Local DNA sequence controls the cooperativity and asymmetry of DNA unwrapping from nucleosome core particles

Authors:
Alexander Mauney, School of Applied and Engineering Physics, Cornell University, Ithaca NY
Joshua Tokuda, School of Applied and Engineering Physics, Cornell University, Ithaca NY
Lisa Gloss, School of Molecular Biosciences, Washington State University, Pullman, WA
Oscar Gonzalez, Department of Mathematics University of Texas Austin, TX
Lois Pollack, School of Applied and Engineering Physics, Cornell University, Ithaca NY

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, in addition to other factors. Here we focus on how DNA sequence affects unwrapping from NCPs, and thus accessibility. Sequence variations impact protein binding, are crucial for understanding transcription, and determine how DNA is organized into chromatin. Small angle x-ray scattering with contrast variation (CV-SAXS) 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-grain Monte-Carlo methods to generate realistic sequence-dependent unwrapped structures for the nucleosomal DNA, and to model thermal variations. These structures are used in conjunction with ensemble optimization methods to determine the composition of the overall ensemble as electrostatic interactions are selectively 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 \( \approx 145 \) base-pairs 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 pseudo-symmetric helical ramp that organizes DNA(2). In the fully wrapped structure, the negatively charged DNA (with a persistence length of \( \approx 500 \) Å) is tightly bent by positively charged residues of the histone core into \( \approx 1.7 \) superhelical turns with a diameter of \( \approx 100 \) 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, chromatin remodelers). However, the inherent biochemical and mechanical properties of nucleosomes themselves are essential to this activity.
Much past work has been devoted to the effect of histone variants or post translational modification on stability(5, 6). The DNA sequence also contributes to overall NCP stability (7, 8), is critical in histone positioning(8, 9) and sliding(10), contains hot spots for protein binding(10–12), and provides sites for epigenetic marks(13). In addition, DNA sequences encode mechanical features to facilitate its packaging and control of 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)(14). Poly-A stretches contribute to especially rigid conformation(15). Finally, dinucleotide steps of pyrimidine-purine (CA, CG, TA, TG) are known to be the most conformationally flexible(16).

Despite the demonstrated importance of DNA sequence, most investigations of nucleosome unwrapping using optical tweezers(17), atomic force microscopy (AFM)(18), Forster resonance energy transfer (FRET)(11, 17, 19), or small angle x-ray scattering (SAXS) (19–21) treat the DNA as a uniform polymer. Notably, recent studies have highlighted the role of DNA flexibility in directing how NCPs unwrap(17). 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 sampled during unwrapping. In particular, this method forces one half of the DNA to fully unwrap before the other can begin(22), as simultaneous unwrapping from both sides would lead to sharp bends in the DNA. Nevertheless, these studies underscore the influence of DNA sequence in tuning the mechanical tendencies of nucleosomes. What remains to be observed is the behavior of nucleosomes in an unconstrained environment.

The combination of small angle x-ray scattering (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 for each of the conformations present in solution(23). In order to characterize the global structural parameters beyond the average radius of gyration or shape ensembles, advanced tools have been developed to determine distributions of conformations that describe the SAXS data. One such approach, ensemble optimization method (EOM), allows selection of a subset of structures (“ensemble”) that best recapitulates the SAXS data from a large pool of possible structures(24). The success and reliability of this approach depends 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(20, 21). Since 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(20, 21). This illustrated a new method for identifying the multiple DNA conformations present in solution, it did not account for the DNA sequence dependent effects that are relevant to regulation or NCP formation in vivo.

Here, we describe a method that explicitly accounts for DNA sequence in ensemble studies of NCPs. Coarse grain 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. We apply this approach to compare the salt-induced unwrapping of two DNA sequences: the SELEX engineered Widom 601(25), and the natural SS rDNA from Lytechinus variegatus(26). Analysis of selected structures reveals newly discovered links between the conformations populated and the underlying mechanical properties of the DNA. These insights may be useful in predicting how and where nucleosomal DNA becomes accessible in vivo.
Results

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 Methods). Because the SAXS signal is proportional to the electron density differences 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(26) or the artificially engineered Widom 601 sequence(25). Static SAXS data were acquired in equilibrium salt titrations, with NaCl concentrations ranging from 0.2 M through 1.8 M. At low salt, the DNA in NCPs is fully wrapped. Partial release occurs with increasing [NaCl]. 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 ensemble optimization method (EOM) to determine which DNA conformations are present in the SAXS data for a particular construct, at a given salt concentration. An overview of the EOM method is illustrated in Figure 1 with a brief description provided in the caption.

Figure 1, Schematic representation of Ensemble Optimization Method (EOM). The first step is computing the theoretical scattering profiles for each structure in a model pool using CRYSTOL. A genetic algorithm is then used to determine the ensemble that best matches the data. The genetic algorithm is repeated multiple times and the final ensembles combined to give an overall distribution of structures.

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 protein data bank, 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(27). Protein-Nucleic Acid complexes are even more complicated; solved
structures may not exist, and MD simulations are even more computationally difficult if available(27). New methods are therefore required to construct complete structural pools. In past work DNA was unwrapped by replacing nucleosomal DNA with linear B-form DNA, limiting the modeling of unwrapped DNA to stiff extensions along the trajectory of the nucleosomal DNA, and 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 1AOI crystal structure(1) (Figure 2a), the location and orientation of each base is found in order to coarse-grain 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 base pair is appended to each end to match the experimental DNA (Figure 2b). A unique stiffness matrix is calculated based on the sequence, using the mechanical properties of each base-pair and base-pair—base-pair junction. This matrix is employed to identify the lowest energy configuration of the appended bases (Figure 2c). This configuration comprises the fully wrapped nucleosome, the starting structure for model generation.

![Figure 2](image)

Figure 2, Representations of steps taken to generate nucleosome models. a) Crystal structure of wrapped nucleosome. b) Extracted locations of bases and backbone with added bases. c) Lowest energy state for 40 bases free on each end. d) 10 thermal variations with 40 free bases on each end.

Our full structural pool is generated from this starting structure by varying the number of free bases on each side of the nucleosome. We designate these sides as “left” and “right” based on a 5’ to 3’ reading of the sequence of interest, and the number of free bases is designated by \(n_L\) and \(n_R\), respectively. Once a base 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 loaded into the pool and classified based on the total number of bases unwound, \(n_T = n_L + n_R\). Finally, atomistic structures of the individual bases are properly placed into the coarse grain model to generate the final structure of nucleosomal DNA.

This process is iterated for values of \(n_L\) and \(n_R\) that span the entire possible unwrapping space, where 5 bases at a time are released. The final pool for each sequence contains 5104 model structures.

With the pool fully in place, EOM can be applied to select an ensemble of DNA conformations whose SAXS profile recapitulates the experimental data. SAXS profiles acquired at each salt concentration were
processed through the algorithm to identify structures. Due to the inherent degeneracy in the model pool, the fitting procedure was run 700 times in order to get a statistical distribution of structures. The fit quality, measured using $\chi^2$, was consistent across these multiple runs.

The selected structures are classified according to: total number of bases unwrapped ($n_T$), bases unwound from the left end ($n_L$) and bases unwound from the right end ($n_R$). The differences between $n_L$ and $n_R$ reveal the asymmetry of unwrapping. This parameter is of great interest in determining any sequence dependent bias. To allow interpretation of the degree of unwinding, as well as any left/right biasing, Figure 3 shows the structures in histogram form. The structures depicted here were selected from the pool that recapitulate one particular data set: 601 DNA sequence at 0.5 M NaCl. At this intermediate salt value, a wide variety of structures are present in the selected ensemble. The x-axis of the histogram represents the total number of bases unwrapped ($n_T$). Interestingly, no structures were selected with fewer than 35 unwound base-pairs. A cluster of structures is found with between 35 and 70 unwind base-pairs, and a second broad group of structures contains more than 85. To compare the structures at each value $n_T$ we 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). If one side has 10 (or more) bases unwound than the other side, the unwinding is considered asymmetric: one side is significantly favored. This threshold for determining if one side is more unwound was selected as it is the smallest that generated consistent results between repetitions with different pools. Because a step size of 5 was used for the unwrapping, differences smaller than 10 are the result of a single step and thus may be spurious. This approach highlights the presence of asymmetry, and indicates which end is favored, if any.
Figure 3. Illustration of Analysis Protocol. Top: histogram of unwrapping states in one condition (601 in 1 M NaCl) colored by dominant side (blue: left dominant, green: no dominant side, yellow: right dominant). Middle: Relation of the Cumulative Unwrapping Distribution to unwrapping histogram. The cumulative graph is the integral of histogram. Bottom: Relation of the Relative Asymmetry Distribution to the dominant side portions of the histogram. The right dominant fractions have been shown as negative to illustrate that the Relative Asymmetry Distribution is the subtraction of the right dominant portion from the left dominant portion.
The representation of Figure 3 (top) makes it challenging to compare across conditions (different constructs, or different salts). To enable such a comparison, we calculate the cumulative unwrapping distribution of the unwrapping histogram by integrating across the number of bases unwrapped ($n_T$). The overall progress of unwrapping becomes more apparent in this form, and changes in unwrapping value are reflected by the slope of the cumulative unwrapping distribution. Conditions where most structures are wrapped move toward a Cumulative Fraction of 1 at low $n_T$, while more unwrapped conditions approach a Cumulative fraction of 1 at higher $n_T$. This approach reports on the total number of bases unwrapped, but does not report on the asymmetry of the unwrapping. In order to determine whether one side is preferentially unwrapped we carry out another analysis: subtracting the fraction of right favored structures from that of the left favored structures at each total unwrapping number, which we define as the relative asymmetry distribution. The presence of peaks in this difference plot indicates portions of the unwrapping process where one side is more likely to unwrap than the other. Thus, the data of Figure 3 (top) can be further processed to determine both a cumulative unwrapping distribution (middle) and relative asymmetry distribution (bottom).

**Widom 601**

![Cumulative Unwrapping Distribution (left) and Relative Asymmetry Distribution (right) for Widom 601. Salt concentration is denoted by color, moving from blue (at low salt) to red (at high salt).](image)

We applied the above approaches to analyze the unwrapping behavior of the Widom 601 sequence, acquired in a static salt series. Each different SAXS profile was analyzed by EOM to identify structures from the full pool. The results are displayed in Figure 4. For the Widom 601 sequence, we can see distinct patterns of behavior. The cumulative unwrapping distribution shows the progressive increase in unwrapping as the salt is increased. At low salt, the unwrapping is highly asymmetric, but seems to alternate between sides. Unwrapping occurs on both sides simultaneously, however peaks in the relative asymmetry distribution indicate that the two sides unwrap by different amounts at a given salt concentration. At a particular point in the unwrapping (about 65 bases in) a left-biased state is much more likely than a right-biased one, and this trend persists over many salt concentrations. For nearly fully unwrapped states there is little difference between left-biased and right-biased states. They are both essentially free DNA with slight curvature from the fixed bases, which leads to the noise that can be seen at high $n_T$.

---

1 One curve (625 mM) was removed from the dataset due to poor fitting, likely due to sucrose mismatch between buffer and sample.
The cumulative unwrapping distribution, provides additional insight into how the DNA is released from the histone core. The curve is relatively flat for 40-60 bases unwound, suggesting that there are very few structures just before this peak in asymmetry. For larger unwinding values, in the range where the bias toward left-biased states is observed (60-70), the slope increases suggesting that there is a region that unwraps all together once it is no longer held on by preceding bases. We propose that this region acts like a spring-loaded latch; it can be held closed by a latch (the earlier bases), however once it is no longer constrained it pops open (off of the nucleosome).

Figure 5, Widom 601 Sequence. The second half of the sequence has been flipped right to left to show symmetry. Flexible regions have been highlighted, and a stiffer region of interest has been underlined.

Examination of the DNA sequence reveals information about the base composition that gives rise to this behavior. Most DNA is relatively stiff; however, pyrimidine-purine (YR) steps give rise to more flexibility[15], and these have been highlighted in Figure 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. Looking at the 601 DNA with those sequences highlighted, we can see that there is one particular spot that has nearly no flexible motifs. The positioning of this region is about 30 bases in, and the region is about 20 bases long (underlined in Figure 5). These numbers are consistent with the position of the peak in asymmetry at around 60-70 total bases unwrapped. We conclude that the unwrapping proceeds symmetrically and relatively steadily up to a point where around 20 bases have been freed from each side. At this point the spring-latch mechanism kicks in and the next 20 bp on the left side is rapidly released. 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

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).

The same analysis can be applied to the 5S DNA sequence. Results are shown in Figure 6. Again, we observe that unwrapping increases with salt concentration. Interestingly, there is a clear region, near the middle of the unwrapping range, with no states represented. From 50 bp unwound to 90 bp unwound the graph is flat for all salt concentrations. For many of the curves this flatness continues until 120 bp are unwrapped. This behavior indicates that the nucleosome transitions rapidly from a mostly wrapped state to a mostly unwrapped state. In contrast to the 601 DNA, the 5S sequence does not display any preference for initial unwrapping from one side or the other as there are no consistent peaks.
in the relative asymmetry distribution. Examining the sequence, we observe that there are two large stiff regions about 35 bp in from each side, underlined in figure 7. These stiff regions are both preceded by relatively flexible regions, and end at 55 bp and 65 bp from each end. These regions and the unwrapping behavior support the idea that the unwrapping proceeds steadily up through the flexible region. Once the flexible region is unwrapped the stiff region rapidly unwraps, leading to near full unwrapping.

5’ ATCCGAGCCCTGTCTGCTTCGACTCTGGTGATGGAGAGAACCCGCTATATTGACATGCCGATGTCGTAAGCCTGCT
3’ TAGACGTGGCTAAATATCCTGGCTACTGCAATATTGAAAGGGATTTAATCGAATTGAAAAAGTTAACGTCTGT

Figure 7, 5S rDNA Sequence. The second half of the sequence has been flipped right to left to show symmetry. Flexible regions have been highlighted, and stiffer regions of interest have been underlined.

Comparing the two sequences we see that the 5S sequence is more unwrapped at any given salt concentration than the 601 sequence. The 5S sequence also has many fewer partially unwrapped states represented compared to the 601 sequence. They both contain asymmetric structures, however the structures that they contain represent distinct unwinding states.

![Figure 8](image)

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 gain additional insight into the differences reported above, 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 report a disruption of the histone core(28). 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, thus corresponds to a fully formed histone core. Decreases in fluorescence correspond to disruption of the core, which is coupled to DNA unwrapping(20). The two constructs behave quite differently when assessed with this metric. 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 than for the 5S sequence, suggesting that 601 release occurs in multiple steps with differing rates (this is most easily seen in the dark blue). 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 sea urchin (Lytechinus variegatus), and is part of a gene cluster that requires dynamic accessibility for transcription. In contrast, the Widom-601 sequence was the product of stringent SELEX selection to achieve unusually high nucleosome affinity, beyond any known biological sequence(25). 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.

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.

**Conclusion**

This study demonstrates the power of sequence-based modeling of nucleic acids to reveal previously unresolvable details. These models allow for the fitting of experimental SAXS data more accurately, and should give a better understanding of the populations present in a sample. Using this method differences in nucleosome unwrapping were observed between 5S rDNA and Widom 601 DNA. These unwrapping profiles are corroborated by FRET data on the same complexes. Based on the structural information implicit in our modeling technique features in the unwrapping profiles could be related back to sequence-based mechanical properties. The locations of specific regions of rapid unwrapping suggest that there may be biologically significance to these structural features. This new 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; furthermore, this methodology, coupled with CVSAXS, could be extrapolated to examine the mechanism of NCP disassembly under the influence of nucleosome chaperones and remodeling complexes.

**Methods**

**SAXS**

SAXS provides low resolution structural information such as the pair-distance function, radius of gyration, and shape envelopes for proteins and/or nucleic acids in solution. The scattering intensity $I$ is measured as a function of the momentum transfer $q = \frac{4\pi}{\lambda} \cdot \sin(\theta)$ where $\lambda$ is the x-ray wavelength and $2\theta$ 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 which reflect contributions from both components(29).
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.

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 (21, 30).

**SAXS Data Collection**

SAXS data were collected at 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 through a semi-transparent molybdenum beamstop was used at 10.5 keV. Scattered x-rays were imaged onto a PILATUS 100K (DECTRIS) detector 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, with the distance calibrated using a silver-behenate standard. 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 (MathWorks) according to established protocols (31). Background subtraction was achieved by subtracting the profile measured for the matching buffer from that of the sample.

**Ensemble Optimization Method**

The Ensemble Optimization Method (EOM) starts by creating a number of random ensembles from the structure pool of a given size. The scattering profiles of these ensembles are then compared to the experimental profile to determine a fitness parameter (χ²). Ensembles with good fitness parameters are then moved into the next generation, where new ensembles are created by randomly changing elements of previous ensembles, or by swapping sections from one ensemble with another. These are then evaluated as before, and the best again advance to the next generation. After many generations, the best fitting ensemble is chosen as the final ensemble. This process was repeated many times with new starting ensembles to determine the probability that a given unwrapping parameter was present.

The pool generating program is based on the cgDNA suite (32). cgDNA provides a direct, explicit prediction of the sequence-dependent free energy and associated equilibrium distribution of an oligomer of B-form DNA of arbitrary sequence in solvent under prescribed environmental conditions. The cgDNA model is of the rigid-base type in which each base on each strand of a DNA oligomer is considered as a rigid entity interacting with its nearest neighbors. The coarse-grain configuration of an oligomer is described by a set of internal coordinates corresponding to the relative, three-dimensional rotation and displacement between neighboring bases both along and across the two backbone strands.

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 which describes energetic couplings between bases; the ground state configuration depends non-locally 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 then 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. While only one
set of solvent conditions could be modeled in this version of cgDNA(32), corresponding to approximately 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 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 are 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(33). Beginning from the 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.

FRET

The protein-protein FRET system has been described previously, as well as sample preparation, data collection and spectral analysis(20, 28).The only significant difference from previous equilibrium studies was that FRET data were collected for individual samples at NCP concentrations of 25 nM, rather than extensive titrations(28). 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.

Funding

A.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 (NIH) (R35- GM122514 to L.P.). CHESS is supported by National Science Foundation (NSF) & National Institutes of Health/National Institute of General Medical Sciences (NIH/NIGMS) via NSF award [DMR-0936384 to CHESS]. Funding for open access charge: NIH R35-GM122514.

Acknowledgements

We acknowledge Yujie Chen for assistance in data acquisition at CHESS and Traci Topping for sample preparation and collection of the initial FRET data.

References


