

Supporting Information

Defining the structure of a protein-spherical nucleic acid conjugate and its counterionic cloud

Kurinji Krishnamoorthy¹, Kyle Hoffman², Sumit Kewalramani², Jeffrey D. Brodin³, Liane M. Moreau², Chad A. Mirkin^{2,3}, Monica Olvera de la Cruz^{1,2,3,4}, Michael J. Bedzyk^{1,2,4}.

¹Applied Physics Program, ²Department of Materials Science and Engineering, ³Department of Chemistry and ⁴Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208

S1. ASAXS Model calculations

Here we demonstrate the feasibility of ASAXS for our Pro-SNA Rb⁺ case. Determination of the structure of the counterion cloud relies on the measurement of subtle changes in the scattered intensity from a DNA nanoparticle-counterion system which arise due to sharp modulations in the effective scattering strength $f(q, E)$ of the counterion in the vicinity of a core-electron binding energy:

$$f(q, E) = f_0(q) + f'(E) + if''(E) \quad (\text{S1})$$

Here $f_0(q)$ is the energy independent form factor for the ion. In the small angle limit, $f_0(q) \sim Z$, the number of electrons in the ion. $f'(E)$ and $f''(E)$ are the energy dependent real and imaginary parts of the dispersion correction. For the case of Rb⁺, the expected variation of $f'(E)$ and $f''(E)$ with energy in the vicinity of a K absorption edge¹ (15.200 keV) is shown in Fig. S1 (A). In order to avoid strong fluorescence above an X-ray absorption edge, ASAXS measurements are typically performed at photon energies below the absorption edge of the targeted ion.²⁻³ At energies below the Rb K absorption edge the imaginary part of anomalous dispersion correction $f''(E)$ is practically zero (Fig. S1 (A)). Thus, the ASAXS effect is primarily dictated by $f'(E)$. We have performed model calculations 5 eV below the edge where $f'(E) = -7.58$.¹ Following Dingenouts et. al.,² the energy dependent scattered intensity from a Pro-SNA-Rb⁺ system above scattering from the salt solution can be written as:

$$I(q, E) - I_{RbCl}(q, E) = \frac{N}{V} [|F_0(q)|^2 + f'(E) (2F_0(q)v(q)) + (f'(E)v(q))^2] \quad (\text{S2})$$

Here $I(q, E)$ and $I_{RbCl}(q, E)$ represent the scattered intensities from the 50 mM RbCl solution with and without the Pro-SNA conjugate respectively. $\frac{N}{V}$ is the concentration of the Pro-SNA. The first term of the above equation is independent of energy and is equal to the scattered intensity from

the Pro-SNA-Rb⁺ system at energies far below the Rb K edge. This term is modeled using a core-shell form factor⁴ as follows:

$$F_0(q) = 3V_p(\rho_p - \rho_D) \frac{[\sin q R_p - q R_p \cos q R_p]}{(q R_p)^3} + 3V_t(\rho_D - \rho_s) \frac{[\sin q R_t - q R_t \cos q R_t]}{(q R_t)^3} \quad (\text{S3})$$

The bare protein is modelled as a sphere with a radius $R_p = 4.5$ nm and a uniform electron density $\rho_p = 403$ e⁻/nm³. The protein electron density (ρ_p) is extracted using CRY SOL⁵ and the corresponding protein data bank entry for the Catalase enzyme⁶ (PDB 4B7F). The DNA shell is modelled as having a thickness $t = 9$ nm which is inclusive of both the ssDNA strand (D) and the linker (L) segment covalently attaching the DNA to the protein core. The model does not take into account the electron density contrast between the linker and DNA segments and instead assigns an effective electron density $\rho_D = 346$ e⁻/nm³ to the DNA shell. The total radius of the conjugate is $R_t = R_p + t = 13.5$ nm. The electron density of the surrounding 50 mM RbCl solution is $\rho_s = 333.28$ e⁻/nm³. V_p and V_t are the volumes of the protein core and the DNA functionalized protein respectively. The second energy dependent term of Eq. S2 additionally contains $v(q)$, the Fourier transform of the excess Rb⁺ density. Assuming that the Pro-SNA-Rb⁺ system is spherically symmetric, $v(q)$ can be written⁷ as:

$$v(q) = 4\pi \int_0^\infty [n_{Rb}(r) - n_B] \frac{\sin qr}{qr} r^2 dr \quad (\text{S4})$$

Here $n_{Rb}(r)$ is the number density of Rb⁺ as a function of radial distance from the center of the conjugate and n_B is the number density in the bulk solution far away from the conjugate. For 50 mM RbCl, $n_B = 0.03$ ions/nm³. In order to compute $v(q)$, a simplified geometric model is used for the excess Rb⁺ density as follows:

$$n_{Rb}(r) - n_B = \begin{cases} -n_B, & r < R_p \\ n_l - n_B, & R_p < r < R_p + L \\ \left(\frac{N_{DNA} \times N_b}{4\pi t r^2}\right) - n_B, & R_p + L < r < R_t \\ 0, & r > R_t \end{cases}$$

(S5)

The Rb⁺ density is set to zero ($n_{Rb}(r) = 0$) in the region occupied by the protein core. Since each linker group (with length L) has a charge of $-3e$, the linker region is assigned a uniform Rb⁺ charge density ($n_l = 0.08661/\text{nm}^3$) equal to the Rb⁺ density required to neutralize the linker's negative charge. The excess Rb⁺ density in the linker region is thus given by $n_l - n_B$. Within the DNA shell it is assumed that the excess Rb⁺ density follows the charge density of the DNA. Namely, each unit of negative charge on the DNA is compensated by a corresponding Rb⁺ which occupies the same radial position as the negative charge it neutralizes. Since the Pro-SNA-Rb⁺ system is assumed to be spherically symmetric, the excess Rb⁺ density is modeled to fall off as the inverse square of r within the DNA shell. This is reflected in Eq. S5 where N_{DNA} refers to the number of DNA strands per protein and N_b refers to the number of bases per DNA strand. The number of DNA strands per protein, N_{DNA} is set to 40 – this loading density was determined using UV-Vis absorption spectroscopy using the known molar extinction coefficients for the protein ($\epsilon_{405} = 324,000 \text{ M}^{-1} \text{ cm}^{-1}$) and DNA ($\epsilon_{260} = 188,300 \text{ M}^{-1} \text{ cm}^{-1}$). The number of DNA bases N_b in our chosen DNA sequence is 18. Finally, the excess Rb⁺ density is set to zero beyond the DNA shell. The calculated partial intensities corresponding to the first, second and third terms of Eq. S2 are shown in Fig. S1 (B). At $q = 0$, the ratio of the second to the first term of Eq. S2, $2|F_0(q)f'(E)v(q)|/[F_0(q)]^2 = 0.073$ which implies that in going from an X-ray energy far below the Rb K-edge to an energy 5 eV below the edge, the change in the second term of Eq. S2 is ~ 7.3% which should be measurable above typical experimental uncertainties. The third term, however, is 3 orders of magnitude lower than the non-resonant term $[F_0(q)]^2$. In particular at $q =$

0, $[f'(E)v(q)]^2 / [F_0(q)]^2 = 0.0013$. This implies that the third term in Eq. S2 changes by only 0.13 % upon going from an X-ray energy far below the K-edge to one 5 eV below the edge. Such changes are lower than the typical statistical uncertainties of the experiment and cannot be measured reliably. Thus, we will neglect the third term $[f'(E)v(q)]^2$ of Eq. S2 in the subsequent analysis. This approximation is further supported by a simple argument and analysis of experimental SAXS data as described below.

We argue that when the scattered intensity changes by 10% due to the x-ray energy dependent ASAXS effect, the quadratic term only contributes 0.2% to the change. Consider the equation for the scattered intensities from the Pro-SNA/Rb⁺ system just below and far below the Rb⁺ edge in Eq. S6 and Eq. S7, respectively.

$$I_{\text{on}}(q, E) = \frac{N}{V} [|F_0(q)|^2 + f'(E) (2F_0(q)v(q)) + (f'(E)v(q))^2] \quad (\text{S6})$$

$$I_{\text{off}}(q, E) = \frac{N}{V} [|F_0(q)|^2] \quad (\text{S7})$$

Therefore, the ratio of the intensities measured on and off the absorption edge can be obtained by dividing Eq.S6 and Eq. S7:

$$\left[\frac{f'(E)v(q)}{F_0(q)} \right]^2 + \frac{2f'(E)v(q)}{F_0(q)} + 1 = \frac{I_{\text{on}}(q,E)}{I_{\text{off}}(q,E)} \quad (\text{S8})$$

For small values of q ($q < 0.3 \text{ nm}^{-1}$), $v(q)$ is expected to be positive. Looking at the inset in Fig. 3A the ratio on the R.H.S of Eq. 3 is ~ 0.9 , which yields $\left[\frac{f'(E)v(q)}{F_0(q)} \right]^2 = 0.0026$ and $2|f'(E)F_0(q)v(q)| = 0.102$. Therefore, when the linear term changes by 10%, the quadratic term changes only by 0.2%. Furthermore, it can be easily shown that as the ratio on the R.H.S. of Eq. 3 moves toward unity, the relative contribution of the linear term to the quadratic term to the ASAXS effect increases.

SAXS profiles from the Pro-SNA/counterion system show that the difference between the scattered intensities at 5 eV and 398 eV below the Rb^+ K edge (Fig. 3A, inset) is less than 10% (maximum difference = 9.1%). Therefore, the above argument that the quadratic term can be ignored in the analysis is valid for our case.

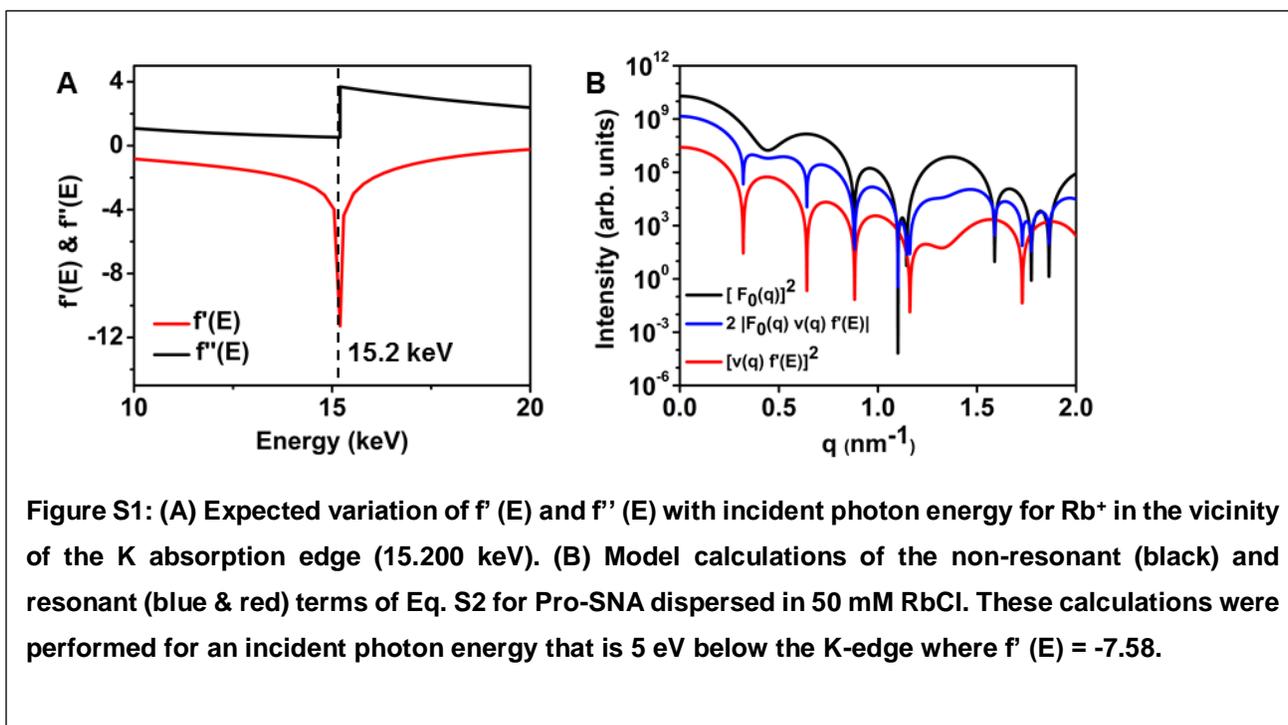


Figure S1: (A) Expected variation of $f'(E)$ and $f''(E)$ with incident photon energy for Rb^+ in the vicinity of the K absorption edge (15.200 keV). (B) Model calculations of the non-resonant (black) and resonant (blue & red) terms of Eq. S2 for Pro-SNA dispersed in 50 mM RbCl . These calculations were performed for an incident photon energy that is 5 eV below the K-edge where $f'(E) = -7.58$.

S2. Determination of Rb K absorption edge and $f'(E)$

In order to calibrate the position of the Rb K-absorption edge and experimentally determine the value of the anomalous dispersion correction $f'(E)$ as a function of incident photon energy in the vicinity of the edge, the transmitted intensity from a 1 M RbCl solution was measured over a 100 eV energy range around the expected Rb K edge (15.200 keV). A plot of the transmitted intensity as a function of incident photon energy is shown in Fig. S2 (A). The position of the Rb K edge was experimentally determined to be at 15.202 keV. The slight deviation in the measured value of the K-edge from its expected value is due to the slightly erroneous calibration

of the monochromator. The real [$f'(E)$] and imaginary [$f''(E)$] parts of the anomalous dispersion correction were determined using CHOOCH⁸, a program that produces $f'(E)$ and $f''(E)$ curves using a user supplied transmission or fluorescence energy scan. Briefly, the program evaluates $f''(E)$ using the optical theorem⁹:

$$f''(E) = \frac{mc\varepsilon_0 E \mu_a}{e^2 \hbar} \quad (\text{S9})$$

Here μ_a is the absorption coefficient of the targeted atom and the physical constants have their usual meaning. CHOOCH determines $f'(E)$ by numerically integrating the Kramers-Kronig transformation:

$$f'(E_0) = \frac{2}{\pi} \oint \frac{E f''(E)}{E_0^2 - E^2} dE \quad (\text{S10})$$

E_0 is the absorption edge of the targeted element which in our case is Rb with a K-edge $E_k = 15.200$ keV. Using these expressions CHOOCH determines the value of $f'(E)$ and $f''(E)$ directly from knowledge of the absorption coefficient as a function of energy. The determined $f'(E)$ and $f''(E)$ are shown in Fig. S2 (B) in red and black respectively.

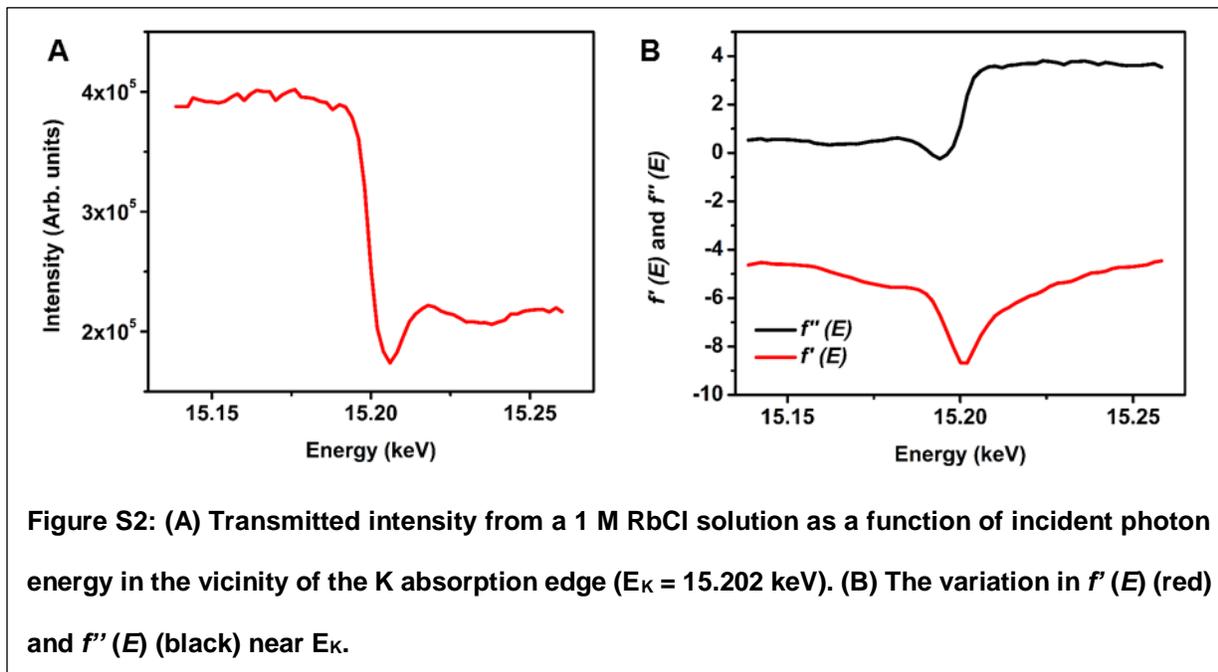
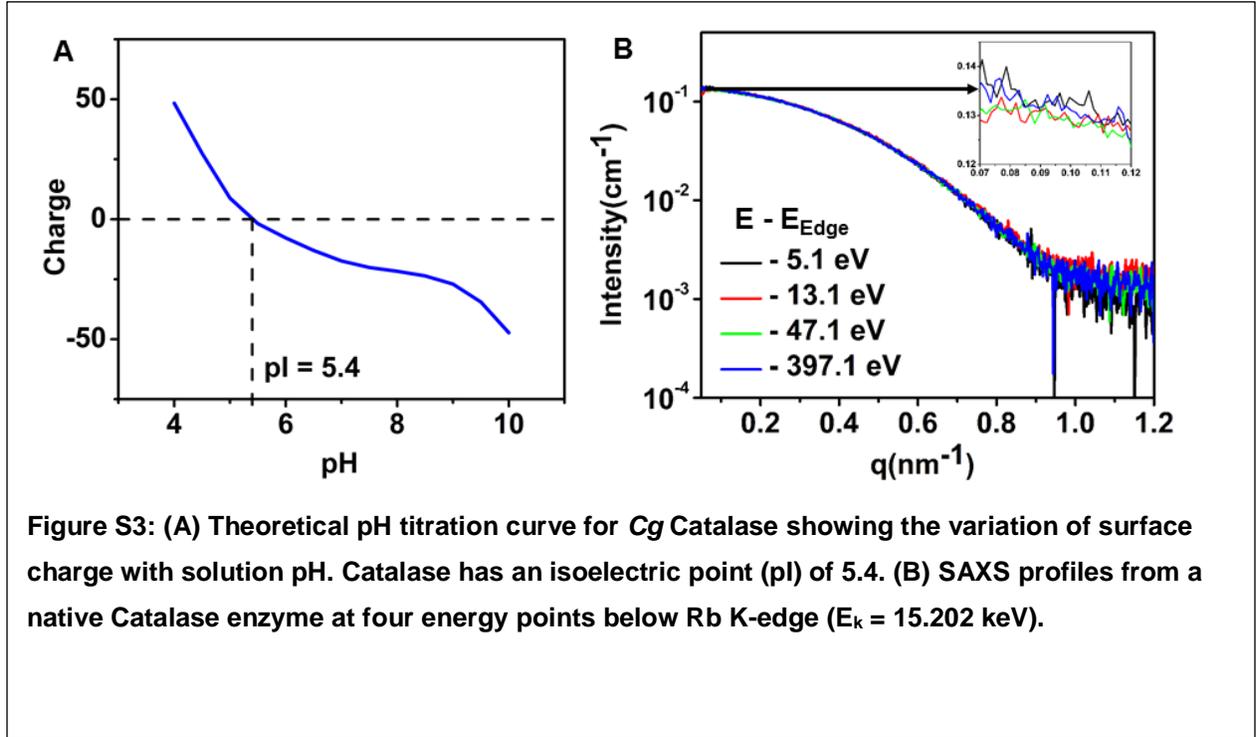


Figure S2: (A) Transmitted intensity from a 1 M RbCl solution as a function of incident photon energy in the vicinity of the K absorption edge ($E_K = 15.202$ keV). (B) The variation in $f'(E)$ (red) and $f''(E)$ (black) near E_K .

S3. ASAXS of bare protein and charge on protein

The isoelectric point (pI) of the native Catalase enzyme is 5.4¹⁰. A theoretically calculated¹¹ pH titration curve showing the surface charge on a Catalase enzyme as a function of solution pH is depicted in Fig. S3 (A). Fig.S3 (A) shows that the surface charge on a native protein in 50 mM RbCl solution is negligible in comparison to the charge of the DNA coating (-840). For instance, at pH 7 the charge on the protein is -17. Accordingly, the native enzyme is not expected to exhibit a significant ASAXS effect in the vicinity of the Rb K edge. This is validated by ASAXS measurements on bare proteins. Figure S3 (B) shows SAXS profiles of a bare protein at a 1 μ M concentration in 50 mM RbCl at four incident photon energies below the Rb K-edge. At $q = 0$, the difference in the scattered intensities for the largest incident energy change (between $E_K - E = 379.1$ eV and $E_K - E = 5.1$ eV) is $\sim 0.4\%$. In comparison, the change in the scattered intensities (at $q = 0$) for a Pro-SNA for the largest incident energy change is $\sim 10\%$. This demonstrates that

the protein core does not have an appreciable influence on the excess Rb^+ ion distribution surrounding the Pro-SNA.



S4. Extraction of resonant and non-resonant terms: Linear fit procedure

In order to extract the first and second terms of Eq.S2, a linear fit to the $I(q, f'(E))$ vs. $f'(E)$ data is carried out for each ' q ' in the range $0.05 - 1.2 \text{ nm}^{-1}$ resulting in the extraction of the resonant $[\frac{N}{V}(2F_0(q)v(q))]$ and non-resonant $(\frac{N}{V}[F_0(q)]^2)$ profiles. This procedure is summarized for the case of $1 \mu\text{M}$ Pro-SNA below. A similar procedure was employed for the case of $4 \mu\text{M}$ Pro-SNA.

SAXS profiles from a $1 \mu\text{M}$ Pro-SNA- Rb^+ system at 4 incident photon energies below the Rb K edge (15.2 keV) are shown in Fig. S4 (A). Intensity data on a linear scale at low q (Fig. S4 (B), inset) exhibits the expected trend, with the scattered intensity lowest for $E - E_{\text{Edge}} = -5.1$ and

increasing with increasing $E - E_{Edge}$ since $F_0(q)$ and $v(q)$ are both positive at low q . At $q = 0.3 \text{ nm}^{-1}$, $v(q)$ crosses the x-axis (Fig. S4 (E), red) as a result of which the measured scattered intensities at the 4 energies are equivalent, as indicated in Fig. S4 (B). Between $q = 0.3 \text{ nm}^{-1}$ and $q = 0.6 \text{ nm}^{-1}$, $v(q)$ is negative [Fig. S4 (E)] as a result of which the trend in the SAXS profiles is reversed. The cross-over point in the SAXS intensity profiles is depicted in Fig. S4 (B) (top inset). Figs. S4 (C) – (D) show examples of linear fits to $I(q, f'(E))$ vs. $f'(E)$ before and after the crossover point: Fig. S4(C) shows a fit at $q = 0.08 \text{ nm}^{-1}$ where the slope ($2F_0(q)v(q)$) and y-intercept ($[F_0(q)]^2$) are both positive. At $q = 0.33 \text{ nm}^{-1}$ (Fig. S4 (D)), the slope becomes negative – this represents the position of the first minima in the resonant term (Fig. S4 (F), red curve). Fig. S4 (F) shows the results of this extraction over an extended q -range ($0.05 - 1.2 \text{ nm}^{-1}$).

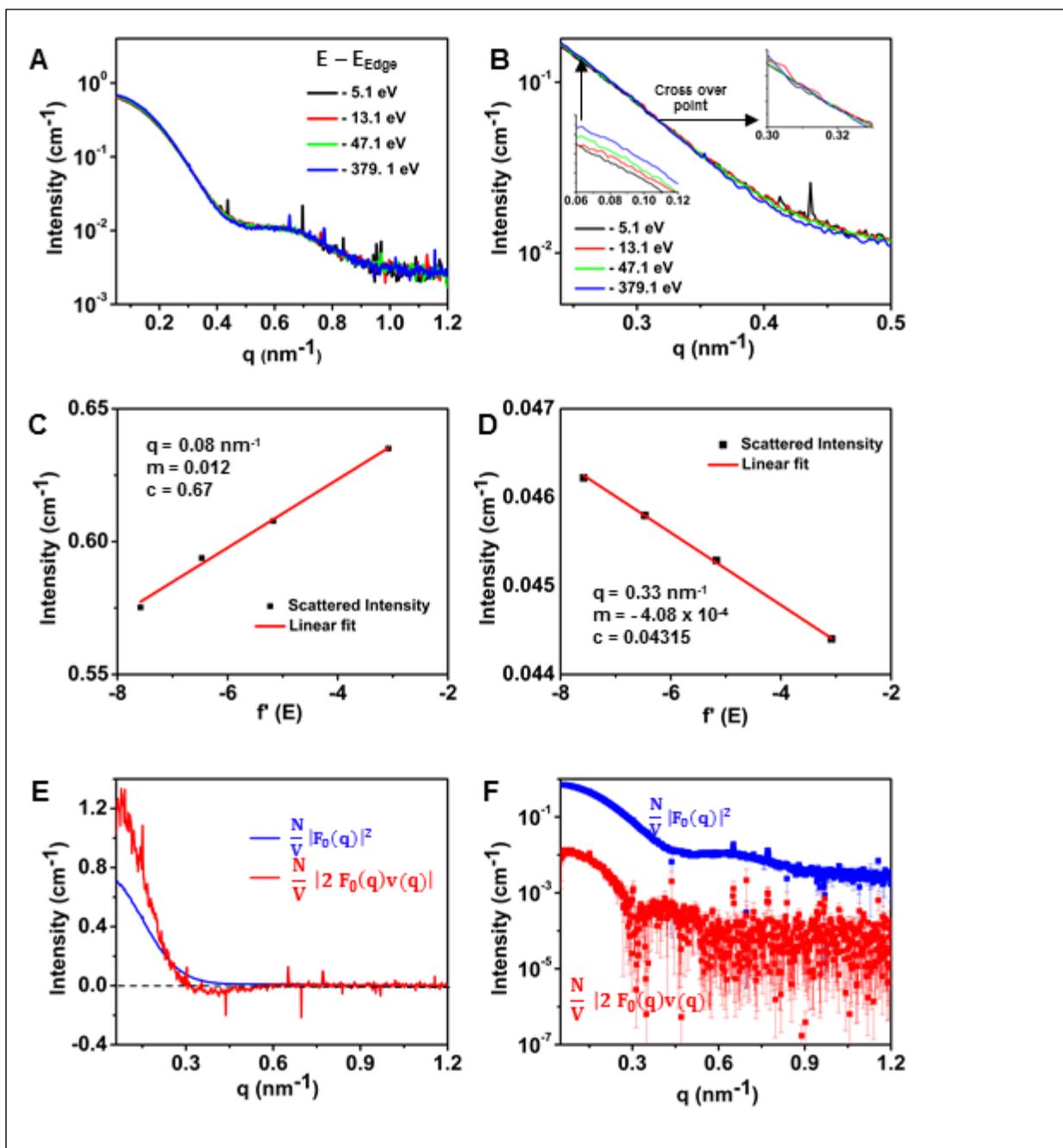


Figure S4: (A) SAXS profiles from 1 μM Pro-SNA-Rb⁺ at 4 incident photon energies below the Rb K edge. (B) SAXS intensity profiles between $q = 0.2 - 0.5 \text{ nm}^{-1}$ depicting the position of the crossover point at $q = 0.3 \text{ nm}^{-1}$. The insets show magnified intensities on a linear scale before and after the crossover point. Examples of linear fits to $I(q, f'(E))$ at $q = 0.08 \text{ nm}^{-1}$ (C) and 0.3 nm^{-1} (D). (E) Linear scaled $\frac{N}{V} |2F_0(q)v(q)|$ (red) and $\frac{N}{V} |F_0(q)|^2$ (blue) profiles depicting the zero crossings of $\frac{N}{V} |2F_0(q)v(q)|$ due to the effect of $v(q)$. (F) Log scaled $\frac{N}{V} |2F_0(q)v(q)|$ (red) and $\frac{N}{V} |F_0(q)|^2$ (blue) intensity profiles over an extended q -range.

S5. Bare Protein: SAXS Profile fitting

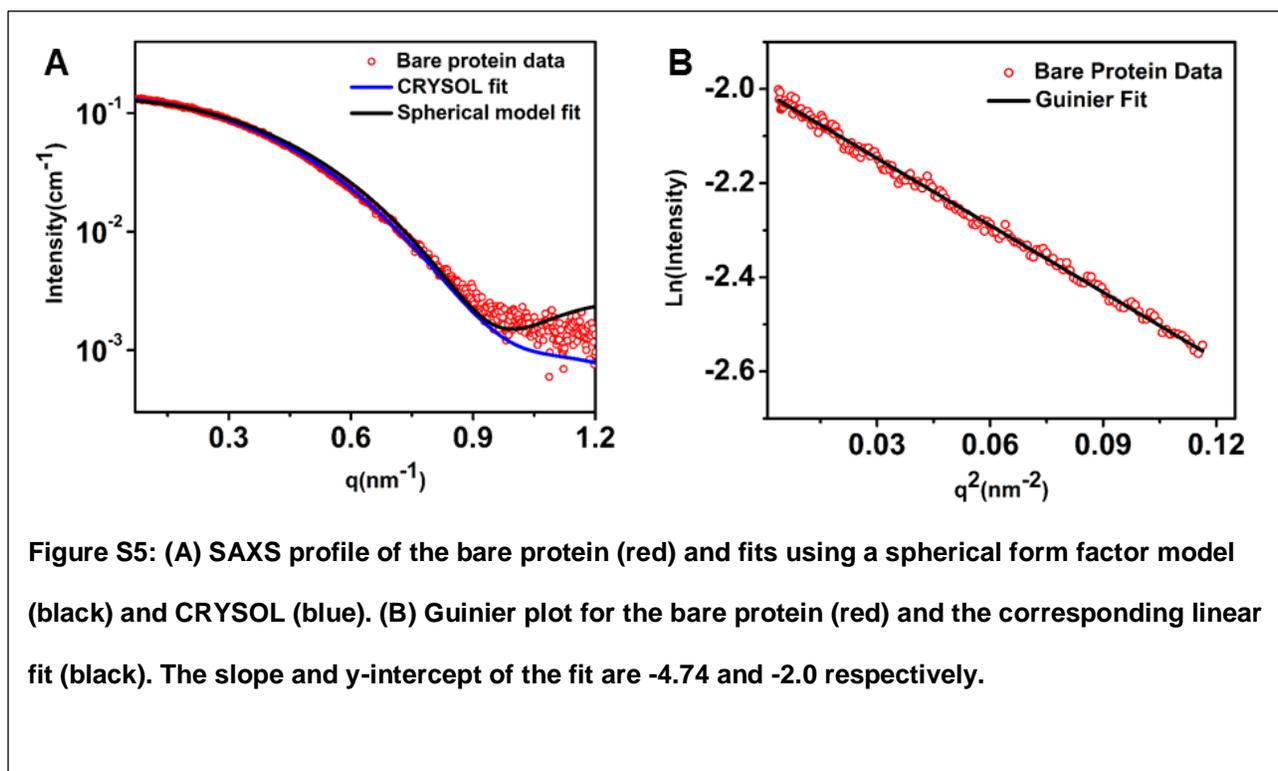
In order to extract the radius of the protein core (R_p) for subsequent fitting, the SAXS profile for an unmodified protein was fit with a model which employs the form factor of a homogenous sphere with a uniform electron density⁴:

$$F_p(q) = -3r_e V_p (\rho_p - \rho_s) \frac{[\sin qR_p - qR_p \cos qR_p]}{(qR_p)^3} \quad (\text{S11})$$

The protein core and surrounding RbCl solution are assigned uniform electron densities $\rho_p = 403 \text{ e}/\text{nm}^3$ and $\rho_s = 333.28 \text{ e}/\text{nm}^3$ respectively. r_e is the classical electron radius given by $r_e = 2.8179 \times 10^{-13} \text{ cm}$ used to convert the intensity to an absolute scale. The model is fit to the data using the protein concentration and radius R_p as parameters. The best fit of the spherical form factor model to the SAXS profile of a bare protein is depicted in Fig.S5 (A). The best-fit value obtained for the protein radius (R_p) is 4.5 nm. The SAXS profile evaluated by CRY SOL⁵ using the atomic coordinates from the protein data bank entry (PDB 4B7F) for Catalase⁶ is also depicted in Fig. S5 (A). Given the experimental range ($q < 1.2 \text{ nm}^{-1}$) where the scattered intensity is measurable above the background, the close shape agreement between the SAXS profiles generated using the spherical form factor model (Eq. S8) and CRY SOL demonstrates the validity of modeling the protein as a spherical core. Further, this corresponds to a radius of gyration $R_g = \sqrt{(3/5)} \times R_p = 3.5 \text{ nm}$ which is close to the value predicted by CRY SOL ($R_g = 3.7 \text{ nm}$). Additionally, Guinier analysis was performed for the bare protein SAXS profile to estimate R_g in a model independent manner. Briefly, in the Guinier region of the data ($qR_g < \sim 1$), the scattered intensity can be approximated by the Guinier equation⁴:

$$\ln I(q) = \ln(I_0) - \frac{q^2 R_g^2}{3} \quad (\text{S12})$$

A linear fit to the $\ln I(q)$ vs q^2 data described by Eq. S9 yields the radius of gyration R_g as the slope. Fig. S5 (B) depicts a Guinier plot for the unmodified protein along with the corresponding linear fit. The radius of gyration determined through this approach is 3.77 nm which is in close agreement with that predicted by CRY SOL.



S6. Core-Shell model for Non-resonant term and fitting

Following the linear fit procedure outlined in section S4, the first two intensity contributions to Eq. S2 were extracted: the non-resonant ($\frac{N}{V} |F_0(q)|^2$) and resonant ($\frac{N}{V} |2F_0(q)v(q)|$) intensity profiles [Figs. 3(C) and (D)]. The non-resonant intensity is equivalent to the scattering from the Pro-SNA, ions and water at X-ray energies away from the Rb K-edge and is modeled using a spherical core-shell model. The protein is modeled as a sphere having a fixed radius of $R_p = 4.5$ nm (section S5) and a uniform electron density of 403 e/nm³. The oligonucleotide shell is composed of linker segments and single stranded DNA, which are modeled as cylindrical rods with a radius of 0.5 nm.¹² The number of electrons in the linker and DNA segments is fixed based

on the chemical composition of the linker and DNA. The number of DNA + linker strands per protein was set to $N_{DNA} = 40$ based on the average loading density experimentally determined through UV-Vis absorption spectroscopy. The lengths of the linker (L) and DNA (D) segments and the concentration of the Pro-SNA conjugates (c) are used as fitting parameters. This concentration is related to the scale factor N/V appearing in both the non-resonant and cross term intensity profiles. The multiplicative scale factor N/V is given by:

$$\frac{N}{V} = c \times 6.022 \times 10^{14} \text{ cm}^{-3} \quad (\text{S13})$$

The non-resonant intensity profile ($\frac{N}{V}|F_0(q)|^2$) was thereafter fit to a spherical core shell model function⁴ given by:

$$I_{NR}(q) = \frac{N}{V} [F_0(q)]^2 + bkg \quad (\text{S14})$$

$$F_0(q) = F_1(q) + F_2(q) + F_3(q) \quad (\text{S15})$$

$$F_1(q) = -3r_e V_p (\rho_p - \rho_s) \frac{[\sin q R_p - q R_p \cos q R_p]}{(q R_p)^3}$$

$$F_2(q) = -r_e \pi r_o^2 N (\rho_L - \rho_s) \int_{R_p}^{R_p+L} \frac{\sin qr}{qr} dr$$

$$F_3(q) = -r_e \pi r_o^2 N (\rho_D - \rho_s) \int_{R_p+L}^{R_p+L+D} \frac{\sin qr}{qr} dr$$

Here, $\frac{N}{V}$ is the particle concentration dependent scale factor described by Eq. S10. V_p and V_t refer to the volume of the protein core and the DNA functionalized protein respectively, calculated using a radius of $R_p = 4.5$ nm for the protein and a radius of $R_t = R_p + L + D$ for the functionalized protein. The electron densities of the solvent, protein core, linker and DNA segments are given by ρ_s , ρ_p , ρ_L and ρ_D respectively. The experimentally extracted non-resonant intensity profile is

then fit to the core-shell model described by Eq. S11 using non-linear least squares curve fitting with constraints placed on the fitting parameters. The best fit of this model to the experimentally extracted non-resonant intensity profile is shown in Fig.4 for the two protein concentrations studied. The optimized values of c , D and L are listed in Table 1.

S7. DFT derived Rb^+ distribution profile to fit cross-term

Density functional theory was used to compute ion density profiles surrounding the ProSNA for a bulk RbCl concentration of 50 mM. The Fourier transform of the excess Rb^+ density profile yielded $v(q)$ which was then multiplied by $F_0(q)$ and the protein concentration-dependent scale factor N/V to obtain the resonant cross term profile ($\frac{N}{V} |2F_0(q)v(q)|$) which was then directly compared to its ASAXS extracted counterpart. The Rb^+ density was computed for various combinations of linker and ssDNA lengths. The linker length was varied between 2.5 nm and 6.5 nm in 0.5 nm increments. The ssDNA length was varied between 2.5 nm and 7.5 nm in 0.5 nm increments. Therefore, $9 \times 11 = 99$ combinations of linker and ssDNA lengths were computed.

Furthermore, the effect of three other parameters was explored: the excluded volume in the linker region, the excluded volume in the ssDNA region, and the DNA diameter. The linker excluded volume and the ssDNA excluded volume represent the volume in that region inaccessible to ions due to the steric repulsions. The ssDNA diameter affects the distribution of DNA charge and excluded volume. Increasing this parameter represents a transition from a more rigid structure where the ssDNA is immobilized in a cylinder to one where the average density is more spread out. The above 99 combinations of linker and ssDNA lengths were computed for each set of parameters listed below in Table . We found that results were not sensitive to these parameter choices for reasonable ranges. The subset of ion density calculations with a moderately low linker excluded volume of 1.3 nm^3 , ssDNA excluded volume of 4.0 nm^3 and ssDNA diameter of 1.2 nm were used for comparison with the experimentally extracted cross term profile.

Table S1. Sets of linker excluded volumes, ssDNA excluded volumes, and ssDNA diameters for which the above 99 combinations of linker and ssDNA lengths were calculated.

Linker excluded volume (nm ³)	ssDNA excluded volume (nm ³)	ssDNA diameter (nm)
0.0	0.0	1.2
1.3	4.0	1.2
1.3	4.0	2.4
1.3	4.0	3.6
1.3	4.0	4.8
5.0	4.0	1.2
9.0	4.0	1.2

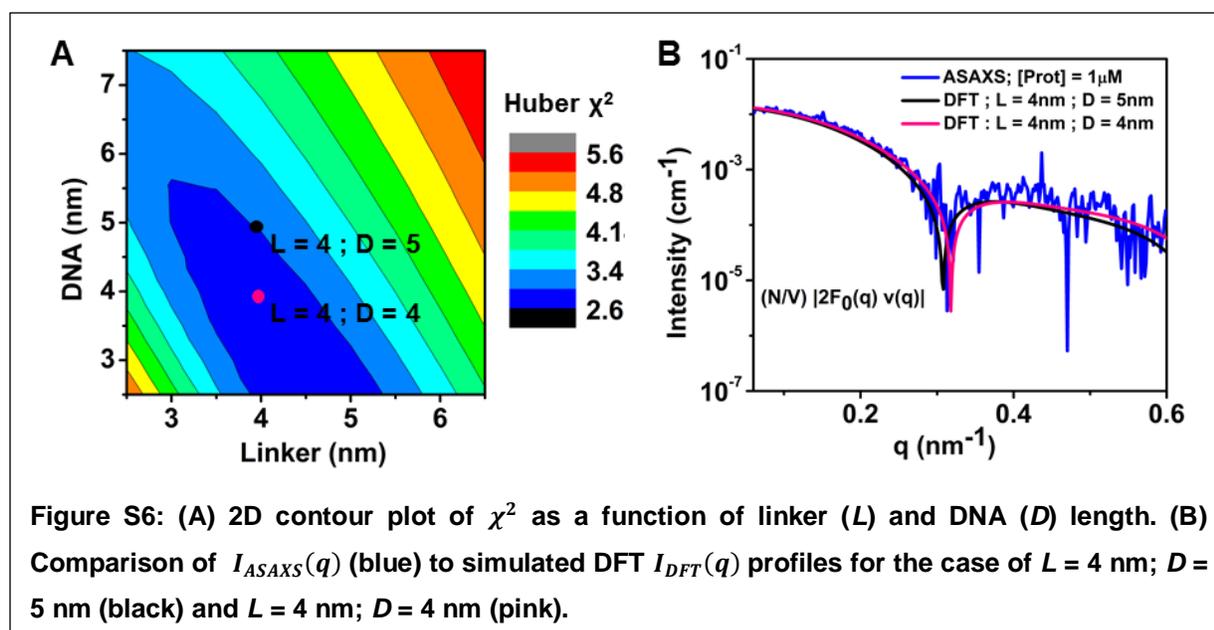
To quantify the agreement with the ASAXS experiments, the resonant intensities were calculated from the predicted cation densities. For each resonant intensity $I(q)$, the residual $r(q)$ was calculated by:

$$r(q) = \frac{I_{\text{experiments}}(q) - I_{\text{simulation}}(q)}{\sigma(q)}$$

(S16)

where $\sigma(q)$ is the standard deviation of the measured intensity at q . The individual residuals were then combined using a Huber loss function with tuning constant=1.345. The Huber loss function is more robust to outliers than the standard least squares measure. It reflects an assumption that points with high z-scores are indicative of data that is not normally distributed, and therefore assigns these points a higher probability than they would from a normal distribution. This prevents shifting the entire curve to fit a few outliers. An effective Huber χ^2 parameter was determined for

each set of parameters (Linker and DNA lengths) by dividing the Huber loss function by the number of data points in the experimental curve. A lower limit of 8 nm was placed on the total length of the DNA shell based on the value obtained from a uniform core-shell model fit to the non-resonant ($\frac{N}{V}|F_0(q)|^2$) profile. The DNA shell length determined through a uniform core shell model (Eq. S3) is effectively a lower limit for the possible length of the DNA shell because the model assigns a uniform electron density to the DNA shell and does account for the $1/r^2$ drop off in the DNA and linker electron density due to their arrangement on a roughly spherical protein core. A contour plot of Huber χ^2 as a function of DNA and linker length is shown in Fig.S6 (A). The lowest value of Huber χ^2 (2.8) was obtained for the case of $L = 4$ nm and $D = 4$ nm. Within a confidence defined by $\chi^2 + 1$, total DNA shell lengths ($L+D$) ranging from 8 – 10.5 nm provide an acceptable match to the experimentally extracted cross term profile. The low resolution of ASAXS for the length of the DNA shell is expected since the cross term profile can be reliably extracted only over a limited range in q (up to 0.6 nm^{-1}). The DFT calculated Rb density profile for the case of the best fit parameters ($L = 4$ nm; $D = 5$ nm) obtained from the core-shell model fit to the non-resonant profile results in a Huber $\chi^2 = 3.0$. Since this value of χ^2 lies within the $\chi^2 + 1$ confidence interval, the parameters $L = 4 (\pm 0.5)$ nm and $D = 5 (\pm 0.6)$ nm obtained from the non-resonant core-shell fit are deemed to reasonably describe the lengths of the linker and DNA segments. Huber χ^2 values corresponding to the DFT best match parameters ($L = 4$ nm and $D = 4$ nm) and the best fit parameters from the core-shell model fit ($L = 4$ nm and $D = 5$ nm) to the non-resonant profile are depicted by the pink and black markers in Fig. S6 (A). Fig. S6 (B) demonstrates that the experimentally extracted cross term profile (blue) is described equally well by both parameter sets.



References

1. Henke, B., *X-ray interactions : photoabsorption, scattering, transmission, and reflection at $E = 50$ - $30,000$ eV, $Z = 1$ - 92* . Academic Press: San Diego, 1993.
2. Dingenots, N.; Patel, M.; Rosenfeldt, S.; Pontoni, D.; Narayanan, T.; Ballauff, M., Counterion distribution around a spherical polyelectrolyte brush probed by anomalous small-angle X-ray scattering. *Macromolecules* **2004**, *37* (21), 8152-8159.
3. Das, R.; Mills, T.; Kwok, L.; Maskel, G.; Millett, I.; Doniach, S.; Finkelstein, K.; Herschlag, D.; Pollack, L., Counterion Distribution around DNA Probed by Solution X-Ray Scattering. *Phys. Rev. Lett.* **2003**, *90* (18), 188103-188103.
4. Guiner, A.; Fournet, G.; Walker, C., *Small angle scattering of X-rays*. J. Wiley & Sons, New York **1955**.
5. Svergun, D.; Barberato, C.; Koch, M. H. J., CRY SOL – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Cryst* **1995**, *28*, 768-773.
6. Candelaresi, M.; Gumiero, A.; Adamczyk, K.; Robb, K.; Bellota-Anton, C.; Sangal, V.; Munnoch, J.; Greetham, G. M.; Towrie, M.; Hoskisson, P. A.; Parker, A. W.; Tucker, N. P.; Walsh, M. A.; Hunt, N. T., A Structural and Dynamic Investigation of the Inhibition of Catalase by Nitric Oxide. *Org. Biomol. Chem.* **2013**, *11*, 7778-7778.
7. Kewalramani, S.; Zwanikken, J. W.; Macfarlane, R. J.; Leung, C.-Y.; Olvera de la Cruz, M.; Mirkin, C. A.; Bedzyk, M. J., Counterion Distribution Surrounding Spherical Nucleic Acid–Au

Nanoparticle Conjugates Probed by Small-Angle X-ray Scattering. *ACS Nano* **2013**, 7 (12), 11301-11309.

8.Evans, G.; Pettifer, R. F., CHOOCH: a program for deriving anomalous-scattering factors from X-ray fluorescence spectra. *J. Appl. Crystallogr.* **2001**, 34 (1), 82-86.

9.James, R. W., *The Optical Principles Of The Diffraction Of X Rays Vol II*. G. Bell And Sons Limited, 1969.

10.Samejima, T.; Yang, J. T., Reconstitution of acid-denatured catalase. *J. Biol. Chem.* **1963**, 238 (10), 3256-3261.

11.Kozlowski, L. P., IPC–Isoelectric Point Calculator. *Biology direct* **2016**, 11 (1), 55.

12.Egli, M.; Saenger, W., *Principles of Nucleic Acid Structure*. Springer: New York, 1984.