5 ± 0.4 |
| 2 | 26.1 ± 0.6 | 23.7 ± 0.2 |
| 5 | 24.5 ± 0.3 | 23.0 ± 0.2 |
| 10 | 23.1 ± 0.2 | N.D. |

***Table S2:*** *$R_g$ values derived from SAXS data.*

*Basic SAXS Metrics – Kratky Plots*

Kratky plots (Figure S5) compare all SAXS profiles acquired at a given salt condition. To enable comparison of the curves at mid and high-q, they have been scaled to match for q<0.05Å$^{-1}$.

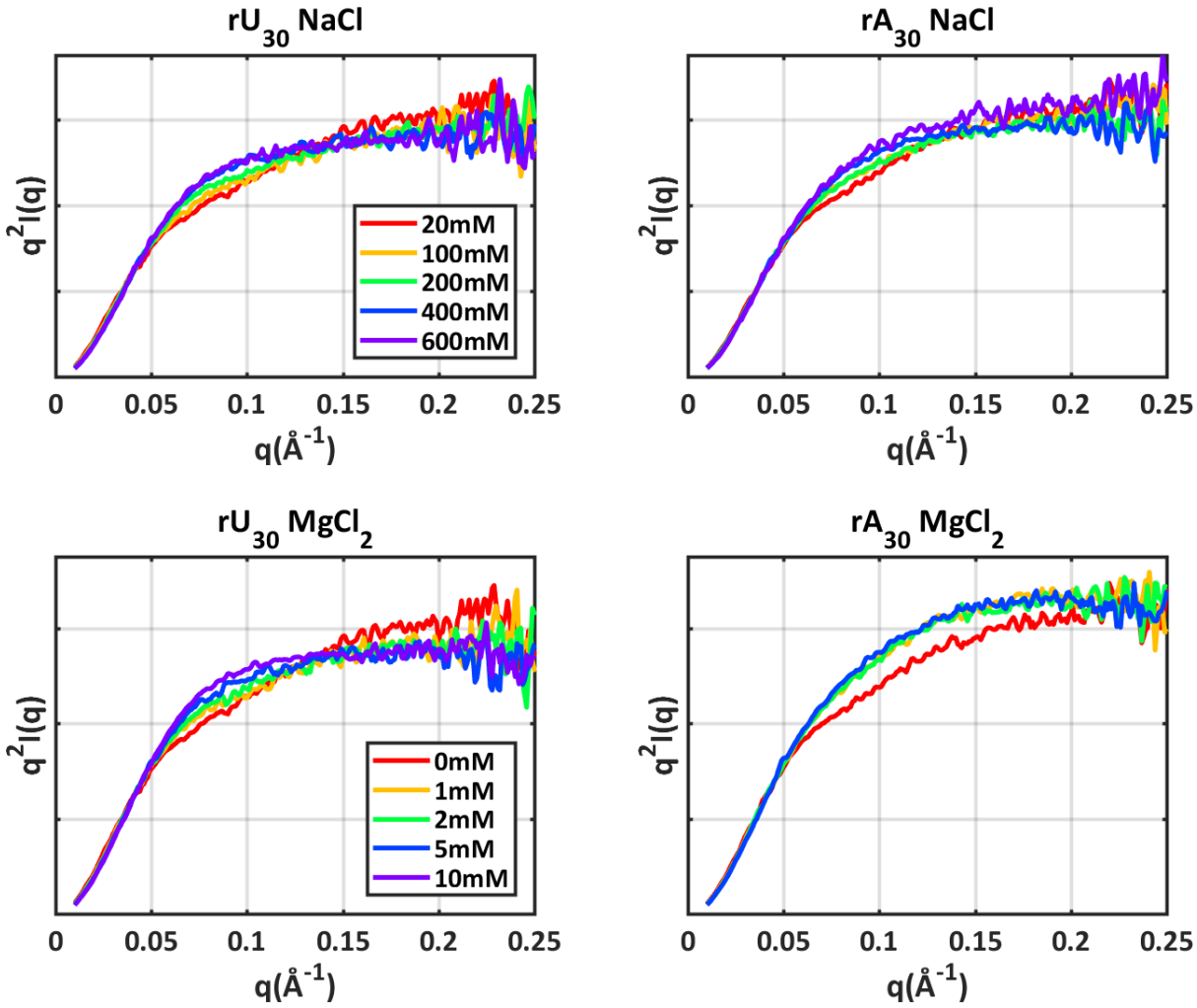



***Figure S5****. More details of chain conformation can be gleaned from the higher angle (larger q) scattering, emphasized in Kratky plots. All SAXS profiles have been scaled to have the same values for q<0.05Å$^{-1}$ to enable comparison of mid and high-q behavior.*

*Second virial calculations.*

Second virial coefficients were extracted from the data as described in reference [2]. Briefly, SAXS curves for each unique experimental condition (construct and salt condition) were scaled to match for q>0.05Å$^{-1}$. The form factor ***S(q)***, which corresponds to the scattering profile for a single scatterer unperturbed by interactions with neighbors, was determined by fitting each SAXS curve in the range q<0.05Å$^{-1}$, to that expected for a random coil:

$$S(q) = \frac{1}{1 + \frac{(qa)^2}{3}} - be^{-(dq)^2} \quad (1)$$

where ***a***,***b*** and ***d*** are fit parameters. This functional form was evaluated at q=0 to provide the intercept for each concentration ***S(q=0,c)***. The second virial coefficients, $B_2$, are then determined by fitting a line to the following equation which quantifies the effect of sample concentration on the interparticle interference:

$$\frac{S(q=0,c=0)}{I(q=0,c)} = 1 + 2AB_2c \quad (2)$$

The use of a coefficient ***A*** ensures that the sample concentration ***c*** is in units of ***M***, and $B_2$ is in units of Å$^3$. The concentration ***c*** for each sample was determined by dividing the scale factor used to match SAXS curves at high-q by the measured high sample concentration for each experimental condition (Table S1).

The extrapolation procedure is illustrated in Figures S6-S22, along with the methodology for deriving a value of $B_2$ from the data. Uncertainties in ***S(q=0,c=0)/I(q=0,c)*** are derived from the errors associated with the extrapolation of the full curve to its value at q=0. Final errors quoted on the second virial coefficients represent one standard deviation intervals on the fit parameter.

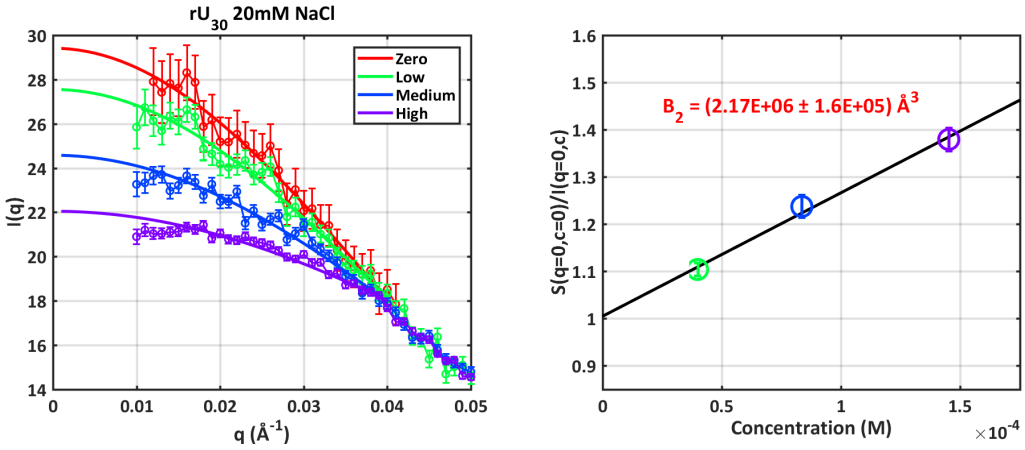



***Figure S6****. Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rU_{30}$ *in 20mM NaCl. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $\boldsymbol{B_2}$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

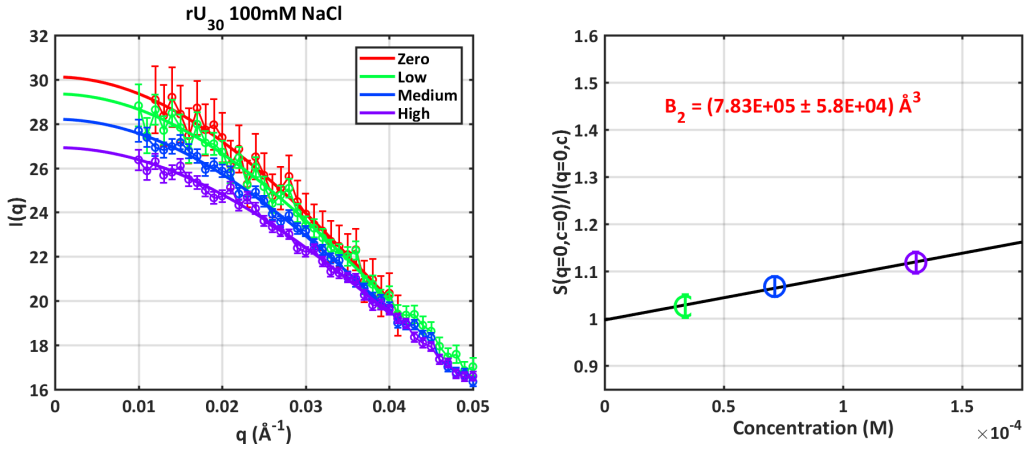



***Figure S7****. Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rU_{30}$ *in 100mM NaCl. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $B_2$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

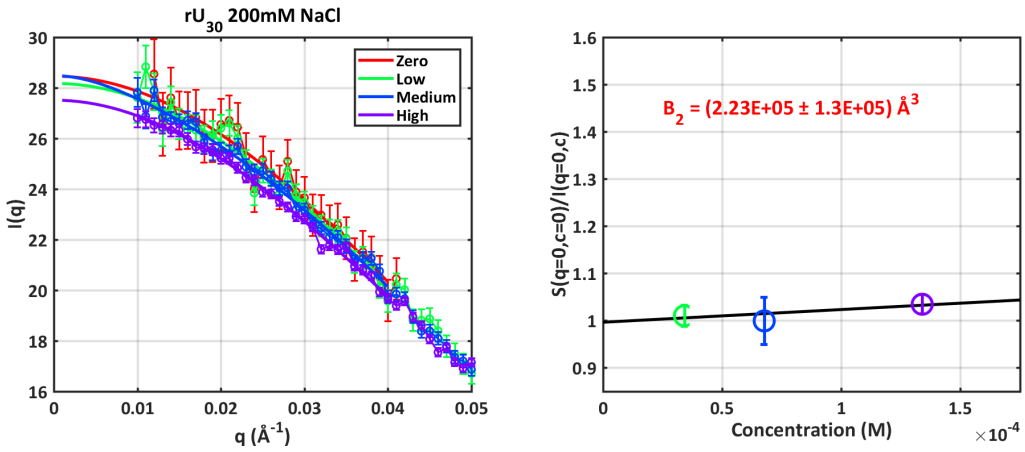



***Figure S8****. Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rU_{30}$ *in 200mM NaCl. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $B_2$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

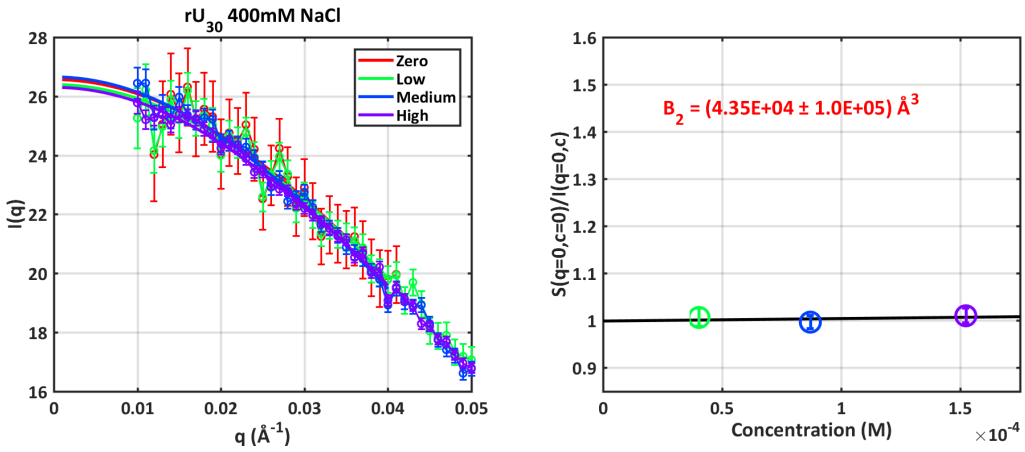



***Figure S9****. Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rU_{30}$ *in 400mM NaCl. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $\boldsymbol{B_2}$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

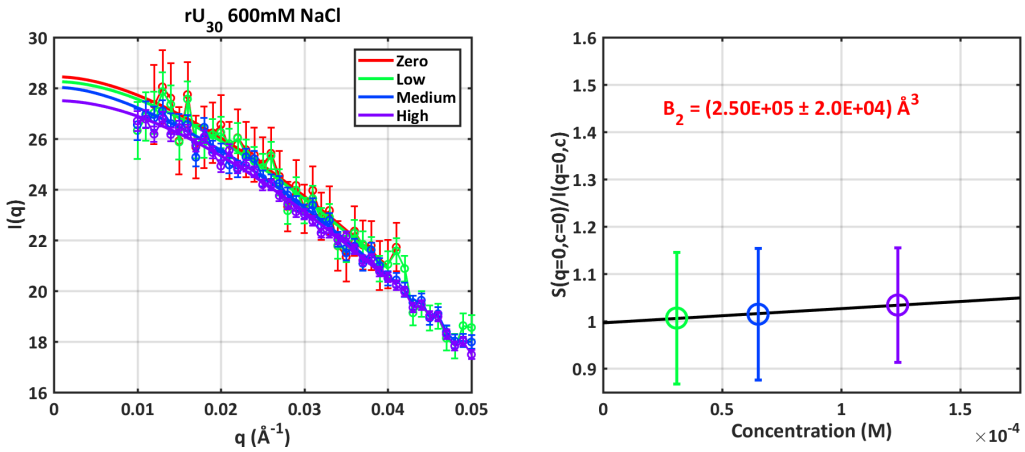



***Figure S10****. Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rU_{30}$ *in 600mM NaCl. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $\boldsymbol{B_2}$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

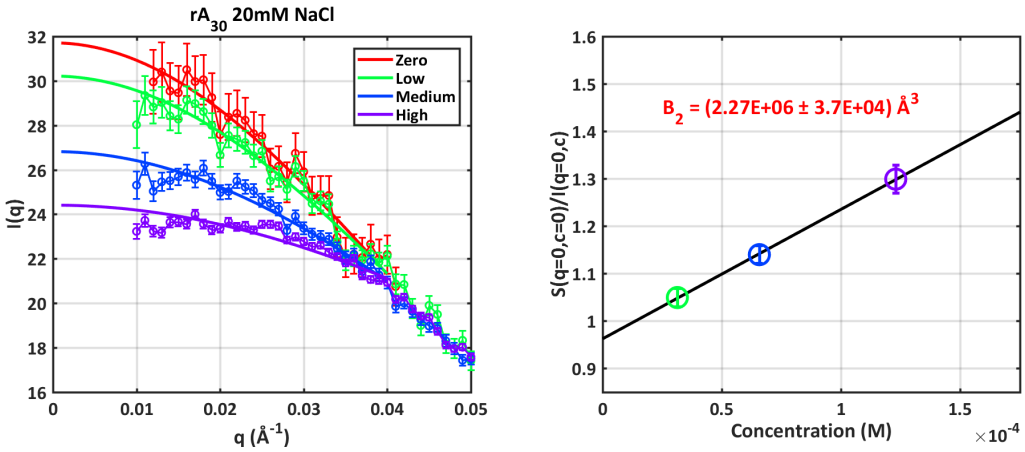



***Figure S11***. *Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for $rA_{30}$ in 20mM NaCl. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* ***$B_2$*** *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

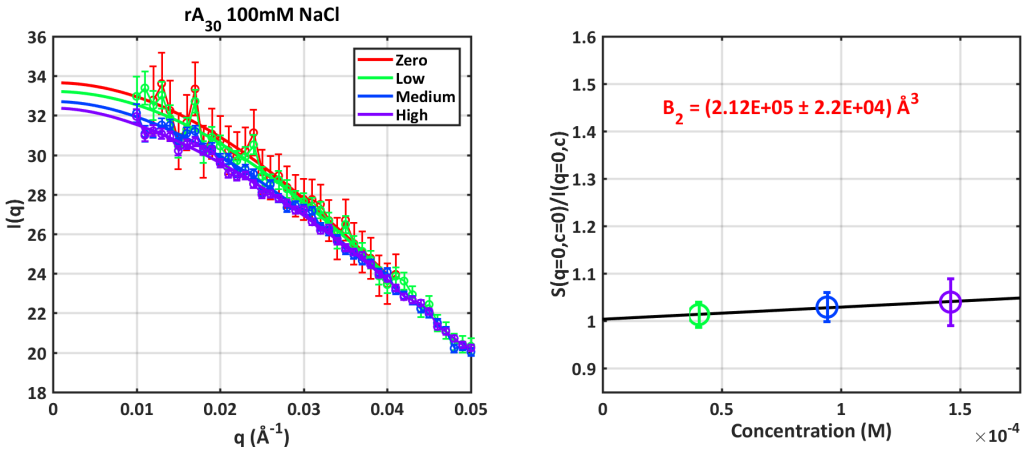



***Figure S12****. Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rA_{30}$ *in 100mM NaCl. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $\boldsymbol{B_2}$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

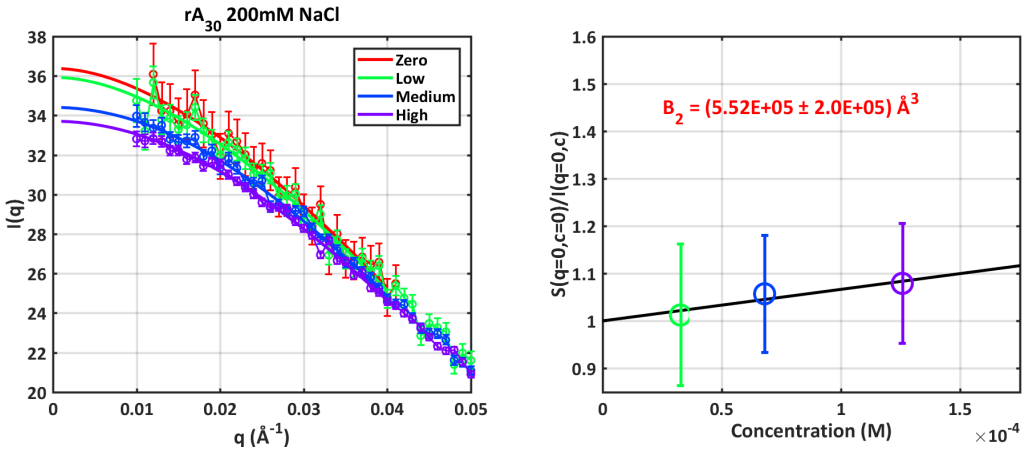



***Figure S13***. *Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rA_{30}$ *in 200mM NaCl. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $\boldsymbol{B_2}$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

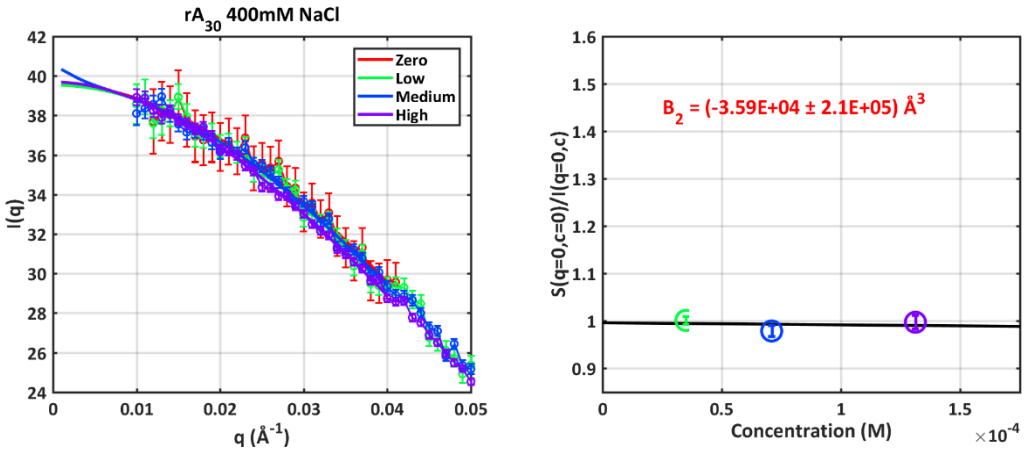



***Figure S14****. Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rA_{30}$ *in 400mM NaCl. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $\boldsymbol{B_2}$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

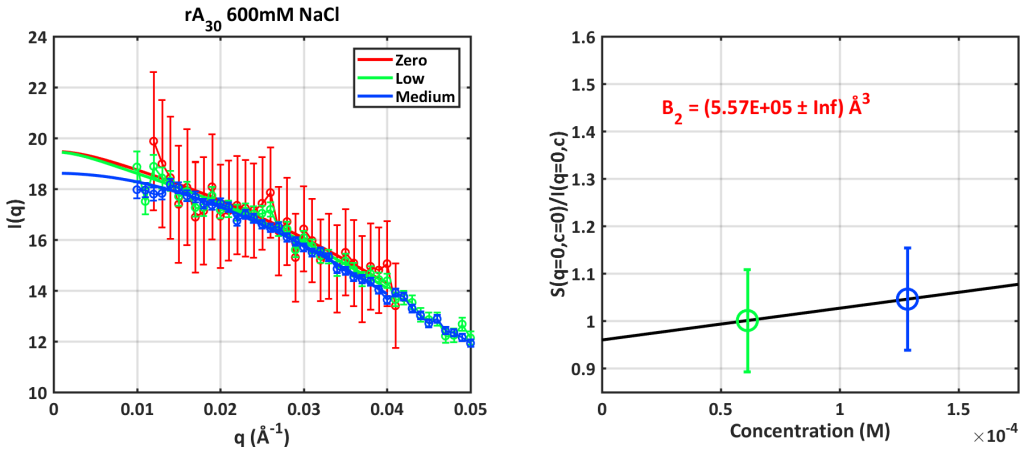



***Figure S15****. Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rA_{30}$ *in 600mM NaCl. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $\boldsymbol{B_2}$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

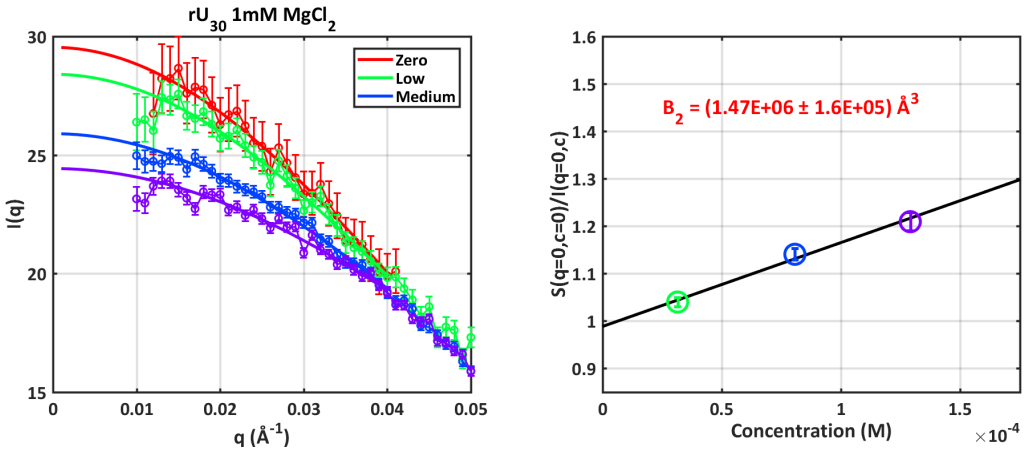



***Figure S16***. *Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for $rU_{30}$ in 1mM $MgCl_2$. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* ***$B_2$*** *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

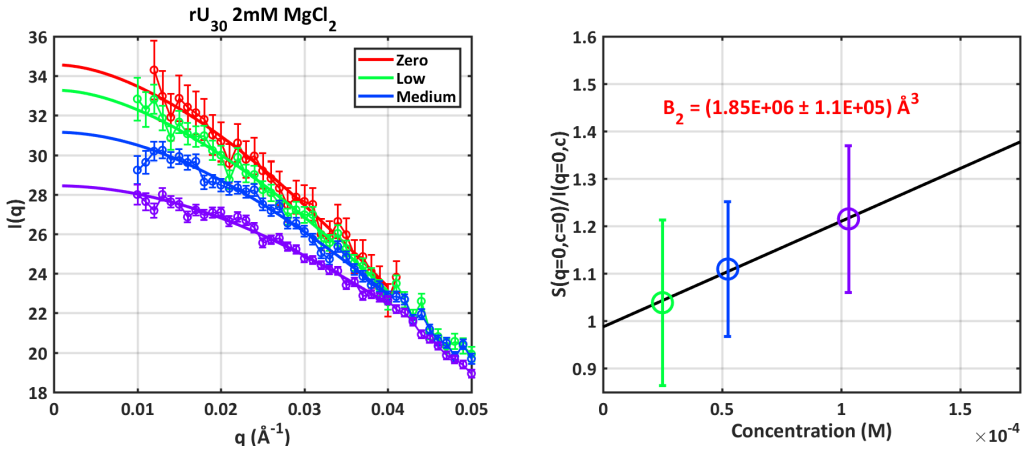



***Figure S17****. Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rU_{30}$ *in 2mM* $MgCl_2$*. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $\boldsymbol{B_2}$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

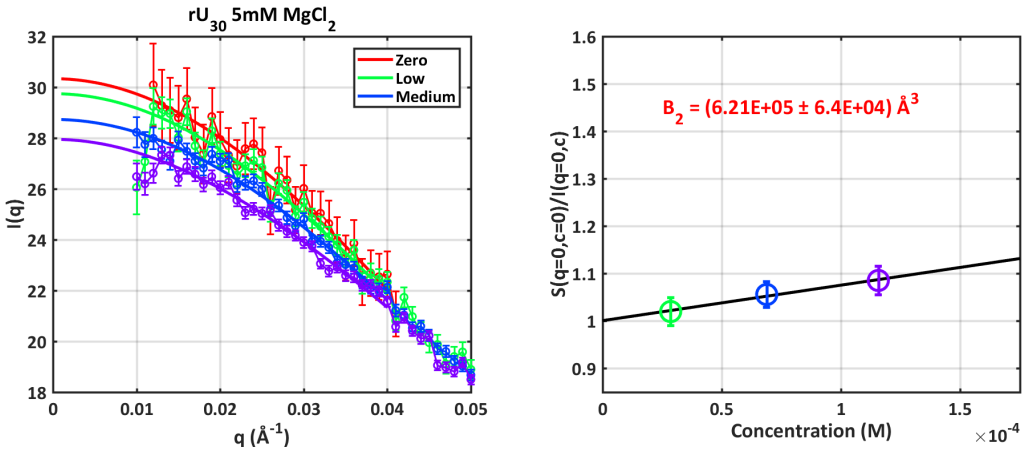



***Figure S18****. Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rU_{30}$ *in 5mM* $MgCl_2$*. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $\boldsymbol{B_2}$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

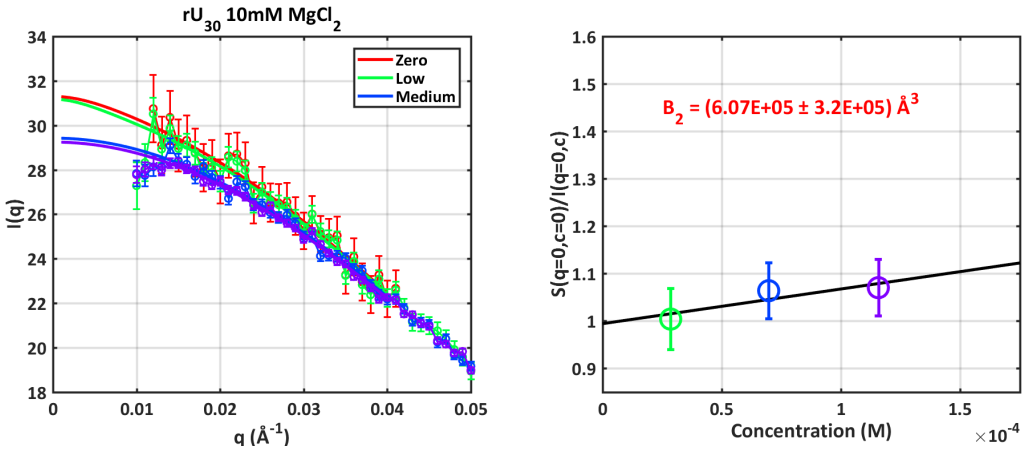



***Figure S19****. Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rU_{30}$ *in 10mM* $MgCl_2$*. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $\boldsymbol{B_2}$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

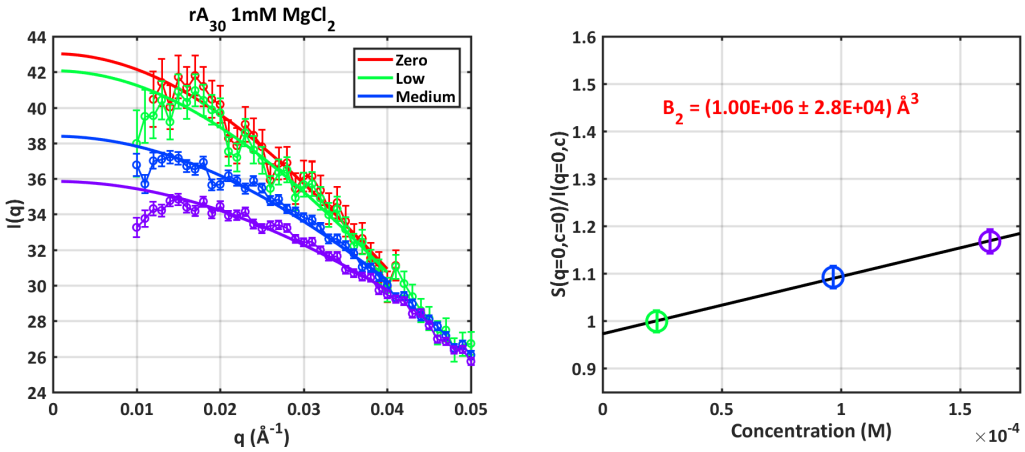



***Figure S20****. Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rA_{30}$ *in 1mM* $MgCl_2$*. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $B_2$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

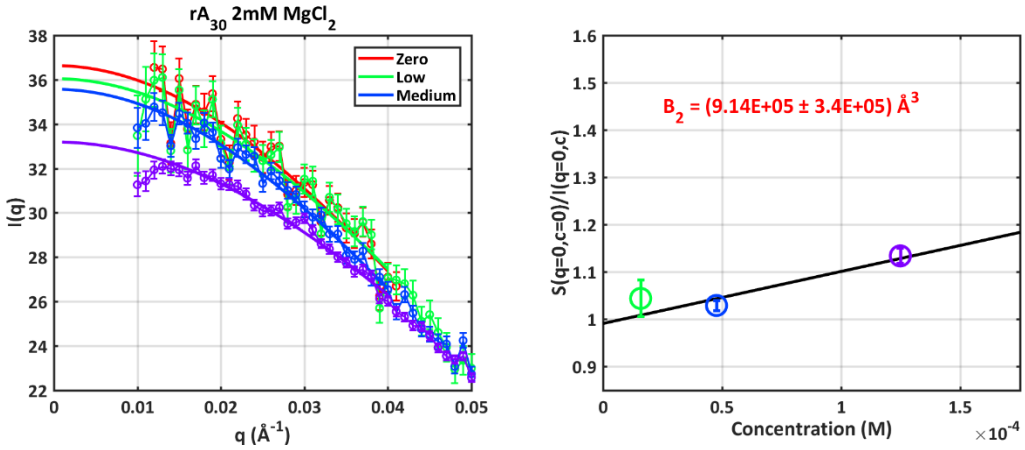



***Figure S21****. Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rA_{30}$ *in 2mM* $MgCl_2$*. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $\boldsymbol{B_2}$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

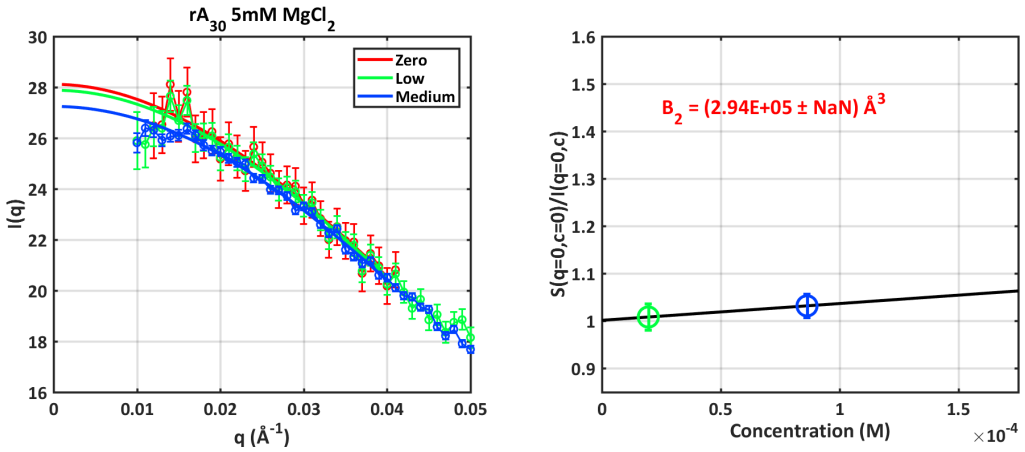



***Figure S22****. Calculation of second virial coefficients.* ***Left:*** *extrapolation of each of the high, medium, low and zero concentration SAXS curves (circles) to the q=0 limit for* $rA_{30}$ *in 5mM* $MgCl_2$*. Fits to the raw SAXS curves (equation 1) are shown as thick solid lines.* ***Right:*** *Linear fit (solid black line) of the ratio* ***S(q=0,c=0)/I(q=0,c)*** *versus nucleic acid concentration* ***c****, yields the second virial coefficient* $\boldsymbol{B_2}$ *through use of equation 2. Errors are propagated from the uncertainties in fit parameters.*

# BE-AES Methods and Analysis

We carried out a total of four separate BE-AES sample preparations, which includes both buffer exchange and dilution. The nucleic acid samples included a 25bp DNA control and two single-stranded RNA constructs ($rU_{30}$ and $rA_{30}$). Buffer samples were also sent for analysis. Each of these preparations occurred on separate days, for a total of four days of preparation.

*Sample preparation 25bp DNA.*

HPLC purified 25 base-pair (bp) DNA oligomers were purchased from Integrated DNA Technologies (Coralville, IA, USA). The strand sequences used for the DNA duplex were: S1 GCATCTGGGCTATAAAGGGCGTCG, and its complement, S2.

The 25bp DNA duplex was formed by resuspending DNA strands S1 and S2 in STE buffer to a final concentration of 400µM. S1 and S2 were then mixed in equimolar concentrations, to yield ~680µl at 200µM concentration. This mixture was annealed at 90°C for 10 minutes, then slow cooled on the bench to room temperature.

Amicon ultra 0.5-10kDa concentrators were used for extensive buffer exchange of the DNA samples with solutions containing the desired amount of $Mg^{2+}$, 20mM NaCl, 10mM NaMOPs (pH 7) and 20µM EDTA. Due to the sample volume used, and the desire to keep nucleic acid concentrations reasonable, two separate concentrators were employed for each buffer condition (buffer conditions are provided in Table S3). All concentrators were pre-washed with the desired buffer before loading nucleic acid sample. To each concentrator used for $Mg^{2+}$ conditions, 85µl of 160µM DNA sample was added, while 70µl of 160µM DNA was added to the $Mg^{2+}$-free concentrators. All samples were then spun for a total of eight rounds. To mitigate effects of evaporation, buffer exchange was conducted at 4°C, and the volume of sample retained in the concentrators was kept above 100µl. After the final round of buffer exchange, the sample and respective flow through was diluted to concentrations appropriate for detection. Dilutions were performed immediately after the final round of buffer exchange. Results for ion numbers are provided in Table S4.

*Sample preparation ssRNA.*

HPLC purified RNA oligomers of $rU_{30}$ and $rA_{30}$ were purchased (deprotected) from Dharmacon (Lafayette, CO, USA).

The lyophilized strands were resuspended in STE buffer to 500µM concentration. Both constructs were annealed at 95°C for 10 minutes, allowed to slow cool on the bench over 20 minutes, and finally placed on ice until loading into concentrators.

Buffer exchange was performed for both constructs in Amicon ultra 0.5-3kdA concentrators. The buffers were identical to those used to prepare the 25 bp construct, described above. Due to the volume of sample used, and the desire to keep nucleic acid concentrations reasonable, two separate concentrators were used for each buffer condition. All concentrators were pre-washed with the desired buffer before loading the nucleic acid sample. To each concentrator, 85µl of 500µM RNA sample was added. All samples were then spun for a total of 8 rounds. To mitigate effects of evaporation, buffer exchange was conducted at 4°C, and the volume of sample retained in the concentrators was kept above 100µl. After the final round of buffer exchange, the sample and respective flow-through was diluted to concentrations appropriate for detection.

Dilutions were performed immediately after the final round of buffer exchange. Results for ion numbers are provided in Tables S5-6.

*Buffer Dilutions*

Buffers were diluted into a volume of 7ml water. Two separate dilutions occurred for each buffer; one set of four samples was diluted by ~200 fold and designed to measure $Na^+$ ppm. A separate set of four was diluted by ~25-60 fold to detect $Mg^{2+}$ in the buffers.

*Sample Dilutions*

Dilution values were assessed gravimetrically using a Miller-Toledo balance. Random repeat weighing was made throughout the process to ensure precision and repeatability in the weighing method. Samples were diluted 35µl into 7ml of water, a ~200 fold dilution. For the sample flow-throughs, identical dilutions to those performed on the sample were used to provide $Na^+$ and $Mg^{2+}$ ppms. After dilution, all tubes were sealed with parafilm and stored on ice.

*Data Acquisition and Analysis.*

In total, four separate dilutions were made for each nucleic acid containing sample (two from each of the two concentrators). Two separate dilutions were prepared for each of the sample flow-throughs (two from each concentrator). For each of these dilutions, three machine readings were carried out. In total therefore, we have twelve readings of the nucleic acid containing sample for each buffer condition, and twelve flow-through measurements.

Concentrations of the counter ions ($Na^+$, $Mg^{2+}$) and nucleic acids (through measurement of phosphorus (P) concentrations) were determined using an Optima 7300DV ICP-AES (Perkin Elmer, Waltham, MA) within the linear detection range of the instrument: multiple emission lines for elements P (two lines), $Na^+$ (three lines) and $Mg^{2+}$ (four lines ), ($Cl^-$ anions were not detected) were monitored, averaged, and integrated intensities were converted to concentration units. This was achieved by calibrating the instrument with seven linearly-spaced standard solutions that spanned the range of concentrations. Two independent calibrations were used when measuring the samples. The number of excess ions ($\Delta N_{Ion}$) per phosphate ($N_P$) was determined for each sample and dilution using:

$$\frac{\Delta N_{ion}}{N_P} = \frac{[Ion]_S - [Ion]_B}{[P]_S - [P]_B} \qquad (3)$$

where the subscripts ***S*** and ***B***, refer to the NA-containing sample and the corresponding buffer flow-through, respectively. [] denote the parts-per-million (ppm) concentration of either ions or phosphates. For this analysis, each reading was treated as an independent measurement.

We used the standard error in the mean of the twelve measurements for each experiment condition as errors for the readings.

Once the excess number of ions per phosphate were determined, we fit a non-cooperative Hill-equation to each sample. For each construct, the excess $Na^+$ and $Mg^{2+}$ numbers were defined by:

$$\frac{\Delta N_{Na}}{N_P} = \frac{F_{Na}}{1 + \left( {[Mg]}/{M_{1/2}} \right)^n} \quad (4)$$

$$\frac{\Delta N_{Mg}}{N_P} = \frac{F_{Mg} \left( {[Mg]}/{M_{1/2}} \right)^n}{1 + \left( {[Mg]}/{M_{1/2}} \right)^n} \quad (5)$$

In these equations, $F_{Na}$ is the number of excess $Na^+$ ions per phosphate in the limit [Mg]→0, $F_{Mg}$ is the number of excess Mg ions per phosphate in the limit [Mg]→∞, $n$ is the Hill coefficients, and $M_{1/2}$ is the competition coefficient.

Equations 4 and 5 were simultaneously fit to Na and Mg ion counts for each construct using non-linear fitting. Data points were weighted by the uncertainty in each measurement, and errors in the fit parameters represent the standard deviation on fit parameters. Results are shown in Figures S23-25, and ion numbers tabulated in Tables S3-6.

Finally, all reported uncertainties pertaining to the number of excess ions are the standard error in the mean (SEM) across 12 measurements.

*Data Availability.*

An excel document containing all measured values (ppms), dilution values and error propagations for every sample catalogued in this work (buffers and nucleic acids) is available from the authors upon reasonable request.

The next section summarizes and tabulates the results of ion counting measurements (Tables S3-6). Comparisons to previously published data are shown in Figures S26-27 and Table S7.

| Buffer | Mean [$Na^+$] (mM) | SEM [$Na^+$] (mM) | Mean [$Mg^{2+}$] (mM) | SEM [$Mg^{2+}$] mM) |
|---|---|---|---|---|
| 20mM NaCl | 22.67 | 0.45 | 0.00 | 0.00 |
| 20mM NaCl + 0.5mM $MgCl_2$ | 23.34 | 0.75 | 0.46 | 0.01 |
| 20mM NaCl + 1mM $MgCl_2$ | 22.62 | 0.61 | 0.96 | 0.01 |
| 20mM NaCl + 2mM $MgCl_2$ | 22.91 | 0.59 | 1.93 | 0.03 |
| 20mM NaCl + 3mM $MgCl_2$ | 22.36 | 0.38 | 2.81 | 0.03 |

***Table S3.*** *The measured mean concentrations of* $Na^+$ *and* $Mg^{2+}$ *of the buffers used in BE-AES experiments with associated standard error in the mean (SEM).*

*Ion numbers 25bp DNA.*

| Buffer | Mean Ions Na | SEM Ions Na | Mean Ions Mg | SEM Ions Mg |
|---|---|---|---|---|
| 20mM NaCl | 35.37 | 3.93 | 0.17 | 0.06 |
| 20mM NaCl + 0.5mM $MgCl_2$ | 21.09 | 1.89 | 9.61 | 0.07 |
| 20mM NaCl + 1mM $MgCl_2$ | 17.64 | 2.49 | 12.07 | 0.12 |
| 20mM NaCl + 2mM $MgCl_2$ | 13.62 | 2.80 | 14.28 | 0.22 |
| 20mM NaCl + 3mM $MgCl_2$ | 11.68 | 1.85 | 15.64 | 0.28 |

***Table S4.*** *The measured mean number of ions per molecule around 25bp DNA with associated standard error in the mean (SEM).*

*Ion numbers $rU_{30}$.*

| Buffer | Mean Ions Na | SEM Ions Na | Mean Ions Mg | SEM Ions Mg |
|---|---|---|---|---|
| 20mM NaCl | 24.21 | 1.28 | 0.01 | 0.01 |
| 20mM NaCl + 0.5mM $MgCl_2$ | 20.99 | 0.75 | 4.29 | 0.04 |
| 20mM NaCl + 1mM $MgCl_2$ | 15.35 | 1.70 | 6.16 | 0.08 |
| 20mM NaCl + 2mM $MgCl_2$ | 11.53 | 1.49 | 7.97 | 0.08 |
| 20mM NaCl + 3mM $MgCl_2$ | 6.26 | 1.57 | 8.97 | 0.13 |

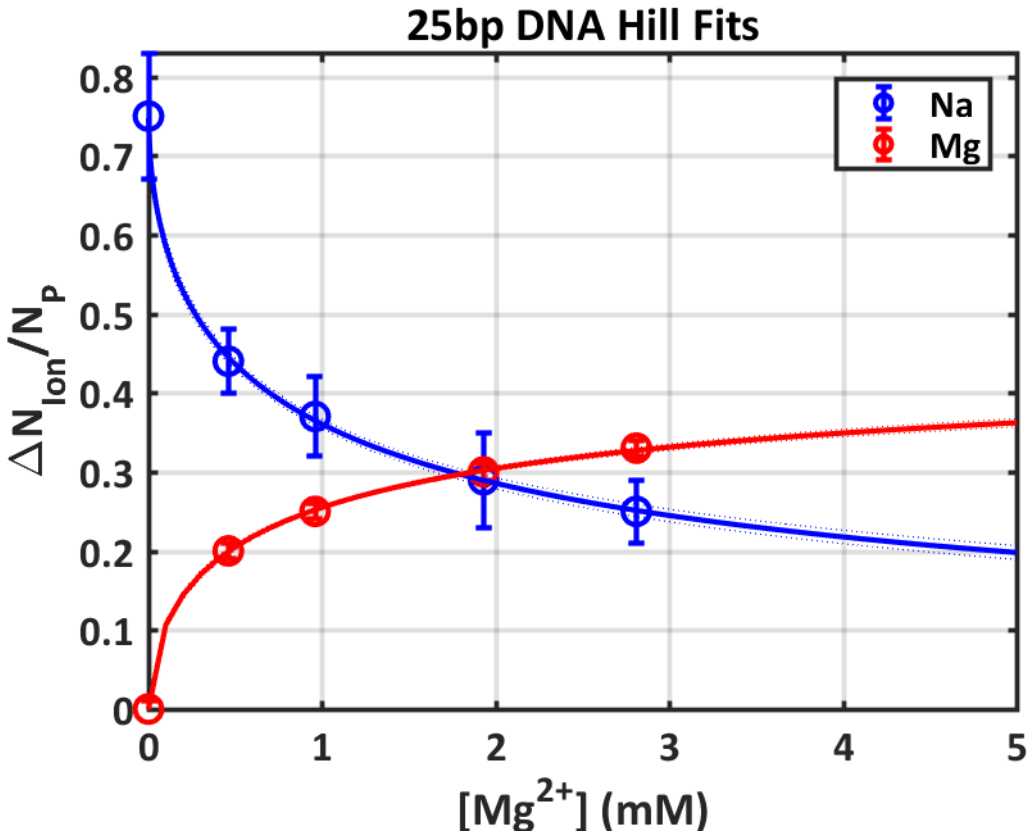
***Table S5.*** *The measured mean number of ions per molecule around $rU_{30}$ with associated standard error in the mean (SEM).*

*Ion numbers $rA_{30}$.*

| Buffer | Mean Ions Na | SEM Ions Na | Mean Ions Mg | SEM Ions Mg |
|---|---|---|---|---|
| 20mM NaCl | 23.86 | 1.43 | 0.11 | 0.07 |
| 20mM NaCl + 0.5mM $MgCl_2$ | 15.32 | 0.89 | 6.34 | 0.08 |
| 20mM NaCl + 1mM $MgCl_2$ | 11.62 | 1.05 | 7.80 | 0.07 |
| 20mM NaCl + 2mM $MgCl_2$ | 8.34 | 1.48 | 9.29 | 0.13 |
| 20mM NaCl + 3mM $MgCl_2$ | 8.72 | 0.88 | 10.46 | 0.16 |

***Table S6.*** *The measured mean number of ions per molecule around $rA_{30}$ with associated standard error in the mean (SEM).*

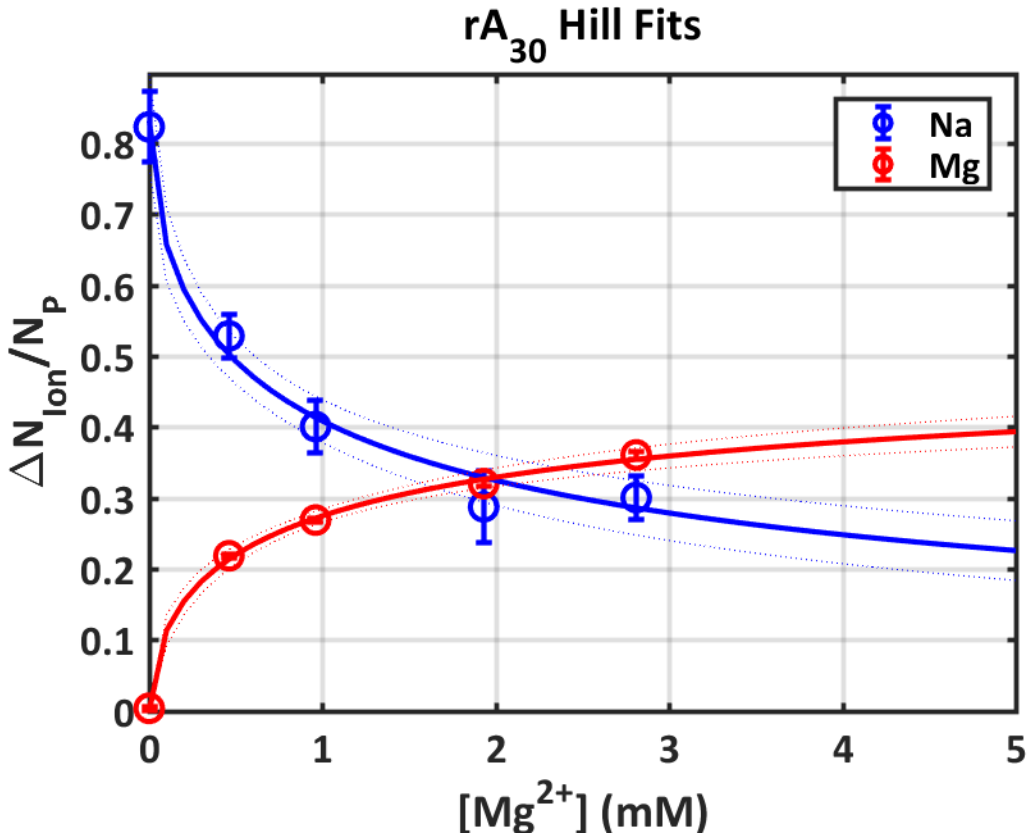
*Hill fits to 25bp DNA.*




| Construct | $F_{Na}$ | $F_{Mg}$ | n | M |
|---|---|---|---|---|
| dsDNA | 0.75±0.01 | 0.49±0.01 | 0.59±0.01 | 0.89±0.03 |

***Figure S23****. 25bp DNA ion counting data (circles) fit simultaneously to equations 4 and 5 (solid lines). Dashed lines represent 95% confidence intervals on fit parameters. Table displays values determined for fit parameters ± one standard deviation.*

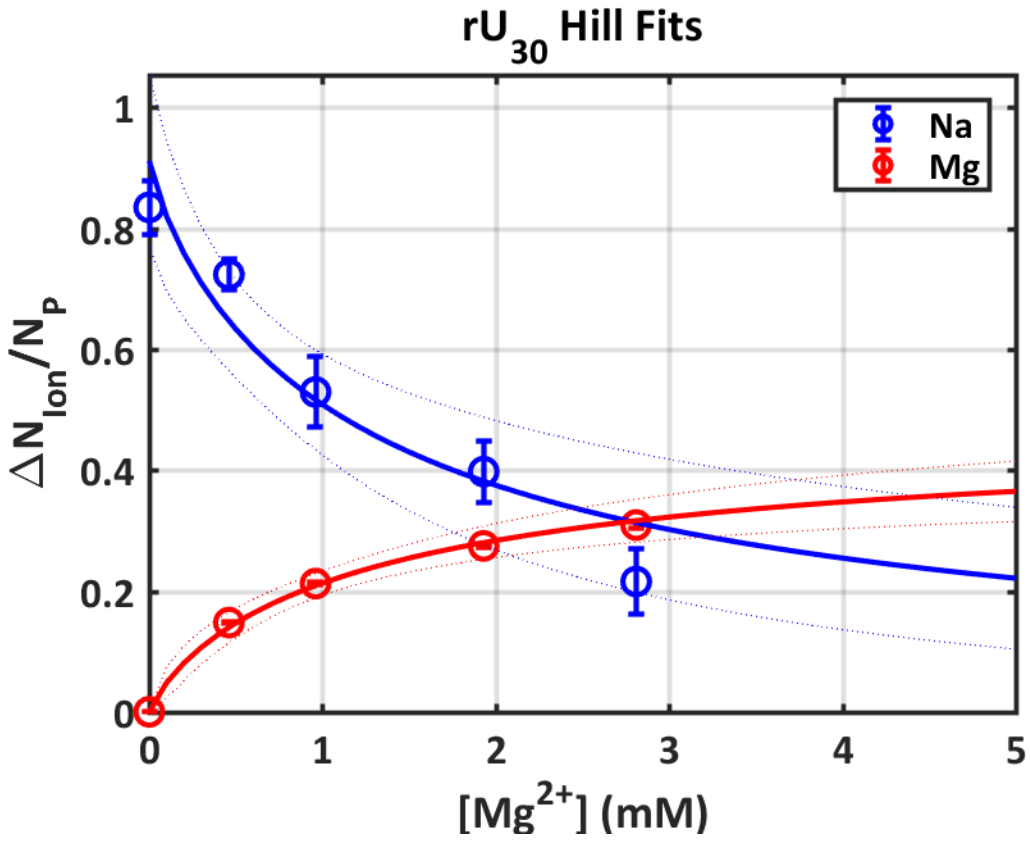
*Hill fits to $rA_{30}$.*




| Construct | $F_{Na}$ | $F_{Mg}$ | n | M |
|---|---|---|---|---|
| $rA_{30}$ | 0.83±0.02 | 0.54±0.01 | 0.59±0.04 | 0.94±0.13 |

***Figure S24**. Ion counting data for $rA_{30}$ (circles) fit simultaneously to equations 4 and 5 (solid lines). Dashed lines represent 95% confidence intervals on fit parameters. Table displays values determined for fit parameters ± one standard deviation.*

*Hill fits to $rU_{30}$*




| Construct | $F_{Na}$ | $F_{Mg}$ | n | M |
|---|---|---|---|---|
| $rU_{30}$ | 0.91±0.05 | 0.48±0.02 | 0.85±0.10 | 1.31±0.32 |

***Figure S25**. Ion counting data for $rU_{30}$ (circles) fit simultaneously to equations 4 and 5 (solid lines). Dashed lines represent 95% confidence intervals on fit parameters. Table displays values determined for fit parameters ± one standard deviation.*

*Comparison to other BE-AES measurements.*

Below, we compare the results of other ion counting measurements on short (24-25bp) DNA using BE-AES. The current work (blue), our former work (green) [3] and measurements from another laboratory (red) [4]. For both our works, the measured data plus minus the SEM are reproduced. For [4], data points were calculated from the fit parameters provided in the supporting information, we refrain from extrapolating errors on these calculated values for clarity (for full information see reference). Both our measurements were performed on 25bp DNA, while other labs were performed on 24bp DNA. Given that we are interested in the excess ions on a per phosphate basis, these measurements should be comparable.

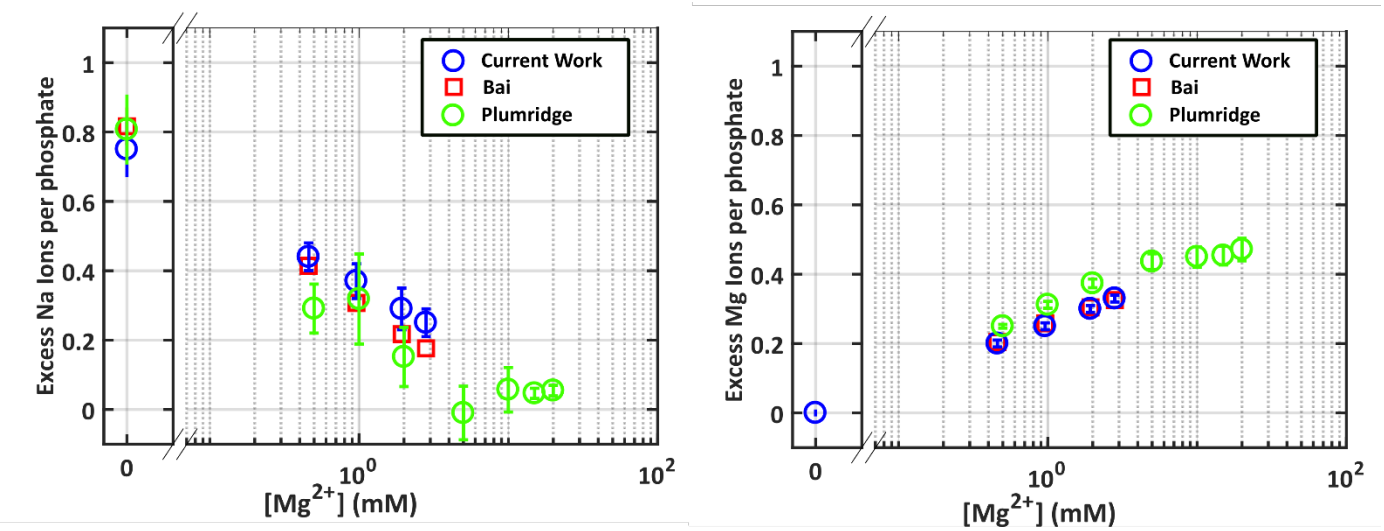



*Figure S26. Comparison of the number of excess ions per phosphate around 25bp DNA; this current work (blue), our former work (green), and around 24bp DNA; measurements from another laboratory (red).*

*Comparison to ssDNA.*

Below we compare the results of our ion counting experiments for ssRNA (this work) to our previous measurements on ssDNA [3]. In the latter identical ion atmospheres between poly dT and poly dA were found. Therefore, we use only one set of data to summarize the ssDNA results.

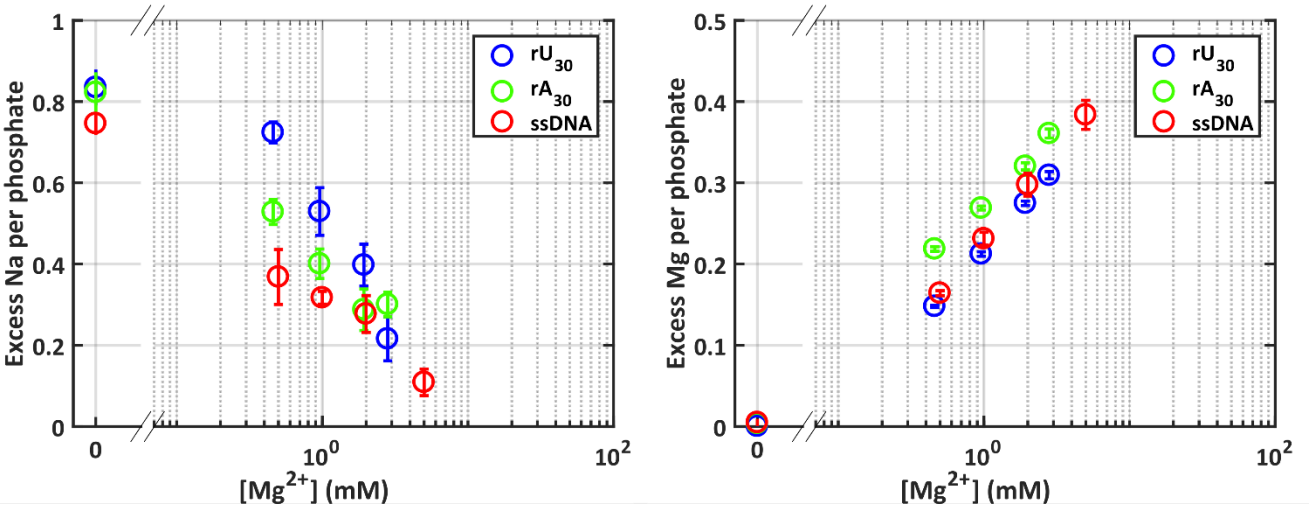



***Figure S27****. Comparison between ion counting results for ssRNA as measured in this study, to values we previously measured for ssDNA. The error bars represent the SEM across all repeats.*

| Construct | $F_{Na}$ | $F_{Mg}$ | n | M |
|---|---|---|---|---|
| dsDNA | 0.75±0.01 | 0.49±0.01 | 0.59±0.01 | 0.89±0.03 |
| $rA_{30}$ | 0.83±0.02 | 0.54±0.01 | 0.59±0.04 | 0.94±0.13 |
| $rU_{30}$ | 0.91±0.05 | 0.48±0.02 | 0.85±0.10 | 1.31±0.32 |
| $*dT_{30}$ | 0.68±0.02 | 0.49±0.01 | 0.88±0.02 | 1.09±0.04 |
| $*dA_{30}$ | 0.71±0.01 | 0.50±0.01 | 0.84±0.01 | 1.13±0.03 |

***Table S7.*** *Comparing fit parameters derived from application of equations 4 and 5 to the ion counting data for 25bp DNA, ssRNA and previously measured ssDNA.*

*Errors in fit parameters calculated through bootstrapping.

*Total compensated charge of ssRNAs.*

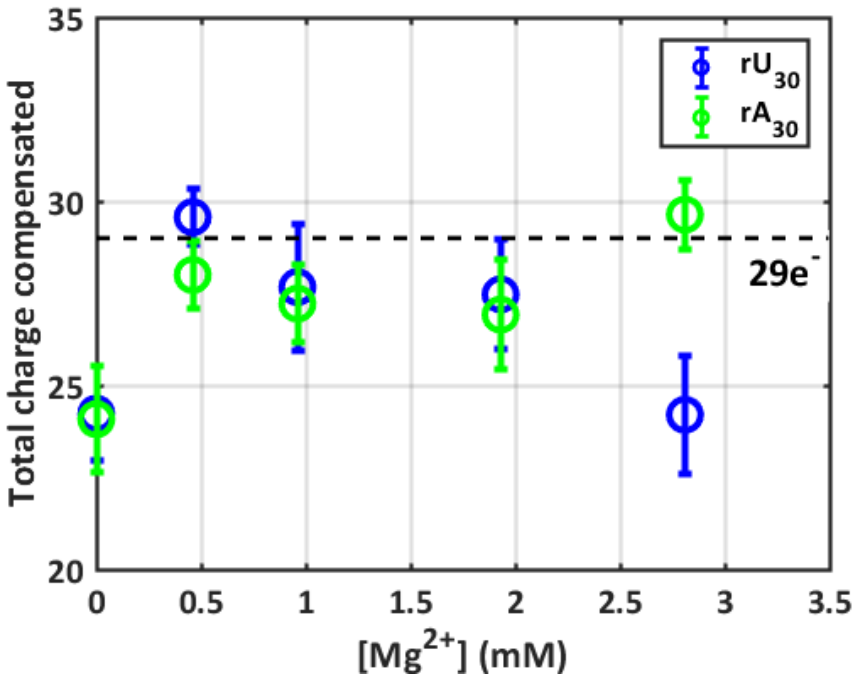



***Figure S28***. *This figure shows the total polymer charge compensated by excess (measured) cations. As is the case for duplex DNA (shown in Figure S23), neutralization of the poly rA/poly rU charge (29e$^-$) also requires exclusion of anions. These data suggest that anion exclusion contributes to overall neutralization of these lower charge density polymers. Consistent with other studies, more anions are excluded in pure monovalent atmospheres than in mixed ion atmospheres (e.g. Supp Ref (4)).*

# Modelling Methods and Analysis

The analysis method is extensively described, outlined and validated in reference [5]. Slight modifications to this procedure and parameters were required to work with ssRNA bases. This section aims to describe the changes made to the method previously prescribed. All remaining details are identical to those previously reported.

*Definition of crystal structure parameters.*

The definition of crystal structure parameters (Figure S28) are consistent with the definitions and terminology in reference [6]. Below we illustrate the torsion angles and nomenclature used here to define the phosphate backbone and base components of a nucleic acid. We describe the phosphate backbone by five torsion angles ***{ε ζ α β γ}***, and the corresponding bases by two ***{δ, χ}***.

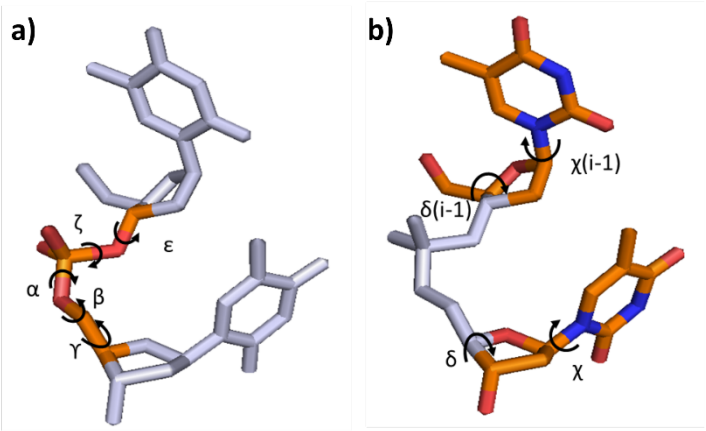



*Figure S29. (a) Torsion angles defining the route of the phosphate backbone, and (b) associated bases.*

*Crystal structure parameters.*

Firstly, basic dinucleotide step parameters, including base sugar puckers ($\delta$) and base torsion angles ($\chi$), were re-defined to be relevant for RNA nucleotides. As for the ssDNA case, we make approximations in the steps to simplify the procedure. The base torsion angles ($\chi$) were taken to always be anti, as it is apparent from the conformational survey of dinucleotides steps [6] that this RNA conformation is overwhelmingly preferred over the syn geometry.

To obtain the specific torsion angles for rA and rU nucleotides (nt), we used a slightly older dinucleotide survey [7], as this paper includes an extensive document containing all the data used to generate the survey. Base torsion angles were set at the average value of all dinucleotide steps for either rA or rU in the survey.

In rA and rU, both sugar puckers are populated frequently, and were therefore modelled. Revised values for sugar puckers were calculated using the survey mentioned above.

The sugar pucker equilibrium constants ($K_{c2}$) between C2' and C3' endo pucker conformations were derived from: [8]. The specific values associated with each change mentioned above are documented in Table S8.

| nt | δ C2'-endo | δ C3'-endo | X C3'-endo | X C2'-endo | $K_{c2}$ |
|---|---|---|---|---|---|
| rA | 146.6 | 81.1 | 199 | 238 | 0.71 |
| rU | 146.6 | 81.1 | 198 | 228 | 0.77 |

***Table S8.*** *Values of torsion angles used to model rA and rU nts. Sugar puckers for C2' endo and C3' endo are given, as well as the associated base torsion angles for each of these sugar puckers. The ratio of C2' to C3' endo sugar pucker conformations is also given.*

The dinucleotide step library for rA and rU is provided in Table S9. Each 'step' contains two nts and we used a suite definition to prescribe the conformations. The 'mnemonics' shown in the chart describe the average step conformations and are derived from [6]. The step definitions from which the torsion angles are derived in the crystal structure survey are tabulated in the second column of the table. The next few columns list the specific values assigned to the torsion angles ***{ε ζ α β ϒ}***. In the next columns (rU,rA) we indicate whether a particular step (row) is present in the refinements for that molecule, including whether the step is sterically allowed. The next column provides all restrictions on sugar puckers, denoting which sugar puckers are permitted for each suite. This information can vary as a result of steric hindrances, e.g. for A1 (A form) geometries, both sugar puckers in the suite are C3' endo. Finally, the rightmost column provides insight into the stacked geometries.

| Mnemonic | Derived from | ε | ζ | α | β | ϒ | rU | rA | Restrictions on {δ(i-1), δ} | Note |
|---|---|---|---|---|---|---|---|---|---|---|
| **A1p** | 5d,6d,6p | 234 | 81 | 65 | 159 | 53 | ✓ | ✓ | None | |
| **A1t** | 5j,5q,6j | 224 | 72 | 67 | 115 | 178 | ✓ | ✓ | None | |
| **A2p** | 4p,4d,7p,3d | 248 | 206 | 72 | 197 | 57 | ✓ | ✓ | None | |
| **P2p** | 4b | 229 | 168 | 292 | 171 | 47 | ✓ | ✓ | None | |
| **P3p** | 2a,2l,1l,1m | 239 | 290 | 292 | 212 | 55 | ✓ | ✓ | None | |
| **P3m** | 2o,1o | 236 | 291 | 292 | 210 | 294 | ✓ | ✓ | None | |
| **P3t** | 2h | 261 | 290 | 296 | 177 | 176 | ✓ | ✓ | None | |
| **S3p** | 1g,1z | 213 | 285 | 182 | 161 | 51 | ✓ | ✓ | {3,2} {2,3} | |
| **S3t** | 1t | 199 | 289 | 180 | 195 | 178 | ✓ | ✓ | {3,2} {2,3} | |
| **A1** | AI | 205 | 285 | 294 | 172 | 55 | × | ✓ | {3,3} | Canonical A-form. |
| **AII** | 1c | 197 | 291 | 153 | 194 | 179 | × | ✓ | {3,3} | Turn in GNRA tetraloop. |
| **4b** | 4b | 245 | 163 | 294 | 172 | 46 | × | ✓ | {3,3} | Exotic stacked geometry. |

***Table S9.*** *Summary of parameters defining the dinucleotide suite library used to model ssRNA conformations.*

*Fitting procedure.*

For each experimental condition, numerous refinement rounds were performed. In each round, a pool of 1000 structures was generated. For every member in the pool, SAXS profiles were calculated with CRYSOL [9], using a maximum harmonic order of 15, Fibonacci grid of order 18 and default hydration parameters.

Selection of structures from this pool that best match the SAXS data was then performed with GAJOE 1.3 [10], using an ensemble size of 20 and with repeat selections allowed. The algorithm was run for 50 generations and repeated 50 times.

We checked for convergence using the reduced chi-square parameter. First, the goodness of fit $\chi 2$ of each individual ensemble is assessed by comparing the ensemble $I_{ens}$ and experimentally derived $I_{exp}$ SAXS curves:

$$\chi^2 = \frac{1}{K-1}\sum_{i=1}^{K}\left(\frac{I_{exp}(q_i) - cI_{ens}(q_i)}{\sigma_{exp}(q_i)}\right)^2 \quad (6)$$

where $K$ is the total number of points in q-space, $\sigma_{exp}$ is the experimental error at each $q$ point and $c$ is a scaling factor. The reduced chi-square is then used to judge the global fit of all ensembles to the experimental data:

$$\chi_{Red}{}^2 = \frac{1}{N}\sum_{j=1}^{N}\chi_j{}^2 \quad (7)$$

Once the value of the reduced chi-square had converged, 10 additional rounds of refinements were performed. These additional rounds account for the variations in step weights that produce chain models that adequately fit the data. The final calculated parameters are derived from these 10 additional rounds once convergence had been met, and differences between fits (assess through eq. 7) to the data deemed negligible. For example, the quoted persistence lengths in the main text are the mean persistence lengths of all ensemble models (selected from the 10 additional rounds), ± the standard deviation across all these models. This process was repeated for each experimental condition independently, and results are shown in Figures S29-38.

The following panels show the refinement procedure and the fit of the selected models to the experimental data. Additionally, a landscape of all selected models is produced cast in a space of the radius of gyration ($R_g$) and end-to-end distance (R). In these landscapes, each selected model represents a point located by its $R_g$ and R value. Heat on the map shows density of structures within a 2Å radius from each point, and 1d projections of the landscape onto $R_g$ and R axis shown. The purple dashed contour represents the bounding states of models available in the pool. Eight randomly selected models refined from the SAXS data are shown as examples. These models are shown to illustrate the themes in each solution condition, and do not represent a 'final' or 'complete' set of structures that fit the data.

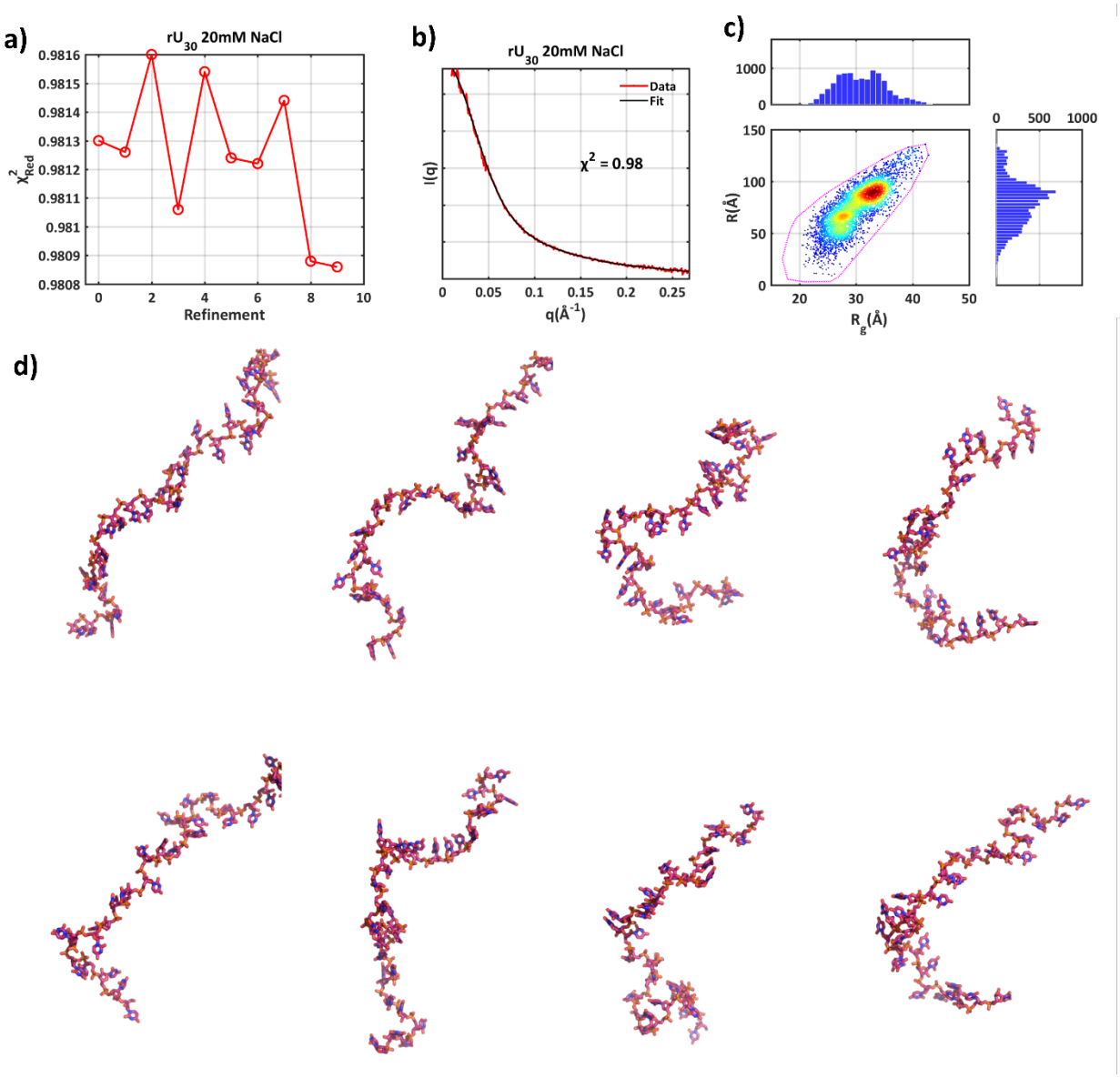



***Figure S30.*** *Ensemble optimization results for $rU_{30}$ in 20mM NaCl.* ***(a)*** *Evolution of $X^2$ values as a function of refinement number* ***(b)*** *Fit of converged ensemble (black) to SAXS data (colored).* ***(c)*** *Map of all selected structures for this solution condition in $R_g$-R space.* ***(d)*** *8 randomly selected models refined from the SAXS data. Note that these models are shown to illustrate the type of structures refined from the data, and do not constitute a complete set. Models derived from refinement rounds 0-9.*

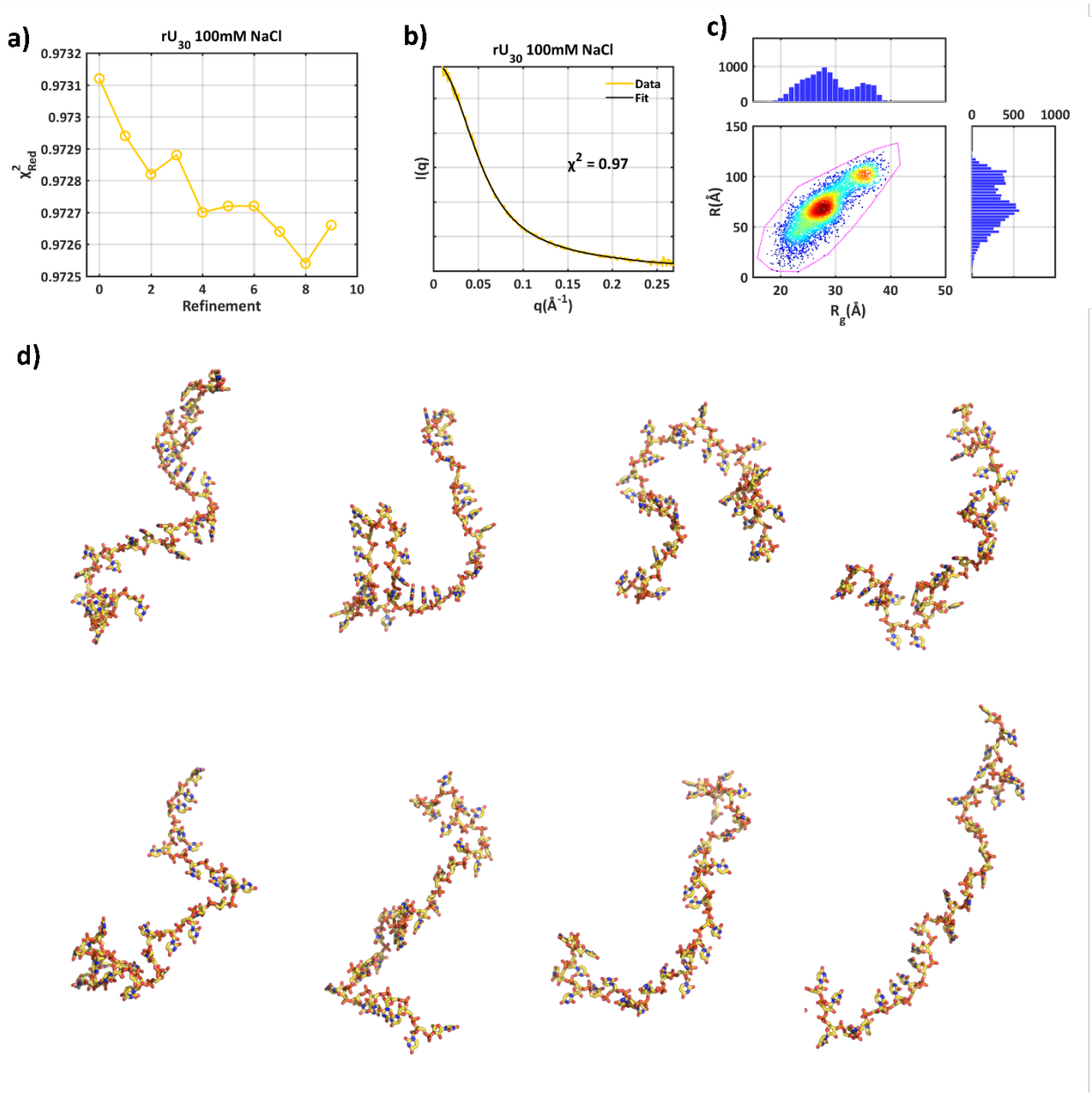



***Figure S31.*** *Ensemble optimization results for $rU_{30}$ in 100mM NaCl.* ***(a)*** *Evolution of $X^2$ values as a function of refinement number* ***(b)*** *Fit of converged ensemble (black) to SAXS data (colored).* ***(c)*** *Map of all selected structures for this solution condition in $R_g$-R space.* ***(d)*** *8 randomly selected models refined from the SAXS data. Note that these models are shown to illustrate the type of structures refined from the data, and do not constitute a complete set. Models derived from refinement rounds 0-9.*

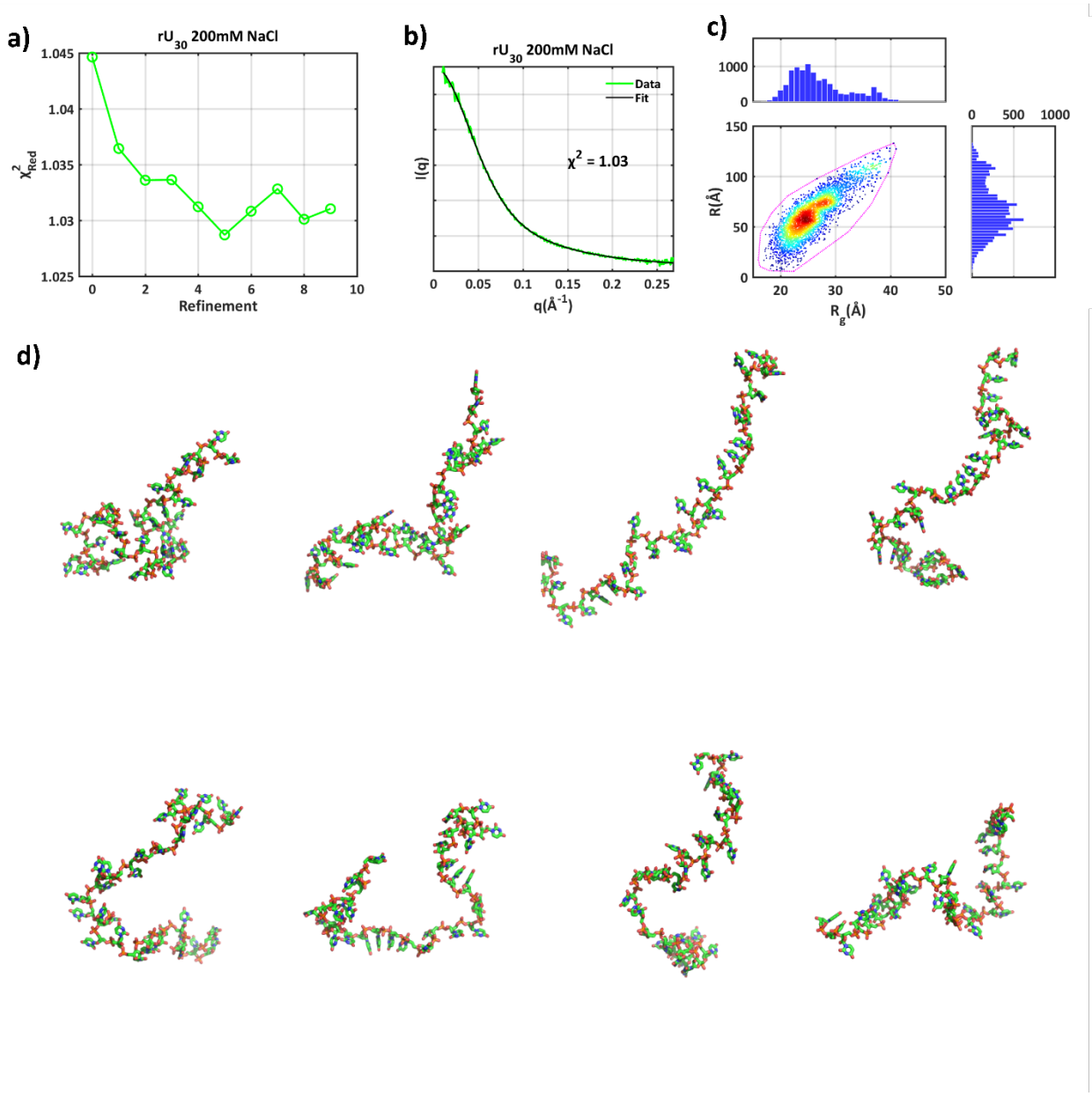



***Figure S32.*** *Ensemble optimization results for $rU_{30}$ in 200mM NaCl.* ***(a)*** *Evolution of $X^2$ values as a function of refinement number* ***(b)*** *Fit of converged ensemble (black) to SAXS data (colored).* ***(c)*** *Map of all selected structures for this solution condition in $R_g$-R space.* ***(d)*** *8 randomly selected models refined from the SAXS data. Note that these models are shown to illustrate the type of structures refined from the data, and do not constitute a complete set. Models derived from refinement rounds 0-9.*

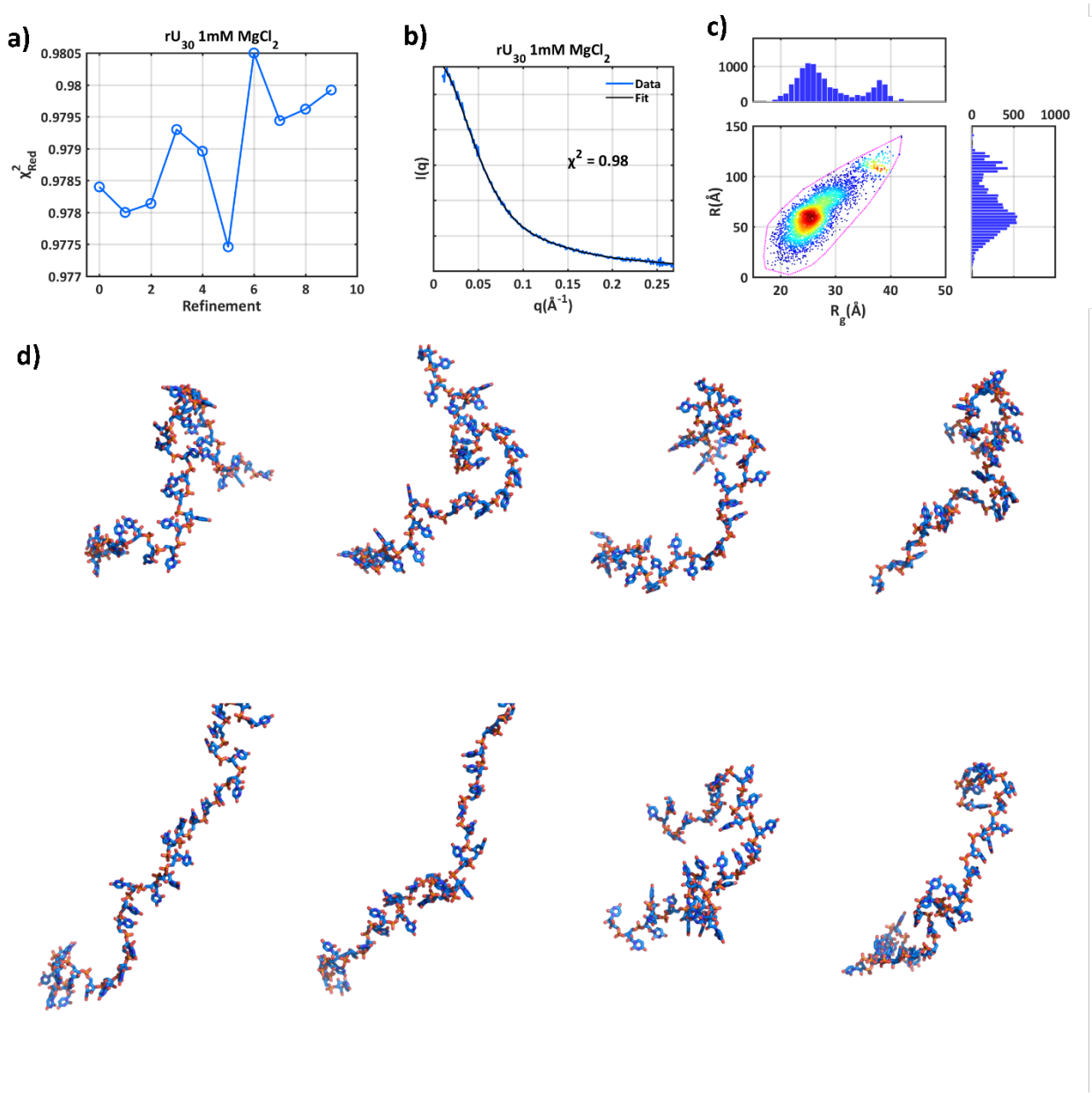



***Figure S33.*** *Ensemble optimization results for $rU_{30}$ in 1mM $MgCl_2$.* ***(a)*** *Evolution of $X^2$ values as a function of refinement number* ***(b)*** *Fit of converged ensemble (black) to SAXS data (colored).* ***(c)*** *Map of all selected structures for this solution condition in $R_g$-R space.* ***(d)*** *8 randomly selected models refined from the SAXS data. Note that these models are shown to illustrate the type of structures refined from the data, and do not constitute a complete set. Models derived from refinement rounds 0-9.*

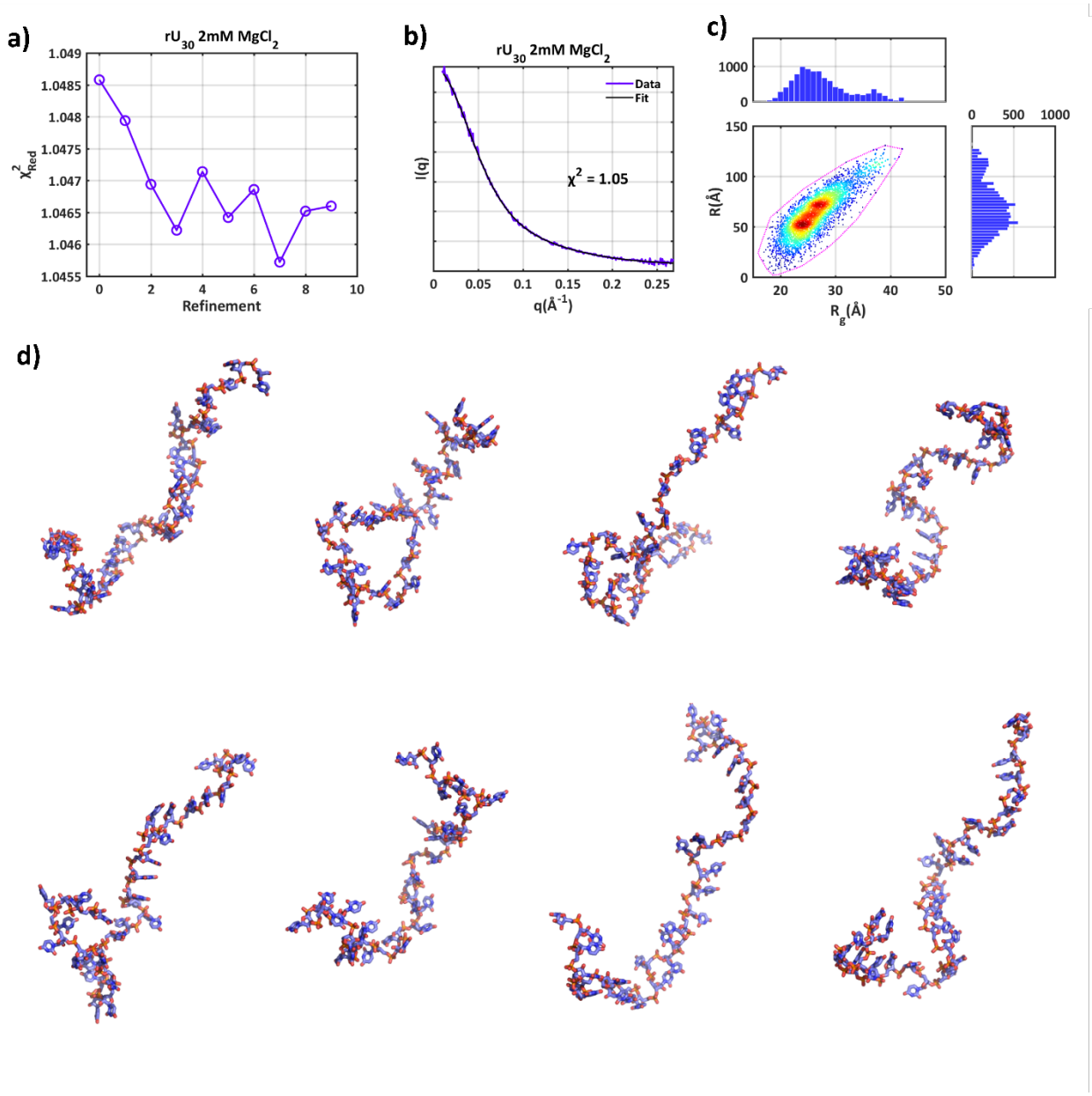



***Figure S34.*** *Ensemble optimization results for $rU_{30}$ in 2mM $MgCl_2$.* ***(a)*** *Evolution of $X^2$ values as a function of refinement number* ***(b)*** *Fit of converged ensemble (black) to SAXS data (colored).* ***(c)*** *Map of all selected structures for this solution condition in $R_g$-R space.* ***(d)*** *8 randomly selected models refined from the SAXS data. Note that these models are shown to illustrate the type of structures refined from the data, and do not constitute a complete set. Models derived from refinement rounds 0-9.*

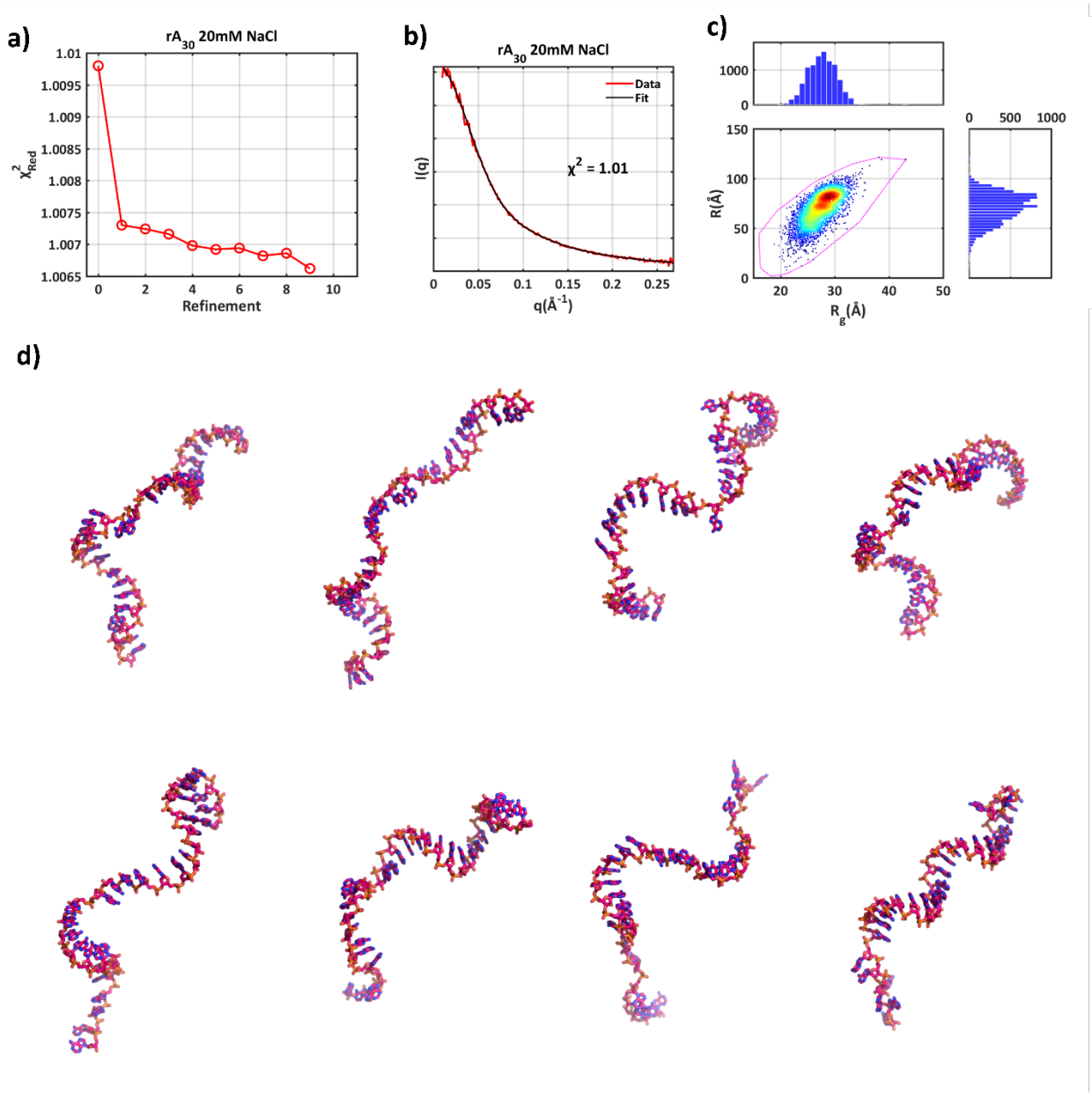



***Figure S35.*** *Ensemble optimization results for $rA_{30}$ in 20mM NaCl.* ***(a)*** *Evolution of $X^2$ values as a function of refinement number* ***(b)*** *Fit of converged ensemble (black) to SAXS data (colored).* ***(c)*** *Map of all selected structures for this solution condition in $R_g$-R space.* ***(d)*** *8 randomly selected models refined from the SAXS data. Note that these models are shown to illustrate the type of structures refined from the data, and do not constitute a complete set. Models derived from refinement rounds 0-9.*

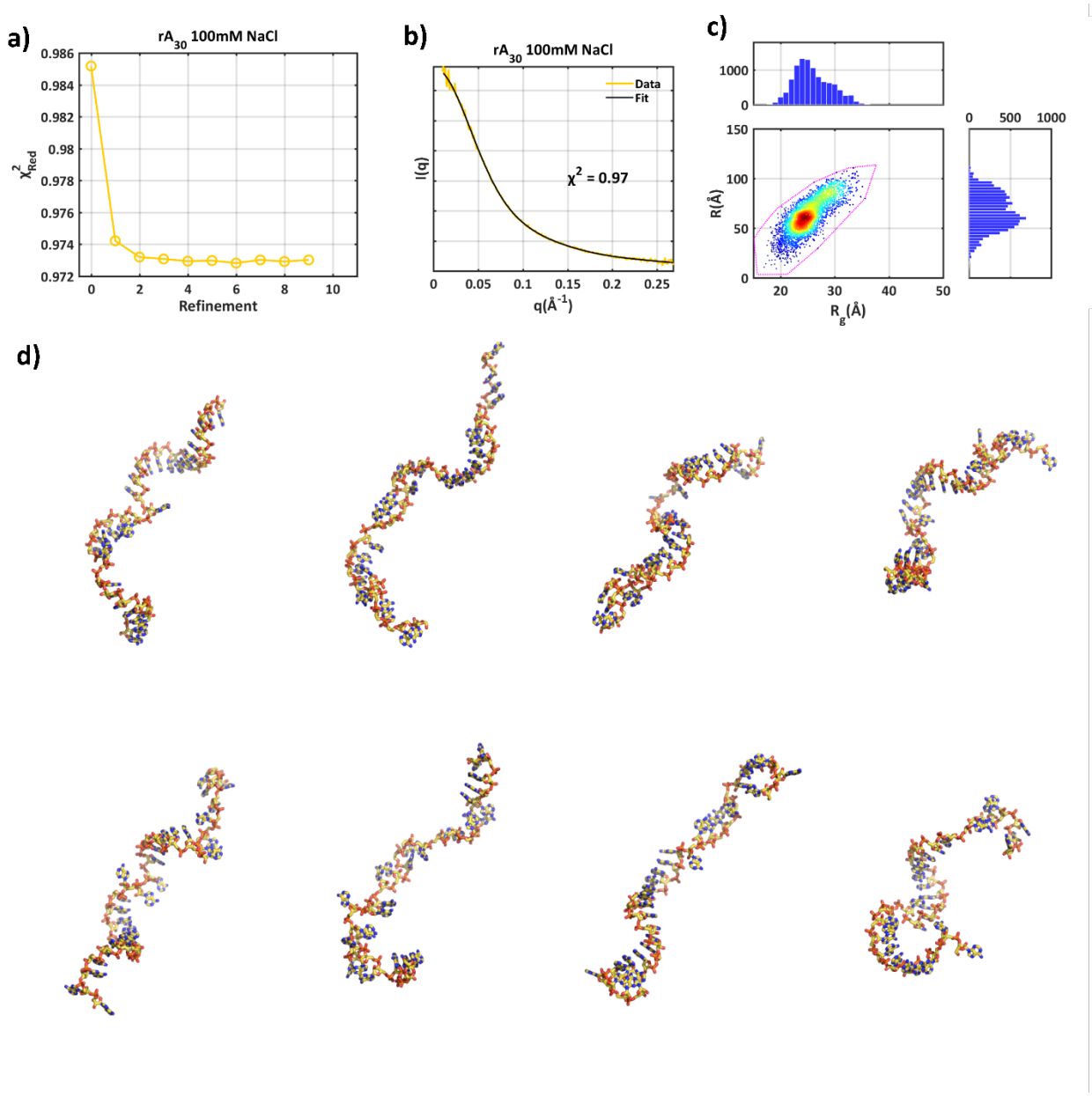



***Figure S36.*** *Ensemble optimization results for $rA_{30}$ in 100mM NaCl.* ***(a)*** *Evolution of $X^2$ values as a function of refinement number* ***(b)*** *Fit of converged ensemble (black) to SAXS data (colored).* ***(c)*** *Map of all selected structures for this solution condition in $R_g$-R space.* ***(d)*** *8 randomly selected models refined from the SAXS data. Note that these models are shown to illustrate the type of structures refined from the data, and do not constitute a complete set. Models derived from refinement rounds 0-9.*

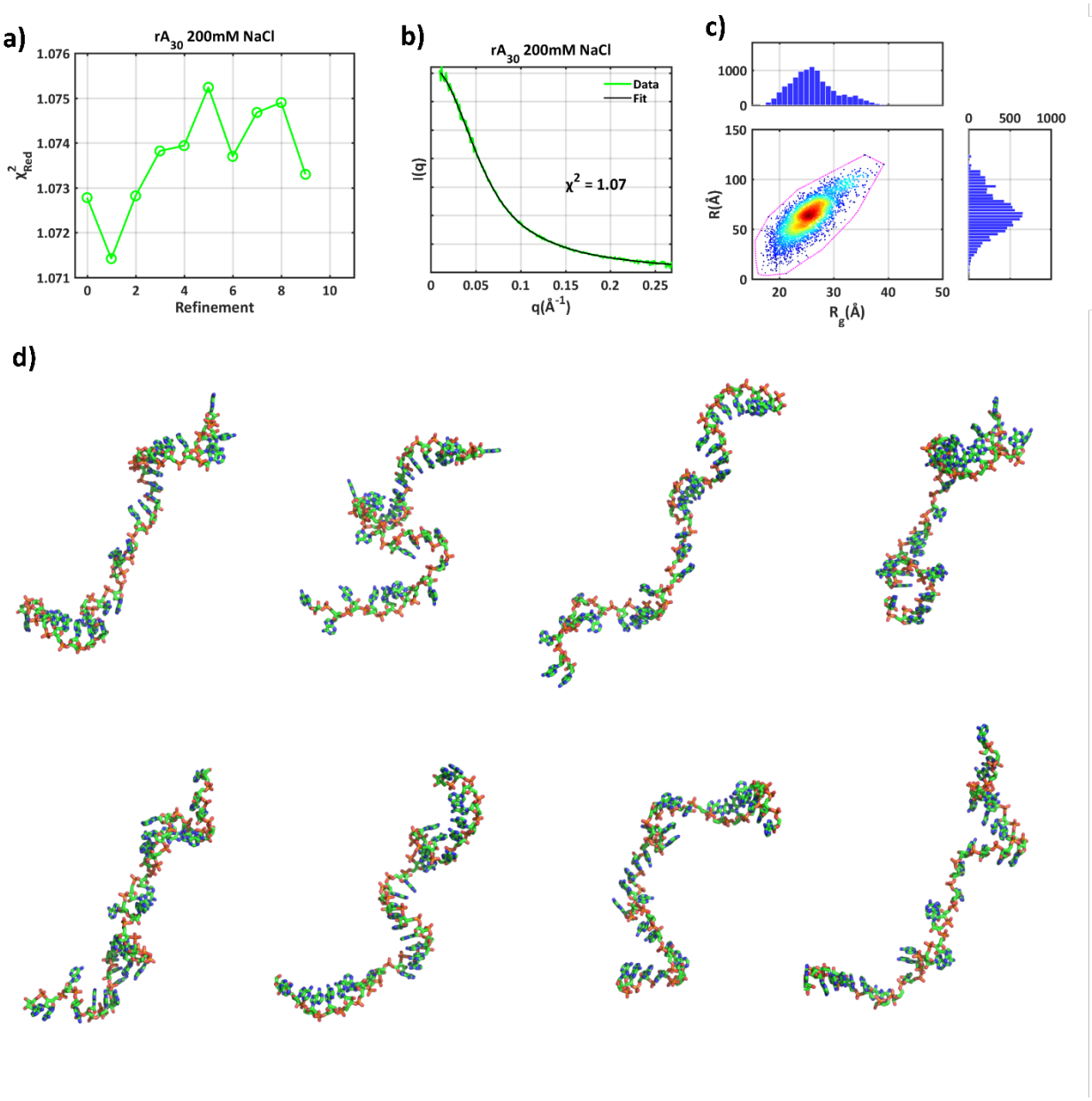



***Figure S37.*** *Ensemble optimization results for $rA_{30}$ in 200mM NaCl.* ***(a)*** *Evolution of $X^2$ values as a function of refinement number* ***(b)*** *Fit of converged ensemble (black) to SAXS data (colored).* ***(c)*** *Map of all selected structures for this solution condition in $R_g$-R space.* ***(d)*** *8 randomly selected models refined from the SAXS data. Note that these models are shown to illustrate the type of structures refined from the data, and do not constitute a complete set. Models derived from refinement rounds 0-9.*

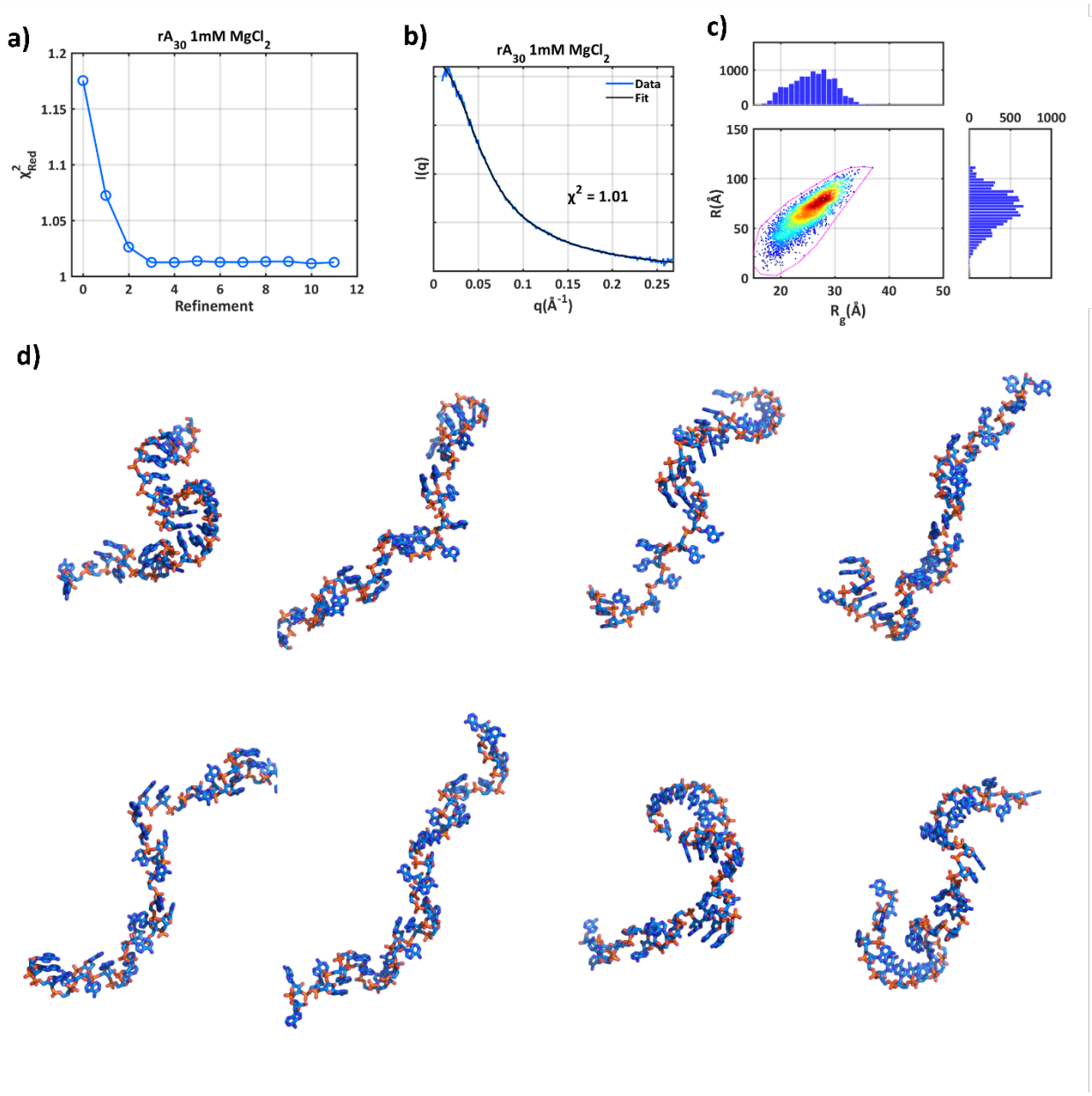



***Figure S38.*** *Ensemble optimization results for $rA_{30}$ in 1mM $MgCl_2$.* ***(a)*** *Evolution of $X^2$ values as a function of refinement number* ***(b)*** *Fit of converged ensemble (black) to SAXS data (colored).* ***(c)*** *Map of all selected structures for this solution condition in $R_g$-R space.* ***(d)*** *8 randomly selected models refined from the SAXS data. Note that these models are shown to illustrate the type of structures refined from the data, and do not constitute a complete set. Models derived from refinement rounds 2-11.*

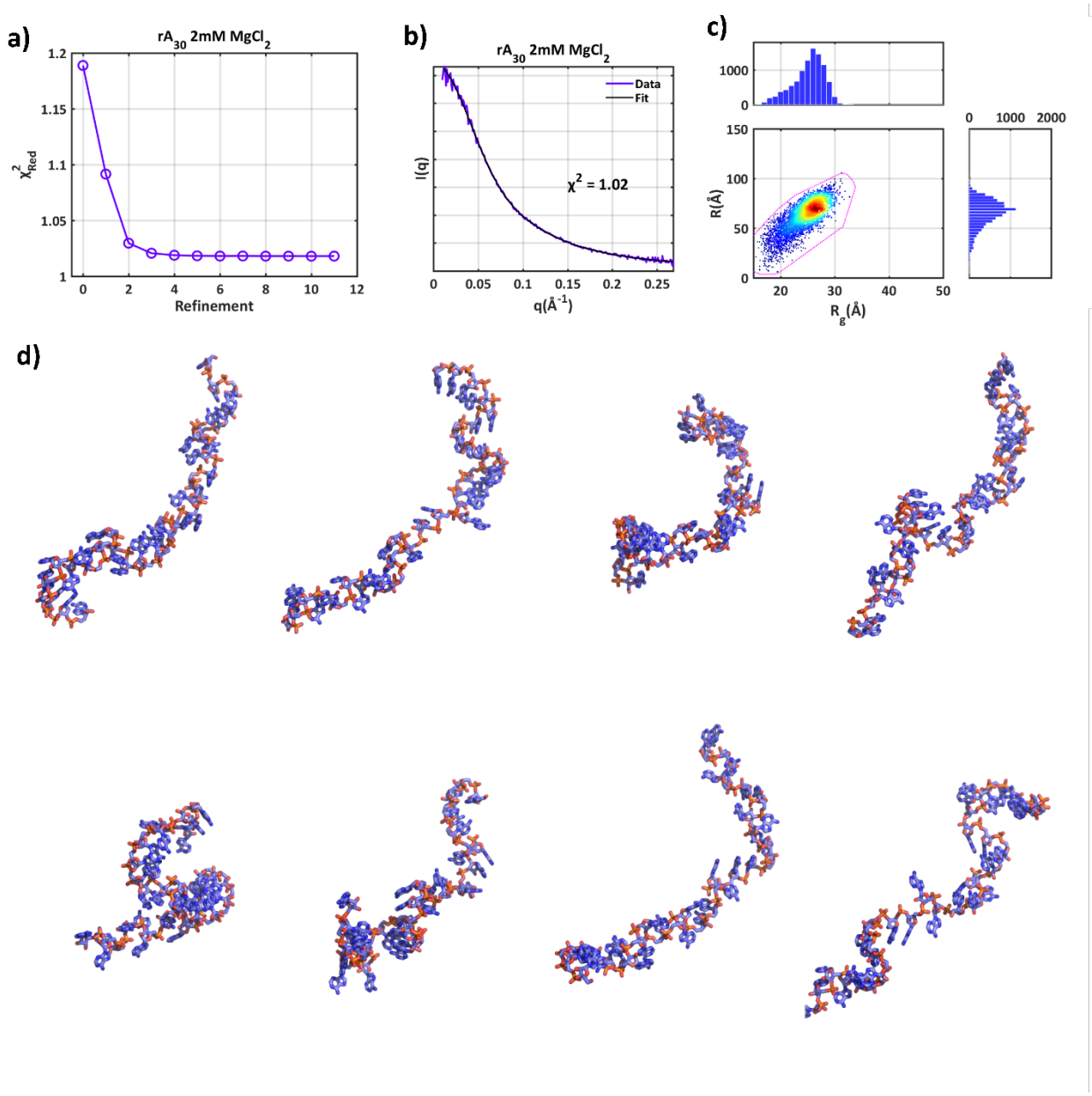



***Figure S39.*** *Ensemble optimization results for $rA_{30}$ in 2mM $MgCl_2$.* ***(a)*** *Evolution of $X^2$ values as a function of refinement number* ***(b)*** *Fit of converged ensemble (black) to SAXS data (colored).* ***(c)*** *Map of all selected structures for this solution condition in $R_g$-R space.* ***(d)*** *8 randomly selected models refined from the SAXS data. Note that these models are shown to illustrate the type of structures refined from the data, and do not constitute a complete set. Models derived from refinement rounds 2-11.*

*Chain parameter calculations – OCFs and correlation lengths.*

To quantify the directional persistence of chains, we calculate the orientation correlation function (OCF), defined as:

$$OCF = < cos\,\theta_{ij} > = < \hat{r}_i \cdot \hat{r}_j > \qquad (8)$$

Here, $\hat{\boldsymbol{r}}_{\boldsymbol{i}}$ is the normalized bond vector between the $i^{th}$ and $(i+1)^{th}$ phosphate in the chain (Figure S39). The average dot product between bond vectors is computed as a function of separation distance along the chain (|i-j|). The average OCF of all members in the selected ensembles is used to interpret the mean shape of the molecule for a given condition.

Correlation lengths were calculated using definition 1 in reference [11], using the OCF in place of persistence lengths traditionally derived from worm-like chain (WLC) models:

$$l_{OCF} = b \sum_{ij}^{N-1} < cos\,\theta_{ij} > \qquad (9)$$

where ***b*** is the bond length (phosphate-phosphate distance for each step).

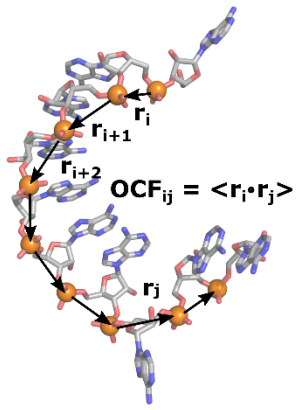



***Figure S40.*** *Calculation of the OCF using bond vectors between neighboring phosphates.*

## Supplemental References


(1) Svergun, D. I. *J. Appl. Crystallogr.* **1992**, *25*, 495–503.

(2) Meisburger, S. P.; Sutton, J. L.; Chen, H.; Pabit, S. A.; Kirmizialtin, S.; Elber, R.; Pollack, L. *Biopolymers* **2013**, *99* (12), 1032–1045.

(3) Plumridge, A.; Meisburger, S. P.; Andresen, K.; Pollack, L. *Nucleic Acids Res.* **2017**, *45* (7), 3932–3943.

(4) Bai, Y.; Greenfeld, M.; Travers, K. J.; Chu, V. B.; Lipfert, J.; Doniach, S.; Herschlag, D. *J. Am. Chem. Soc.* **2007**, *129* (48), 14981–14988.

(5) Plumridge, A.; Meisburger, S.; Pollack, L. *Nucleic Acids Res.* **2017**, *45* (9), e66.

(6) Richardson, J. S.; Schneider, B.; Murray, L. W.; Kapral, G. J.; Immormino, R. M.; Headd, J. J.; Richardson, D. C.; Ham, D.; Hershkovits, E.; Williams, L. D.; Keating, K. S.; Pyle, A. M.; Micallef, D.; Westbrook, J.; Berman, H. M. *RNA* **2008**, *14* (3), 465–481.

(7) Schneider, B.; Morávek, Z.; Berman, H. M. *Nucleic Acids Res.* **2004**, *32* (5), 1666–1677.

(8) Lee, C. H.; Ezra, F. S.; Kondo, N. S.; Sarma, R. H.; Danyluk, S. S. Biochemistry 1976, 15 (16), 3627–3639.

(9) Svergun, D.; Barberato, C.; Koch, M. H. J. Appl. Crystallogr. 1995, 28 (6), 768–773.

(10) Bernado, P.; Mylonas, E.; V. Petoukhov, M.; Blackledge, M.; I. Svergun, D. J. Am. Chem. Soc. 2007, 129 (17), 5656–5664.

(11) Ullner, M.; Woodward, C. E. *Macromolecules* **2002**, *35* (4), 1437–1445.