Biophysical Journal
Article

Biophysical Society

Local DNA Sequence Controls Asymmetry of DNA Unwrapping from Nucleosome Core Particles

Alexander W. Mauney,[1] Joshua M. Tokuda,[1] Lisa M. Gloss,[2] Oscar Gonzalez,[3] and Lois Pollack[1,*]
[1]School of Applied and Engineering Physics, Cornell University, Ithaca, New York; [2]School of Molecular Biosciences, Washington State University, Pullman, Washington; and [3]Department of Mathematics, University of Texas, Austin, Texas

ABSTRACT DNA is tightly wrapped around histone proteins in nucleosome core particles (NCPs) yet must become accessible for processing in the cell. This accessibility, a key component of transcription regulation, is influenced by the properties of both the histone proteins and the DNA itself. Small angle x-ray scattering with contrast variation is used to examine how sequence variations affect DNA unwrapping from NCPs at different salt concentrations. Salt destabilizes NCPs, populating multiple unwrapped states as many possible unwrapping pathways are explored by the complexes. We apply coarse-grained Monte Carlo methods to generate realistic sequence-dependent unwrapped structures for the nucleosomal DNA with thermal variations. An ensemble optimization method is employed to determine the composition of the overall ensemble as electrostatic interactions are weakened. Interesting DNA-sequence-dependent differences are revealed in the unwrapping paths and equilibrium constants. These differences are correlated with specific features within the nucleic acid sequences.

INTRODUCTION

In eukaryotic cells, DNA must be efficiently packaged for storage yet readily accessible for processes including transcription and repair. DNA is hierarchically packaged first into fundamental units known as nucleosome core particles (NCPs), which consist of ~145 basepairs of DNA tightly wrapped around a histone protein core (1). The canonical histone core consists of two copies each of the histone proteins H2A, H2B, H3, and H4, which form a pseudosymmetric helical ramp that organizes DNA (2). In the fully wrapped structure, the negatively charged DNA (with a persistence length of ~500 Å) is tightly bent by positively charged residues of the histone core into ~1.7 superhelical turns with a diameter of ~100 Å (1). NCPs are dynamic entities that populate diverse conformations to regulate DNA accessibility (3–5). A full understanding of how DNA is processed requires knowledge of these conformations and the interplay of factors that coordinate their formation. In the cell, reorganization of NCP structure is driven mostly by the activity of extrinsic proteins (e.g., polymerases, histone chaperones, and chromatin remodelers). However, the inherent biochemical and mechanical properties of nucleosomes themselves are essential to this activity. In particular, there is recent interest in the formation of asymmetric intermediates during transcription and remodeling processes that may be relevant in vivo (6).

Much past work has been devoted to the effect of histone variants or post-translational modification on NCP stability (5,7). The DNA sequence also contributes to stability (8,9), is critical in histone positioning (9,10) and sliding (11), contains hot spots for protein binding (11–13), and provides sites for epigenetic marks (14). In addition, DNA sequences encode mechanical features to facilitate its packaging and to control nucleotide access. For example, the CG content of a sequence correlates with an increase in short-range (~3 bases) bending and a decrease in longer-range bending (~10 bases) (15), and poly-A stretches contribute to especially rigid conformations (16). Finally, dinucleotide pyrimidine-purine steps (YR, such as CA, CG, TA, and TG) are known to be the most conformationally flexible (17).

Despite the demonstrated importance of DNA sequence, most investigations of nucleosome unwrapping using optical tweezers (18), atomic force microscopy (19), Forster resonance energy transfer (FRET) (12,18,20), or small angle x-ray scattering (SAXS) (20–22) treat the DNA as a uniform polymer. Notably, recent studies have highlighted the role of DNA flexibility in directing how NCPs unwrap (18). However, geometric constraints imposed by the optical tweezers requires that the free DNA on either side of the nucleosome remain co-linear, which restricts the orientations that can be

Submitted February 12, 2018, and accepted for publication July 5, 2018.
*Correspondence: lp26@cornell.edu
Editor: Tamar Schlick.
https://doi.org/10.1016/j.bpj.2018.07.009

sampled during unwrapping. In particular, the collinearity condition causes the bending-energy landscape of the DNA to have a kinetic barrier against unwrapping the final turn (23,24) and to favor asymmetric unwrapping (25). Nevertheless, these studies underscore the influence of DNA sequence in tuning the mechanical tendencies of nucleosomes. Studies on unconstrained nucleosomal dynamics have also been performed using restriction assays (26,27) and FRET (8,14,28); however, these techniques are limited to probing the behavior of a single location at a time and do not report on the global structural changes.

The combination of SAXS and ensemble modeling is a powerful method for characterizing the structures of free NCPs in solution. For polydisperse systems, the experimentally measured SAXS profile represents a linear combination of the scattering profiles of each of the conformations present in solution (29). To characterize the global structural parameters beyond the average radius of gyration or shape envelopes, advanced tools have been developed to determine distributions of conformations that describe the SAXS data. One such approach, the ensemble optimization method (EOM), allows selection of a subset of structures (the "ensemble") from a large pool of possible structures (30) whose computed scattering profile best recapitulates the SAXS data. The success and reliability of this approach depend on the quality of the conformational pool from which the ensemble is selected.

Recent applications of this approach enabled elucidation of DNA conformations within NCPs during salt-induced disassembly (21,22). Because the DNA-protein interactions are stabilized by electrostatics, increasing salt concentrations were used to weaken these interactions and trigger DNA unwrapping from the histone core in both static and dynamic studies (21,22). These studies illustrated a new method for identifying the multiple DNA conformations present in solution but did not account for the DNA-sequence-dependent effects that are relevant to regulation or NCP formation in vivo.

Here, we describe what is to our knowledge a novel method that explicitly accounts for DNA sequence in ensemble studies of NCPs. Coarse-grained simulations generate conformational pools that incorporate the known mechanical properties of DNA based on its sequence. Ensembles are then selected from these pools using experimental SAXS data as a guide. We apply this approach to compare the salt-induced unwrapping from NCPs of two DNA sequences: the SELEX (systematic evolution of ligands by exponential enrichment)-engineered Widom 601 (31) and the natural 5S ribosomal DNA (rDNA) from *Lytechinus variegatus* (32). We find that the inclusion of sequence information allows us to distinguish the ends of the nucleosome, which was previously impossible using SAXS. Analysis of selected structures reveals newly discovered links between the conformations populated and the underlying mechanical properties of the DNA to our knowledge. These insights may be useful in predicting how and where nucleosomal DNA becomes accessible in vivo.

MATERIALS AND METHODS

SAXS

SAXS provides low-resolution, global structural information for proteins and/or nucleic acids in solution. SAXS profiles report scattering intensity I as a function of the momentum transfer $q = 4\pi\sin(\theta)/\lambda$, where λ is the x-ray wavelength and 2θ is the scattering angle. For a multiple-component system, such as a protein-nucleic acid complex like the nucleosome, both the components contribute to the overall scattering. The contributions are summed to result in a scattering amplitude, A; however, intensity (which is $|A|^2$) is measured, containing cross-term scattering that reflects contributions from both components (33).

The presence of this cross-term complicates interpretation of scattering profiles. The contribution of each component (which reflects its conformation) cannot be readily extracted from the full scattering intensity without prior knowledge of the shape of one component. The use of contrast variation, which increases the electron density of the solvent to match the protein density, removes the protein contribution to the scattering profile, leaving only the contribution from the DNA. Fig. S2 shows the effect of sucrose on the measured scattering profiles of both proteins and DNA. To match the histone proteins, 50% w/w of electron dense sucrose was added to the buffers. Sucrose is ideal for contrast matching, as it negligibly affects electrostatic interactions and nucleosome stability (22,34). SAXS data acquired at the contrast-matched condition are shown in Fig. S3.

SAXS data collection

SAXS data were collected at the Cornell High Energy Synchrotron Source G1 station. Monochromatic x rays at either 9.97 or 10.5 keV were incident on the samples, with normalization of the intensity achieved using counts from a PIN diode beamstop at 9.97 keV, whereas the transmitted beam intensity measured after transmission through a semitransparent molybdenum beamstop (Goodfellow, Coraopolis, PA) was used at 10.5 keV. Scattered x rays were imaged onto a PILATUS 100K (DECTRIS, Baden-Dättwil, Switzerland) detector located 1.5 m (q = 0.009–0.289 Å^{-1}) or 2.0 m (q = 0.007–0.250 Å^{-1}) away from the sample. This calibration was established using a silver-behenate standard.

Nucleosomes incorporating 149-basepair DNA were produced using the same method as previous work (22) (detailed in the Supporting Materials and Methods), and samples were manually prepared and equilibrated for ~10 min before being loaded into a flow cell at room temperature and oscillated during exposure to reduce radiation damage. Scattering profiles were integrated and processed in MATLAB (The MathWorks, Natick, MA) according to established protocols (35). Background subtraction was achieved by subtracting the profile measured for the matching buffer from that of the sample.

Ensemble optimization method

The EOM creates a number of random ensembles of a given size (up to 50 models) from a much larger pool. The scattering profiles of these ensembles are then compared to the experimental profile, and the quality of the fit is assessed by a fitness parameter (χ^2). Ensembles with good fitness parameters progress into the next generation, in which new ensembles are created by randomly changing elements of previous ensembles or by swapping sections from one ensemble with another. These new ensembles are then evaluated as discussed above, with the best advancing to the next generation. After 1000 generations, the best-fitting ensemble is chosen as the final

ensemble. This process is repeated many (~700) times using different starting ensembles. The resulting structures are mined to extract the unwrapping parameters described in the text.

The pool of structures is generated using the cgDNA suite (36). This program provides a direct, explicit prediction of the sequence-dependent free energy and associated equilibrium distribution of a B-form DNA oligomer of arbitrary sequence in solvent under prescribed environmental conditions. Within this model, each base on each strand of a DNA oligomer is considered to be a rigid entity interacting with its nearest neighbors. The coarse-grained configuration of a given oligomer is described by a set of relative coordinates denoting both the rotational and translational displacement between adjacent basepairs and within each basepair.

The free energy is modeled by a shifted quadratic function, which is defined by the equilibrium or ground-state configuration of the oligomer, together with an equilibrium stiffness matrix that describes energetic couplings between bases. The ground-state configuration depends nonlocally on sequence, whereas the stiffness matrix depends locally on sequence at the level of dinucleotides. Once the free energy for an oligomer is constructed, the configurational statistics of the oligomer are described by an associated Gaussian probability density on the space of internal coordinates, which provides a model for the thermal distribution of oligomer configurations in the solvent. Although the current version of cgDNA (36) employs a background of 150 mM of KCl salt, we assumed that the range of configurations of the oligomer ends would be similar across the concentrations of NaCl salt considered here and used cgDNA to generate candidate structures for all salt concentrations.

As described, the cgDNA model provides a Gaussian distribution on the space of internal coordinates for an unconstrained oligomer in solvent. When any subset of the internal coordinates is held fixed, as for example in a model of the wrapped region of a nucleosome, there is then an implied distribution on the remaining free coordinates. This implied (or conditional) distribution is also Gaussian and can be determined using well-known formulae (37). Beginning from the Protein Data Bank (PDB): 1AOI structure, we computed the internal coordinates for a fully wrapped nucleosome and then constructed and sampled the sequence-dependent conditional distribution on the free coordinates for different numbers of free bases at each end of the 601 and 5S oligomers. More details about the parameters used, as well as fitting results, are provided in the Supporting Materials and Methods and shown in Figs. S4–S22.

FRET

Two FRET pairs were used to monitor H2A-H2B dissociation from the 601- and 5S- containing NCPs: H3-78W/H2B-109Cys-AEDANS and H4-60W/H2A-108Cys-AEDANS. This system, including protocols for sample preparation, data collection, and spectral analysis, is described in previous publications (21,38). FRET data reported here were collected at a single NCP concentration of 25 nM.

RESULTS AND DISCUSSION

We applied contrast-variation SAXS (CVSAXS) to measure the salt-dependent conformations of DNAs within NCPs. In CVSAXS, the electron density of the surrounding solvent is raised to match (equal) that of the lower-density protein component of a protein-DNA complex (see Materials and Methods). Because the SAXS signal is proportional to electron-density differences of solutes relative to the solvent, only the higher-density DNA component contributes to the SAXS profile.

Data were acquired on two different NCP constructs. Wild-type histones packaged either the naturally occurring 5S (32) or the artificially engineered Widom 601 sequence (31). Static SAXS data were acquired in equilibrium salt titrations, with NaCl concentrations ranging from 0.2 to 1.8 M. At low salt, the DNA in NCPs is fully wrapped. Partial release occurs with increasing [NaCl], and full release is observed at the highest salt concentrations. The use of contrast variation allows us to focus strictly on the changing DNA conformations during this salt titration.

To analyze the data, we employed an EOM to ascertain which DNA conformations are present in the SAXS data for a particular construct at a given salt concentration. An overview of this method is illustrated in Fig. 1, with a brief description provided in the caption.

A critical step in the EOM process involves the generation of a pool of structures that reflect realistic conformations of the macromolecule. For protein systems, this process is relatively straightforward: multiple structures can be extracted from the PDB or can be generated using molecular dynamics (MD) simulations. For nucleic acids, the situation is more complicated; many fewer solved structures exist, MD simulations of nucleic acids are more challenging, and solvent interactions raise the computational cost greatly (39). Protein-nucleic acid complexes introduce more complexity; MD simulations are even more computationally difficult, if available (39). New methods are therefore required to construct complete structural pools for compounds that include nucleic acids. In past works, models of unwrapped DNA were generated by replacing curved,

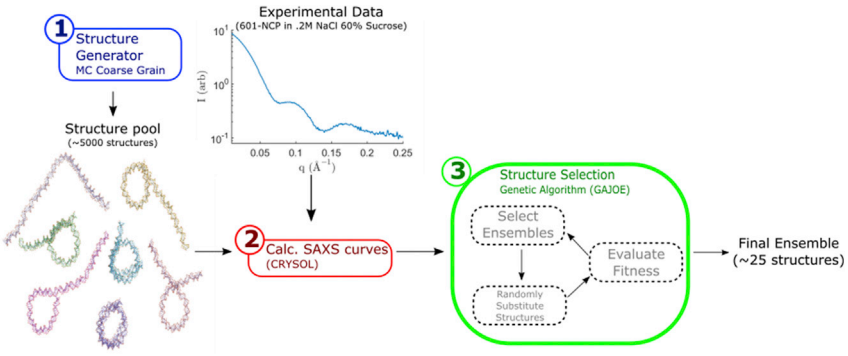

FIGURE 1 Schematic representation of ensemble optimization method (EOM). Briefly, pools of many possible structures are generated from Monte Carlo coarse-grained simulations (*step 1*). The theoretical scattering profile for each structure in the pool is computed using CRYSOL (*step 2*) for comparison with experimental data. A genetic algorithm determines the ensemble of models whose summed scattering profiles best match the data. The genetic algorithm is repeated multiple times (*step 3*), and the final ensembles are combined to provide an overall distribution of structures that best recapitulates the measurement. To see this figure in color, go online.

bound nucleosomal DNA with linear B-form structures: the unwrapped DNA was modeled as a stiff extension directed along the trajectory of the nucleosomal DNA. This approach did not leverage the mechanical information present in the DNA sequence.

Here, we describe a method for generating improved structure pools that vastly expand the knowledge gained from EOM by introducing thermally allowed motions that reflect sequence dependent flexibility. Beginning with the PDB: 1AOI crystal structure (1) (146 nucleotides) (Fig. 2 *a*), the location and orientation of each base is found as a first step in coarse graining the structure. Once this spatial information is extracted and logged, the nucleic acid sequence is converted to either the 601 or 5S DNA sequence, and an additional basepair is appended to each end to match the length of the DNA used in SAXS experiments (149 nucleotides). A unique stiffness matrix is calculated based on the sequence, using the mechanical properties of each basepair and basepair-basepair junction. This matrix is employed to identify the lowest-energy configuration of the appended bases. Once identified, this configuration comprises the fully wrapped nucleosome, the starting structure for model generation (Fig. 2 *b*).

Our full structural pool is generated from this starting structure by varying the number of free basepairs on each side of the nucleosome to mimic all possible degrees of DNA unwrapping from the histone core. We designate these sides as "left" and "right" based on a 5′–3′ reading of the sequence of interest, and the number of free basepairs on each side is designated by n_L and n_R, respectively. Once a basepair is designated "free," a new lowest-energy conformation is determined based on the stiffness matrix, which is also used to find an envelope of realistic thermal displacements. Representative structures, sampled within this envelope, are added to the pool and classified based on the total number of basepairs unwound, $n_T = n_L + n_R$. Finally, atomistic structures of the individual bases are properly placed into the coarse-grained models to generate the final structure of nucleosomal DNA used by EOM. This process is iterated by releasing five basepairs at a time to access values of n_L and n_R that span the space of all possible unwrapped structures. The final pool for each sequence contains 5104 model structures.

With the pool fully in place, EOM is applied to select an ensemble of DNA conformations whose summed SAXS profiles recapitulate the experimental data. This process was repeated for SAXS profiles acquired at each experimental concentration. Because of the inherent degeneracy in the model pool, the fitting procedure was run 700 times to produce a statistical distribution of structures. The fit quality measured using χ^2 was consistent across these multiple runs.

We classify selected structures by three parameters: the total number of basepairs unwrapped (n_T), basepairs unwound from the left end (n_L), and basepairs unwound from the right end (n_R). Normally, SAXS is unable to discern which end of the nucleosomal DNA is released; however, the mechanical differences between the ends (a result of differing sequences) lead to distinct envelopes mapped out by the thermal variations and allow the ends to be differentiated with this EOM approach. Comparison of the n_L and n_R values of the chosen models reveals the asymmetry of unwrapping, a parameter of great interest in determining any sequence-dependent bias.

To establish both the degree of unwrapping as well as any left/right biasing, we group the selected models together in histograms (Fig. 3). The *x* axis of the histogram reports the total number of basepairs unwrapped (n_T); the *y* axis shows the fraction of the selected structures containing n_T unwrapped basepairs. The histograms of Fig. 3 represent selected structures that recapitulate SAXS data acquired on NCPs with the 601 DNA sequence at 1.0 M NaCl. At this intermediate salt value, a wide variety of structures are present in the selected ensemble. The following information is readily gleaned from the distribution shown in Fig. 3 (*top*): none of the selected structures have fewer than 35 unwound

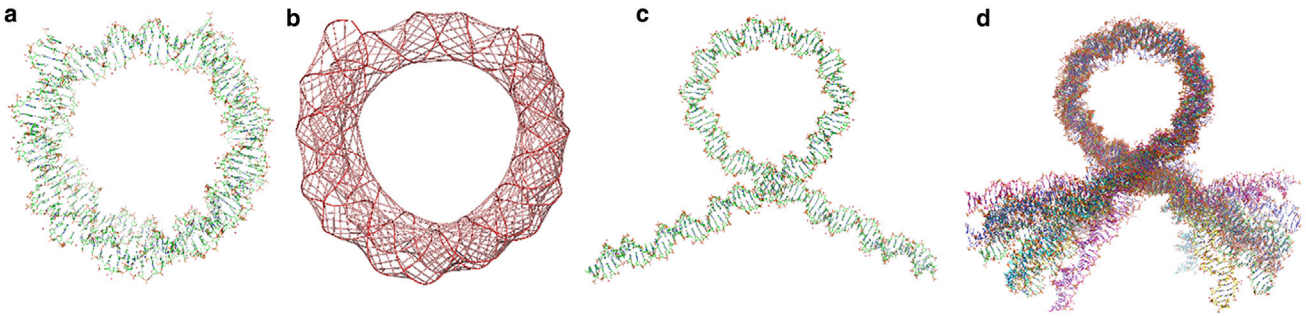

FIGURE 2 Representations of steps taken to generate nucleosome models. (*a*) shows the crystal structure of DNA within a wrapped nucleosome. (*b*) shows the locations of bases and the backbone extracted from (*a*) along with the added bases at the end that extend the length of the DNA to match the samples used in SAXS experiments. (*c*) illustrates the lowest-energy state when 40 basepairs are extended on each end of the DNA (40 free basepairs on each end), and (*d*) shows 10 structures that represent thermal variations of the state shown in (*c*). Conformational differences that reflect the DNA sequence are visible at the two ends. To see this figure in color, go online.

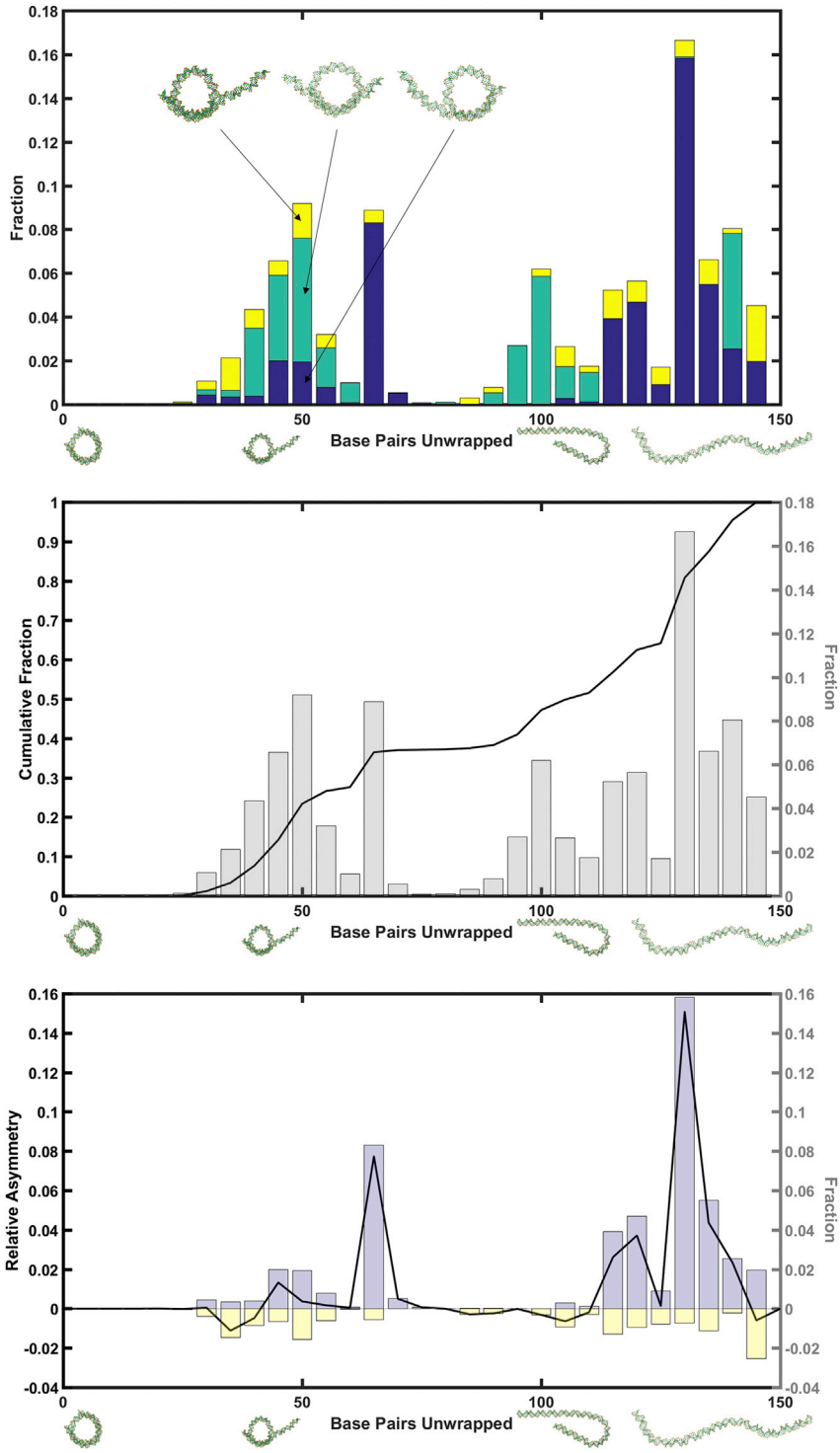

FIGURE 3 Illustration of analysis protocol. The top panel shows the unwrapping histogram of the DNA structures identified when NCPs containing 601 DNA are placed in 1.0 M NaCl. Multiple structures have the same number of basepairs unwrapped, parameterized by n_T; however, these structures can be distinguished by identifying the dominant side for release (*blue* (darkest): left dominant, *green*: no dominant side, *yellow* (lightest): right dominant), with example models of each type shown for $n_T = 50$. Example structures have also been placed along the axis to illustrate other unwrapping amounts. The middle panel demonstrates the relationship between the cumulative unwrapping histogram and the unwrapping histogram shown above. The former is the integral of the latter and is sensitive only to the total number of bases unwrapped, not to the side of release. The bottom panel shows the relative asymmetry distribution of this set of structures: the difference between right- versus left-dominated structures at each value of n_T. Only asymmetrically unwrapped structures are included in this representation. To highlight the asymmetry, we establish the following sign convention: the right-dominant fraction of structures is treated as negative, whereas the left-dominant fraction is treated as positive. Thus, a net positive value represents a tendency for left-dominant fractions in the models. To see this figure in color, go online.

basepairs, one cluster of structures contains between 35 and 70 unwound basepairs, and a second broad group of structures contains more than 85. To further mine the selected ensemble, we examined the structures at each value of n_T and classified them in three ways: the fraction that are symmetrically unwound ($n_L = n_R$, plotted in *green*), the fraction that favor unwinding from the left end ($n_L > n_R$, plotted in *blue*), and those that favor unwinding from the right end ($n_R > n_L$, plotted in *yellow*). Because a step size of five was used for the unwrapping, differences smaller than 10 are the result of a single step and thus may be spurious. Thus, if one side has 10 (or more) basepairs unwound

than the other side, the unwinding is considered asymmetric. This approach highlights asymmetry and indicates which end is favored for unwrapping, if any.

To facilitate comparisons across conditions (different constructs or different salts), we integrate the unwrapping histogram (Fig. 3, *top*) to derive the cumulative unwrapping histogram (Fig. 3, *middle*), which readily displays patterns of unwrapping. In this representation, the slope contains information about the distribution of chosen structures; a flat portion of the curve represents a region of n_T that contains few additional unwrapped structures, whereas a sloped segment indicates a region of n_T that contains a larger population. For sets of models that are mostly wrapped (the majority of structures populating the low-n_T side of the histogram), the cumulative fraction approaches 1 at low n_T. In contrast, for sets that contain mostly unwrapped models (the majority of structures populating the large-n_T side of the histogram) the cumulative fraction approaches 1 at higher n_T. This approach focuses only on the total number of bases unwrapped; it contains no information about the asymmetry of the unwrapping. To evaluate preferential unwrapping, we create the relative asymmetry distribution, (Fig. 3, *bottom*) by subtracting the fraction of right-dominant structures from left-dominant structures at each total unwrapping number. A peak in this difference plot indicates a region in the unwrapping distribution where one side is more likely to unwrap than the other. With these analysis tools in place, we now turn to the full salt titrations for both constructs.

Widom 601

Fig. 4 illustrates the salt-dependent unwrapping of the Widom 601 DNA sequence. Static SAXS profiles, acquired at the different salt values indicated in the figure, were analyzed by EOM to select structural models from the full pool as described above. Analysis of these structures, displayed in Fig. 4, reveals distinct and salt-dependent unwrapping patterns. The cumulative unwrapping distribution (*left*) shows the progressive increase in unwrapping as the salt is increased. Many structures populate the low-n_T side of the histogram at low salt, whereas the opposite is true at high salt. The relative asymmetry distribution (*right*) shows a highly asymmetric unwrapping at low salt that appears to alternate between sides but with no consistent trend or dependence on salt. Both sides appear to unwrap simultaneously, and the varying heights of the peaks in this curve indicate that the two sides unwrap by different amounts at a given salt concentration. However, when ~65 bases have unwrapped, a consistent trend emerges and persists through the salt concentrations: a left-dominant state is much more likely than a right-dominant one. At the highest salt concentrations, at which the DNA is nearly fully unwrapped, there is little difference between left-biased and right-biased states. These high salt structures represent essentially free DNA with slight curvature from the fixed bases, which leads to the noise at high n_T. This strong asymmetry at low salt is consistent with previous indications of asymmetry detected in 601 DNA unwrapping during pulling experiments (18,40), when the sequences are aligned.

Taken together, the information in these curves provides insight into how this DNA is released from the histone core. The cumulative unwrapping distribution curve (*left*) reports that mostly wrapped structures are released at the ends (total numbers of basepairs unwrapped less than ~40), consistent with breathing modes previously reported (41). For n_T values of ~40–60, the curve is relatively flat, suggesting that there are few structures in this region. However, for larger unwinding values, in the range in which the bias toward left-biased states is observed from the relative asymmetry distribution (n_T between 60 and 70), the slope increases, signaling that a short stretch or region of the DNA is released all together. We propose that this region acts like a spring-loaded latch; it can be held closed by the earlier bases, but once it is no longer constrained, it pops open (DNA is released from the nucleosome).

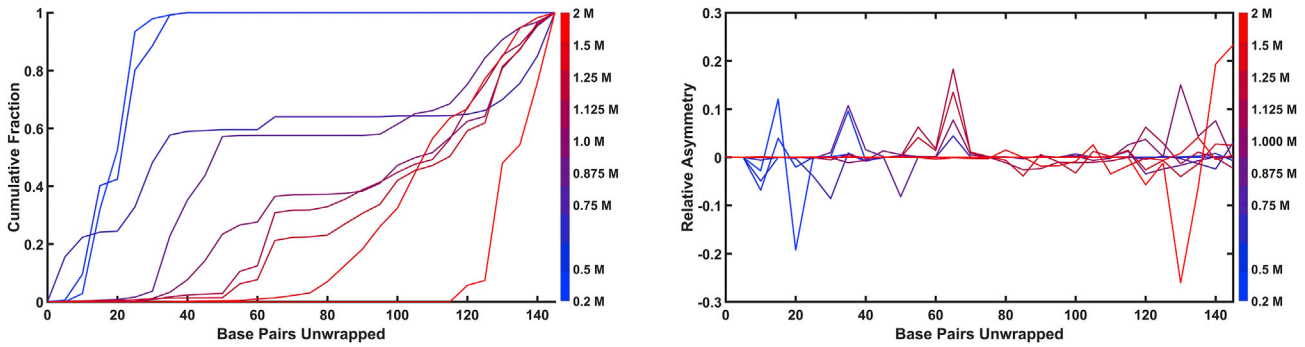

FIGURE 4 Cumulative unwrapping distribution (*left*) and relative asymmetry distribution (*right*) for Widom 601 DNA. Salt concentration is denoted by color: blue curves correspond to low salt, whereas red curves correspond to high salt. One curve (625 mM NaCl) was removed from the data set because of poor fitting, likely due to sucrose mismatch between buffer and sample.

5′ATCGGGATGTATATATCTGACACGTGCCTGGAGACTAGGGAGTAATCCCCTTGGCGGTTAAAACGCGGGGGACAG
3′TAGCCTCTTAGGGCCACGGCTCCGGCGAGTTAACCAGCATCTGTCGAGATCGTGGCGAATTTGCGTGCATGCGC

FIGURE 5 Widom 601 Sequence. The second half of the sequence is flipped right to left to align DNA ends. Flexible YR dinucleotides are highlighted, and a stiffer region of interest is underlined. The "left" end is the 5′ end in this reading, whereas the "right" end is the 3′ end.

A close examination of the DNA sequence reveals how the 601 sequence could contribute to this behavior. Most DNA is relatively stiff; however, pyrimidine-purine (YR) steps give rise to more flexibility (17). These steps are highlighted in Fig. 5. It is worth noting that these locations do not depend on which strand is being analyzed, as the complement of a YR step is also a YR step. Fig. 5 shows that there is one particular region that contains significantly fewer flexible motifs. This region (*underlined*) is positioned ∼30 bases from one end and is ∼20 bases long. The relative asymmetry distribution suggests that unwrapping proceeds symmetrically and relatively steadily up to a point when around 20 bases have been freed from each side. At this point, the spring-latch mechanism kicks in, and the next 20 basepairs on the left side are rapidly released. These numbers are consistent with the position of the peak in asymmetry at an n_T value around 60–70. The right side does not have a similar rigid region, and the DNA unwraps progressively until the structures are symmetric before the remaining DNA is released.

5S rDNA

The same analysis can be applied to quantify the salt-dependent unwrapping of the naturally occurring 5S DNA sequence (Fig. 6). For this construct, a large population of fully unwrapped (120+ basepair) structures is present at all salt concentrations, which we attribute to free DNA in the sample. The cumulative unwrapping distribution for this 5S DNA differs substantially from the one shown in Fig. 4 for the Widom 601 DNA. At the lowest salt concentrations (*blue*), DNA in the former NCPs is already more unwrapped than the latter. For all salt concentrations, the curve is flat for n_T between 50 and 90 basepairs; for many of the curves, this trend persists until n_T reaches 120. Thus, the nucleosomal DNA transitions rapidly from mostly wrapped to mostly unwrapped states. The relative asymmetry distribution for this 5S DNA also contrasts with the curve for 601 DNA. The absence of persistent peaks suggests that the 5S sequence has no side preference for initial unwrapping (small values of unwrapping number). Examining the 5S sequence, we note two large, stiff regions ∼35 bases in from each side, underlined in Fig. 7. These regions are both preceded by relatively flexible regions and end 55 and 65 bases from each end. The experimentally observed unwrapping behavior is consistent with the idea that the unwrapping proceeds steadily through the flexible regions on each side. Once the stiff regions are exposed, the bending energy overcomes the binding energy and the nucleosome "jumps" to a nearly fully unwrapped state.

Comparing Figs. 4 and 6, we see that the 5S sequence is more unwrapped at any given salt concentration than the 601 sequence. The models representing the 5S DNA also have many fewer partially unwrapped states compared to those representing the 601 sequence.

To gain additional insight into the differences reported above and validate the EOM predictions, FRET experiments were carried out to monitor NCP dissociation as a function of salt for the two different DNA constructs. Because the labels are on the proteins, not the DNA, changes in the energy transfer reflect a disruption of the histone core (38). Importantly, many of the conclusions of the EOM-based analysis of DNA dissociation are validated by FRET data. The high-FRET state occurs when the labels are in close proximity, corresponding to a fully formed histone core. Decreases in fluorescence correspond to disruption of the core, which is coupled to DNA unwrapping (21). The two constructs behave quite differently when assessed with this metric.

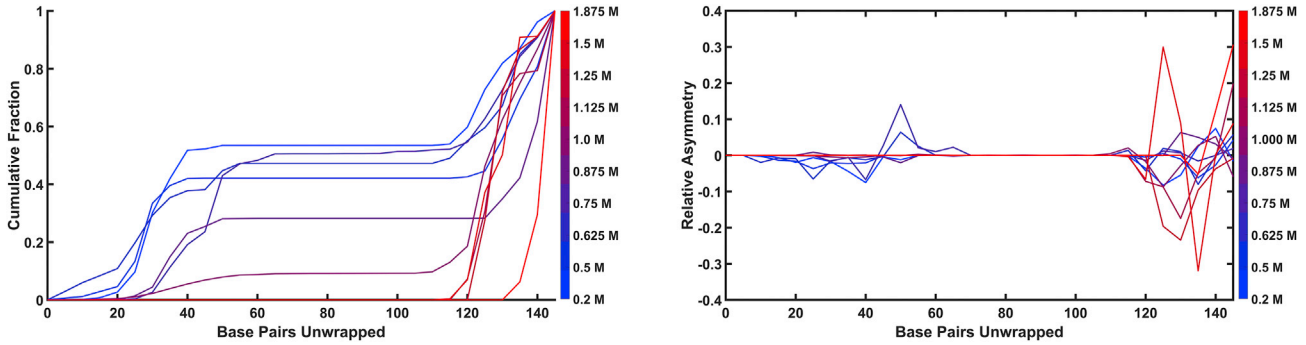

FIGURE 6 Cumulative unwrapping distribution (*left*) and relative asymmetry distribution (*right*) for 5S DNA. Salt concentration is denoted by color, moving from blue (at low salt) to red (at high salt).

5′ ATCCGAGCCCTATGCTGCTTGACTTCGGTGATCGGACGAGAACCGGTATATTCAGCATGGTATGGTCGTAGGCT C
3′ TAGAGGTCCCTAAATATTCGGCTACTGCAGTATTGTAGGGACTGGGAAATTTATCGAATTGAAAGTAGTTCGTT

FIGURE 7 5S rDNA Sequence. The second half of the sequence has been flipped right to left to align DNA ends. Flexible regions have been highlighted, and stiffer regions of interest have been underlined. The "left" end is the 5′ end in this reading, whereas the "right" end is the 3′ end.

The FRET decrease is detected at lower salt for the 5S sequence compared to the 601 and in addition is much sharper, consistent with fewer intermediate states. Lastly, the FRET curve for the 601 sequence is less cooperative (with increasing salt) than for the 5S sequence, suggesting that 601 release occurs in multiple steps with differing rates (this is most easily seen in the *blue curve* of Fig. 8). This sequential release of the histone components could result from the asymmetric unwrapping of the DNA at these salt concentrations.

The variations in unwinding reflect the different origin and function of the two DNA sequences. The 5S rDNA is a natural nucleosome-positioning sequence from a sea urchin (*L. variegatus*) and is part of a gene cluster that requires dynamic accessibility for transcription. In contrast, the Widom-601 sequence is the product of stringent SELEX (systematic evolution of ligands by exponential enrichment) selection to achieve unusually high nucleosome affinity beyond any known biological sequence (31). Thus, the apparent K_{eq} for the 601 sequence for the nucleosome octamer is 150 times greater than for the 5S sequence, consistent with the requirement for higher salt concentrations to induce DNA unwrapping (31).

The mechanics of the 5S unwrapping suggest a biological significance for the spring-latch mechanism in regulating gene availability. By allowing unwrapping to complete more easily after a tougher initiation, the ability of enzymes to remove the DNA from the nucleosome may be enhanced.

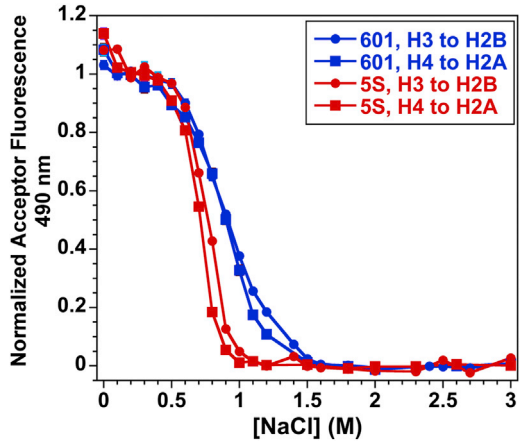

FIGURE 8 Salt-induced dissociation of H2A-H2B histone dimers from nucleosome core particles with 5S (*red*) and 601 (*blue*) DNA constructs. Two FRET pairs were used to monitor the transitions: H3 donor to H2B acceptor (*circles*) and H4 donor to H2A acceptor (*squares*). To see this figure in color, go online.

CONCLUSIONS

This study demonstrates that the inclusion of sequence information can enhance our understanding of how the mechanical properties of DNA affect its function. Using coarse-grained models that explicitly account for sequence in conjunction with SAXS data and ensemble-modeling tools, we detected and explained differences in nucleosome unwrapping observed for two different DNA sequences. Differences in DNA flexibility allow the two sides to be distinguished from one another for the first time using SAXS to our knowledge. This approach revealed the locations of stretches of DNA bases that rapidly unwrap from the histone core, suggesting a biological significance for these regions. Importantly, our findings are corroborated by FRET data on the same complexes. This new to our knowledge, technique provides a powerful tool to dissect the intrinsic effects on NCP dynamics of variations in NCP composition such as different DNA sequences and incorporation of histone variants. In future applications, this methodology has the potential to examine the mechanism of NCP disassembly under the influence of nucleosome chaperones and remodeling complexes.

SUPPORTING MATERIAL

Supporting Materials and Methods and 22 figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(18)30812-9.

AUTHOR CONTRIBUTIONS

A.W.M. implemented the cgDNA method for this system, carried out all of the computations described, and interpreted the results. J.M.T. carried out the SAXS experiments and helped analyze the data. L.M.G. provided the samples and supervised the FRET experiments. O.G. helped implement cgDNA. L.P. designed the research. All authors contributed to writing the manuscript.

ACKNOWLEDGMENTS

We acknowledge Yujie Chen for assistance in data acquisition at the Cornell High Energy Synchrotron Source (CHESS) and Traci Topping for sample preparation and collection of the initial FRET data.

A.W.M. is supported by the National Science Foundation through DGE-1650441. Additional financial support for this work was provided by the National Institutes of Health (R35- GM122514 to L.P.). CHESS is supported by the National Science Foundation and National Institutes of Health/National Institute of General Medical Sciences via National Science Foundation award DMR-0936384 to CHESS.

SUPPORTING CITATIONS

References (42–46) appear in the Supporting Material.

REFERENCES

1. Luger, K., A. W. Mäder, …, T. J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature.* 389:251–260.
2. Davey, C. A., D. F. Sargent, …, T. J. Richmond. 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. *J. Mol. Biol.* 319:1097–1113.
3. Li, G., and D. Reinberg. 2011. Chromatin higher-order structures and gene regulation. *Curr. Opin. Genet. Dev.* 21:175–186.
4. Lai, W. K. M., and B. F. Pugh. 2017. Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat. Rev. Mol. Cell Biol.* 18:548–562.
5. North, J. A., J. C. Shimko, …, M. G. Poirier. 2012. Regulation of the nucleosome unwrapping rate controls DNA accessibility. *Nucleic Acids Res.* 40:10215–10227.
6. Ramachandran, S., K. Ahmad, and S. Henikoff. 2017. Transcription and remodeling produce asymmetrically unwrapped nucleosomal intermediates. *Mol. Cell.* 68:1038–1053.e4.
7. Brehove, M., T. Wang, …, M. G. Poirier. 2015. Histone core phosphorylation regulates DNA accessibility. *J. Biol. Chem.* 290:22612–22621.
8. Li, G., M. Levitus, …, J. Widom. 2005. Rapid spontaneous accessibility of nucleosomal DNA. *Nat. Struct. Mol. Biol.* 12:46–53.
9. Widom, J. 2001. Role of DNA sequence in nucleosome stability and dynamics. *Q. Rev. Biophys.* 34:269–324.
10. Radman-Livaja, M., and O. J. Rando. 2010. Nucleosome positioning: how is it established, and why does it matter? *Dev. Biol.* 339:258–266.
11. Bowman, G. D. 2010. Mechanisms of ATP-dependent nucleosome sliding. *Curr. Opin. Struct. Biol.* 20:73–81.
12. Kelbauskas, L., N. Woodbury, and D. Lohr. 2009. DNA sequence-dependent variation in nucleosome structure, stability, and dynamics detected by a FRET-based analysis. *Biochem. Cell Biol.* 87:323–335.
13. Widlund, H. R., J. M. Vitolo, …, J. J. Hayes. 2000. DNA sequence-dependent contributions of core histone tails to nucleosome stability: differential effects of acetylation and proteolytic tail removal. *Biochemistry.* 39:3835–3841.
14. Gansen, A., K. Tóth, …, J. Langowski. 2009. Structural variability of nucleosomes detected by single-pair Förster resonance energy transfer: histone acetylation, sequence variation, and salt effects. *J. Phys. Chem. B.* 113:2604–2613.
15. Vinogradov, A. E. 2003. DNA helix: the importance of being GC-rich. *Nucleic Acids Res.* 31:1838–1844.
16. Nikolova, E. N., G. D. Bascom, …, H. M. Al-Hashimi. 2012. Probing sequence-specific DNA flexibility in a-tracts and pyrimidine-purine steps by nuclear magnetic resonance (13)C relaxation and molecular dynamics simulations. *Biochemistry.* 51:8654–8664.
17. Packer, M. J., M. P. Dauncey, and C. A. Hunter. 2000. Sequence-dependent DNA structure: tetranucleotide conformational maps. *J. Mol. Biol.* 295:85–103.
18. Ngo, T. T., Q. Zhang, …, T. Ha. 2015. Asymmetric unwrapping of nucleosomes under tension directed by DNA local flexibility. *Cell.* 160:1135–1144.
19. Lohr, D., R. Bash, …, S. Lindsay. 2007. Using atomic force microscopy to study chromatin structure and nucleosome remodeling. *Methods.* 41:333–341.
20. Andresen, K., I. Jimenez-Useche, …, X. Qiu. 2013. Solution scattering and FRET studies on nucleosomes reveal DNA unwrapping effects of H3 and H4 tail removal. *PLoS One.* 8:e78587.
21. Chen, Y., J. M. Tokuda, …, L. Pollack. 2017. Asymmetric unwrapping of nucleosomal DNA propagates asymmetric opening and dissociation of the histone core. *Proc. Natl. Acad. Sci. USA.* 114:334–339.
22. Chen, Y., J. M. Tokuda, …, L. Pollack. 2014. Revealing transient structures of nucleosomes as DNA unwinds. *Nucleic Acids Res.* 42:8767–8776.
23. Mihardja, S., A. J. Spakowitz, …, C. Bustamante. 2006. Effect of force on mononucleosomal dynamics. *Proc. Natl. Acad. Sci. USA.* 103:15871–15876.
24. Kulić, I. M., and H. Schiessel. 2004. DNA spools under tension. *Phys. Rev. Lett.* 92:228101.
25. de Bruin, L., M. Tompitak, …, H. Schiessel. 2016. Why do nucleosomes unwrap asymmetrically? *J. Phys. Chem. B.* 120:5855–5863.
26. Polach, K. J., and J. Widom. 1995. Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. *J. Mol. Biol.* 254:130–149.
27. Anderson, J. D., and J. Widom. 2000. Sequence and position-dependence of the equilibrium accessibility of nucleosomal DNA target sites. *J. Mol. Biol.* 296:979–987.
28. Koopmans, W. J., R. Buning, …, J. van Noort. 2009. spFRET using alternating excitation and FCS reveals progressive DNA unwrapping in nucleosomes. *Biophys. J.* 97:195–204.
29. Blanchet, C. E., and D. I. Svergun. 2013. Small-angle X-ray scattering on biological macromolecules and nanocomposites in solution. *Annu. Rev. Phys. Chem.* 64:37–54.
30. Tria, G., H. D. Mertens, …, D. I. Svergun. 2015. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ.* 2:207–217.
31. Lowary, P. T., and J. Widom. 1998. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* 276:19–42.
32. Simpson, R. T., and D. W. Stafford. 1983. Structural features of a phased nucleosome core particle. *Proc. Natl. Acad. Sci. USA.* 80:51–55.
33. Tokuda, J. M., S. A. Pabit, and L. Pollack. 2016. Protein-DNA and ion-DNA interactions revealed through contrast variation SAXS. *Biophys. Rev.* 8:139–149.
34. Blose, J. M., S. A. Pabit, …, L. Pollack. 2011. Effects of a protecting osmolyte on the ion atmosphere surrounding DNA duplexes. *Biochemistry.* 50:8540–8547.
35. Skou, S., R. E. Gillilan, and N. Ando. 2014. Synchrotron-based small-angle X-ray scattering of proteins in solution. *Nat. Protoc.* 9:1727–1739.
36. Petkevičiūtė, D., M. Pasi, …, J. H. Maddocks. 2014. cgDNA: a software package for the prediction of sequence-dependent coarse-grain free energies of B-form DNA. *Nucleic Acids Res.* 42:e153.
37. Eaton, M. L. 2007. Multivariate Statistics: A Vector Space Approach. Institute of Mathematical Statistics, Beachwood, OH.
38. Hoch, D. A., J. J. Stratton, and L. M. Gloss. 2007. Protein-protein Förster resonance energy transfer analysis of nucleosome core particles containing H2A and H2A.Z. *J. Mol. Biol.* 371:971–988.
39. Mackerell, A. D., Jr., and L. Nilsson. 2008. Molecular dynamics simulations of nucleic acid-protein complexes. *Curr. Opin. Struct. Biol.* 18:194–199.
40. Li, M., and M. D. Wang. 2012. Unzipping single DNA molecules to study nucleosome structure and dynamics. *Methods Enzymol.* 513:29–58.
41. Culkin, J., L. de Bruin, …, H. Schiessel. 2017. The role of DNA sequence in nucleosome breathing. *Eur. Phys. J. E Soft Matter.* 40:106.
42. Luger, K., T. J. Rechsteiner, …, T. J. Richmond. 1997. Characterization of nucleosome core particles containing histone proteins made in bacteria. *J. Mol. Biol.* 272:301–311.
43. Gloss, L. M., and B. J. Placek. 2002. The effect of salts on the stability of the H2A-H2B histone dimer. *Biochemistry.* 41:14951–14959.
44. Banks, D. D., and L. M. Gloss. 2003. Equilibrium folding of the core histones: the H3-H4 tetramer is less stable than the H2A-H2B dimer. *Biochemistry.* 42:6827–6839.
45. Dyer, P. N., R. S. Edayathumangalam, …, K. Luger. 2004. Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* 375:23–44.
46. Franke, D., M. V. Petoukhov, …, D. I. Svergun. 2017. *ATSAS 2.8*: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Cryst.* 50:1212–1225.