In silico evidence for sequence-dependent nucleosome sliding

Joshua Lequieu[a], David C. Schwartz[b,c,d], and Juan J. de Pablo[a,e,1]

[a]Institute for Molecular Engineering, University of Chicago, Chicago, IL 60637; [b]Laboratory for Molecular and Computational Genomics, Department of Chemistry, University of Wisconsin–Madison, Madison, WI 53706; [c]Laboratory of Genetics, University of Wisconsin–Madison, Madison, WI 53706; [d]UW-Biotechnology Center, University of Wisconsin–Madison, Madison, WI 53706; and [e]Materials Science Division, Argonne National Laboratory, Argonne, IL 60439

Edited by David A. Weitz, Harvard University, Cambridge, MA, and approved September 11, 2017 (received for review April 8, 2017)

Nucleosomes represent the basic building block of chromatin and provide an important mechanism by which cellular processes are controlled. The locations of nucleosomes across the genome are not random but instead depend on both the underlying DNA sequence and the dynamic action of other proteins within the nucleus. These processes are central to cellular function, and the molecular details of the interplay between DNA sequence and nucleosome dynamics remain poorly understood. In this work, we investigate this interplay in detail by relying on a molecular model, which permits development of a comprehensive picture of the underlying free energy surfaces and the corresponding dynamics of nucleosome repositioning. The mechanism of nucleosome repositioning is shown to be strongly linked to DNA sequence and directly related to the binding energy of a given DNA sequence to the histone core. It is also demonstrated that chromatin remodelers can override DNA-sequence preferences by exerting torque, and the histone H4 tail is then identified as a key component by which DNA-sequence, histone modifications, and chromatin remodelers could in fact be coupled.

nucleosome repositioning | chromatin dynamics | molecular simulation | advanced sampling techniques

The basic building block of eukaryotic chromatin is the nucleosome, a DNA–protein complex containing 147 bp of DNA wrapped around a disk-like protein complex known as the histone octamer (1). Since nucleosomal DNA is inaccessible to other DNA-binding proteins, such as transcription factors and polymerases (2–4), the locations of nucleosomes represent an important mechanism by which cellular processes are controlled. Notably, nucleosome positions are dynamic, with changes in transcription levels, cellular state, and environmental factors resulting in different packagings of chromatin (5, 6). Proper packaging of genomic DNA is critical to cellular function, and a wide range of human diseases have been associated with defects in chromatin structure (7, 8). Understanding the molecular factors that control the locations of nucleosomes, and how they are dynamically modulated, therefore represents a central goal of molecular biology and biophysics.

It is now appreciated that the DNA sequence itself represents a key factor that governs the locations of nucleosomes. Different DNA sequences exhibit different affinities for the histone proteins, and as such, they form nucleosomes with probabilities that can differ by several orders of magnitude (9, 10). The dependence of nucleosome locations on DNA sequence originates from subtle differences in the intrinsic shape and flexibility of a specific DNA sequence, which lead to favorable electrostatic interactions between the DNA backbone and residues on the histone surface (11). In fact, this pronounced dependence on DNA sequence has led several authors to propose that a genetic code exists (12, 13) where the positions of 50% of nucleosomes in vivo are dictated by DNA sequence alone. Such a view, however, is not without controversy (14, 15), and a better understanding of the underlying processes must be developed.

Given that DNA sequence is largely constant throughout the life of a cell, other mechanisms must also be at play to achieve the dynamic nucleosome repositioning necessary for cellular function. One part of this dynamic regulation is accomplished by chromatin remodelers, ATP-dependent proteins within the nucleus that actively reposition nucleosomes along the genome (16). Chromatin remodelers can facilitate many different modifications to chromatin by positioning (17) or removing (18, 19) nucleosomes from promoters or by evenly spacing nucleosomes across sections of DNA (20–22). They are central to genetic compaction because, in their absence, nucleosome diffusion is extremely slow, typically on the time scale of hours (23–25). Chromatin remodelers accelerate the packaging of DNA by moving nucleosomes away from their sequence-directed binding locations into metastable positions. The mechanism by which these metastable nucleosomes relax back to their equilibrium locations represents an important process through which DNA sequence and chromatin dynamics are coupled.

Several studies have sought to elucidate the mechanism by which DNA repositions around the nucleosome. One group of studies has led researchers to propose a "loop propagation" model, where DNA loops are introduced into one side of the nucleosome and then move along the histone core in an inchworm-like manner (26–31). Another group of studies has led authors to propose a "twist diffusion" model, in which a twist defect in the natural helicity of nucleosomal DNA is first introduced and then diffuses around the nucleosome in a

Significance

The dynamic compaction of DNA into chromatin is essential for gene expression. Errors during compaction are associated with numerous diseases. Several molecular factors are known to affect chromatin dynamics, but their relative importance and the interplay between them are poorly understood. A detailed molecular model is used here to examine chromatin dynamics at the level of its most fundamental building block, namely the nucleosome. Nucleosome dynamics are demonstrated to be encoded in the DNA sequence itself, and key fundamental factors are uncovered that can significantly alter these dynamics at the molecular level. The results serve to complete a hitherto unavailable description of nucleosome dynamics by introducing previously unappreciated molecular processes, with the potential to influence macroscopic chromatin structure and genetics.

Author contributions: J.L., D.C.S., and J.J.d.P. designed research; J.L. performed research; J.L. contributed new reagents/analytic tools; J.L. and J.J.d.P. analyzed data; and J.L. and J.J.d.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: depablo@uchicago.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1705685114/-/DCSupplemental.

corkscrew-like motion (32–35). Several single-molecule measurements suggest a repositioning mechanism that cannot be classified into any of these models (36, 37). There is considerable evidence in support of each model, and it is likely that these repositioning mechanisms are not mutually exclusive but arise depending on other factors, including the DNA sequence (38). Because chromatin is inherently dynamic, elucidating the mechanisms by which nucleosomes are rearranged, both in the presence and absence of chromatin remodelers, is of considerable scientific interest.

Both sequence and dynamics do matter, but the relationship between these factors has been difficult to elucidate by relying exclusively on experiments. A major obstacle to developing a comprehensive picture of nucleosome repositioning is the lack of sequence-dependent studies on the mechanism of chromatin repositioning. Some of the effects of DNA sequence on nucleosome mobility have been explored (21, 36, 39). However, the vast majority of studies on nucleosome repositioning (22, 37, 40–44) have only considered the 601 positioning sequence (9), which exhibits a particularly strong affinity for histones. Such studies have led to valuable insights, but it remains unclear if or to what extent observations pertaining to the 601 sequence can be generalized to other, naturally occurring, nonsynthetic, DNA sequences. Note that the lack of sequence-dependent studies even extends to high-resolution crystal structures of the nucleosome, which are only available for several DNA sequences (45). Several models now assume that a bias exists toward nucleosome affinity depending on sequence, but little work (38, 46) has been done to understand the effects of sequence on the dynamics of nucleosome positions and how DNA sequence might impact the underlying mechanism of nucleosome rearrangement.

In this work, we investigate the interplay of DNA sequence and nucleosome repositioning dynamics in detail, by using a molecular model of the nucleosome. By relying on a variety of advanced simulation techniques, we characterize the effect of DNA sequence on both the free energies of nucleosome arrangement as well as the time scales over which they occur. Our results indicate that different DNA sequences do indeed rely on different mechanisms to reposition, through pathways reminiscent of both the loop propagation and twist diffusion models described above. However, our results demonstrate that the original formulation of these previously proposed mechanisms is incomplete and identify several molecular details of the histone surface that play crucial roles in repositioning. Lastly, we examine the effect of applied forces on nucleosomal dynamics and suggest that nucleosome remodelers can override certain sequence-based positioning preferences by applying torque to nucleosomal DNA. Taken together, our results serve to provide a more comprehensive picture of the effect of DNA sequence on nucleosome repositioning dynamics than was previously available and will help develop an understanding of the dynamic molecular processes that occur within chromatin.

Results

To examine nucleosome repositioning, we rely on a coarse-grained model of the nucleosome that combines the "3 Site Per Nucleotide" (3SPN) model of DNA with the "Atomic-Interaction-based Coarse-Grained" (AICG) model of proteins (47, 48). By combining detailed, fully validated models of DNA (49) and proteins (50), the 3SPN-AICG combination has been demonstrated to accurately reproduce experimental measurements of both the tension-dependent and sequence-dependent binding free energies of nucleosome formation without introducing adjustable parameters (48). Importantly, no information from the nucleosome crystal structure (51) or locations of DNA–histone contacts are rigidly fixed in our model; instead, all energies, conformations, and dynamics arise naturally from the DNA and protein force fields, which include explicit Coulombic interactions between the negatively charged DNA phosphate groups and the charged residues on the histone proteins (see *Materials and Methods*). As such, the 3SPN-AICG model can be viewed as truly predictive and is an appropriate choice for a molecular-level investigation of DNA repositioning.

Two representative 3SPN-AICG molecular configurations of relevance to nucleosome sliding are shown in Fig. 1*A*. They contain 223 bp of DNA, of which 147 bp are initially incorporated into the nucleosome, with 38 flanking base pairs on each side. Since we are primarily concerned with the mechanism of small rearrangements of DNA ($<$20 bp), 38 flanking bases were used to minimize boundary effects that might arise from the free DNA ends. To characterize the degree of nucleosome sliding, an order parameter is defined, S_T, which represents the *translocational* position of the central base pair relative to the histone dyad (Fig. 1*B* and *Materials and Methods*). Here S_T is given in units of base pairs (bp), with $S_T \approx 0$ bp corresponding to the native binding position of a DNA sequence and $S_T = \pm 10$ corresponding to a translocation of the DNA by one helical turn forward (+) or backward (−) relative to the histone dyad.

To characterize the time scales at play during nucleosome repositioning, we first compute the mean-squared displacement (over short to intermediate times), $\langle(\Delta S_T)^2\rangle$, of DNA around the histone core for two different DNA sequences (Fig. 2). In this analysis, time is reported in units of τ, the characteristic time scale of DNA unwrapping (see *Materials and Methods*). The first sequence, denoted "601," contains the strongly positioning 601 sequence discovered by Lowary and Widom (9). In contrast, the second sequence is a "TTAGGG" repeat, a sequence found in the telomeres of human chromosomes (52, 53) that positions nucleosomes poorly (39, 54). Both sequences are characterized by anomalous subdiffusion, with scaling exponents ranging from 0.4 to 0.69 for all time scales below τ. Subdiffusion here was expected; entire nucleosomes are known to reposition on the time scale of minutes to hours (23–25) and would not be

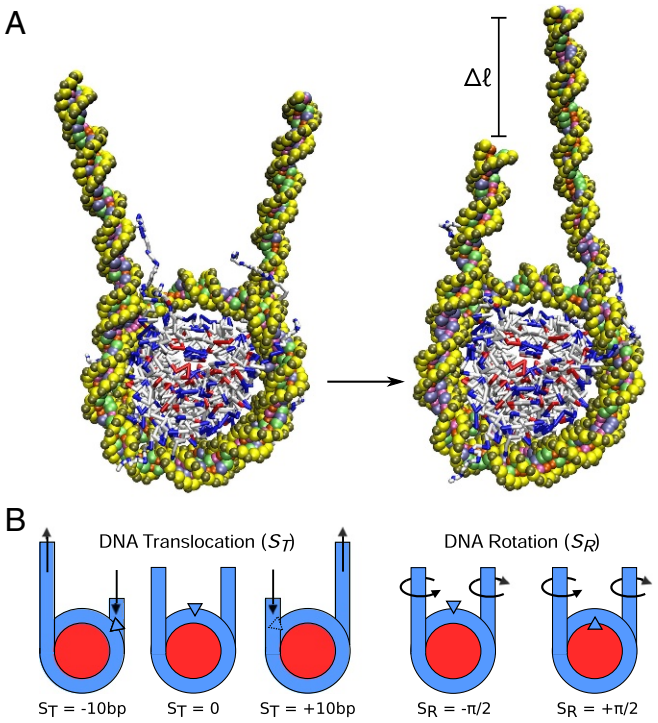

Fig. 1. Molecular configurations of nucleosome repositioning. (*A*) Molecular representation of repositioning using coarse-grained nucleosome model. The chief aim of this study is to characterize the molecular mechanism by which this repositioning occurs. (*B*) Order parameters used to characterize DNA translocation, S_T, and rotational position, S_R, relative to the histone proteins.

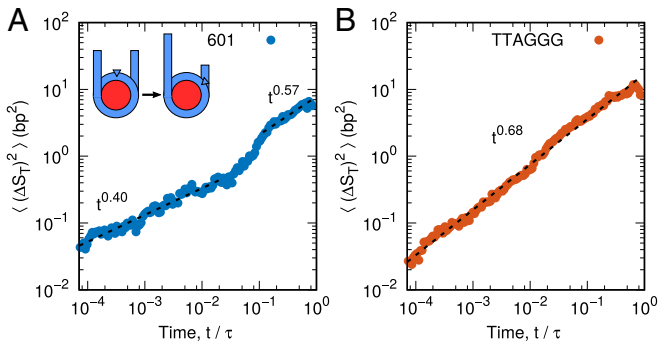

Fig. 2. Mean-squared displacement of nucleosomal DNA around histone proteins for (*A*) 601 positioning sequence and (*B*) TTAGGG repeat. Both DNA sequences exhibit anomalous diffusion, but the 601 positioning sequence contains two dynamic regimes, while the TTAGGG repeat only contains one. This difference suggests that DNA with different sequences might translate using different mechanisms. Averaging was performed over 100 independent molecular trajectories.

expected to exhibit diffusive motion ($\langle (\Delta S_T)^2 \rangle \sim t$) on the time and length scales considered here.

What was not expected, however, was the qualitative differences between the dynamics of the two different sequences. Whereas the TTAGGG repeat is characterized by a single dynamic regime (i.e., $\sim t^{0.68}$ for all times), the 601 sequence is observed to demonstrate two dynamic regimes: a strongly subdiffusive regime ($\sim t^{0.40}$) for short time scales $< 10^{-2}\tau$, and a weakly subdiffusive regime ($\sim t^{0.57}$) at longer times $> 10^{-1}\tau$. These results suggest that the 601 sequence might reposition via a slow mode at short times, characterized by little DNA motion, followed by a faster mode at longer times, where DNA repositions more rapidly. More generally, the fact that motion of the TTAGGG repeat exhibits one time scale, whereas the 601 sequence exhibits two, leads us to hypothesize that these two DNA sequences reposition via different mechanisms.

Sequence-Dependent Nucleosome Sliding To explore this possibility, we focus on the free energy surface of DNA repositioning. An additional order parameter is introduced, S_R, which corresponds to the *rotational* position of DNA relative to the histone proteins. S_R measures what side of the DNA double helix is facing the histone core, with $S_R = \pm\pi/2$ corresponding to the major groove (+) or the minor groove (−) facing the protein core (see Fig. 1*B* and *Materials and Methods*). Combined with S_T, which quantified the translocational position of DNA, these two order parameters provide a relatively complete description of the position of DNA wrapped around the histone core.

Fig. 3 shows the free energy surfaces as a function of S_T and S_R for the 601 sequence and the TTAGGG repeat. As with the mean-squared displacement measurements, the two free energy surfaces are considerably different. The free energy surface of the 601 sequence reveals a strong tendency for both translocational and rotational positioning. More specifically, the free energy minimum corresponds to translocational DNA positions with the 601 sequence centered on the nucleosome ($S_T \approx 0$) and with the minor groove facing the histone core ($-\pi/2 > S_R > 0$). DNA movement away from these strongly bound configurations leads to large free energy penalties. The location of this minimum corresponds well with the experimentally established position of the 601 sequence (55) (see Fig. S1) and serves as further validation that the 3SPN-AICG model can accurately capture DNA sequence effects within the nucleosome. The free energy surface of the TTAGGG repeat lacks the pronounced translocational and rotational positioning preferences of the 601 sequence and is characterized by a diffuse and relatively flat free energy surface.

Given these differences, it is reasonable to anticipate that these two DNA sequences would reposition through different mechanisms. To infer such mechanisms from the free energy surfaces, we rely on the "String Method," as implemented in the SSAGES package (see *Materials and Methods*), to calculate the minimum energy path corresponding to a 20 bp rearrangement from ξ_0 to ξ_1 (Fig. 4). Since this minimum-energy path corresponds to the most probable transition from ξ_0 to ξ_1, the resulting string can be used to infer the most likely mechanism of DNA rearrangement.

For the 601 sequence (Fig. 4*A*), we observe that the mechanism of DNA repositioning is almost independent of rotational position. The minimum energy path is characterized by two distinct modes: One is characterized by increasing S_T at constant S_R, and the other is characterized by changing S_R at near-constant S_T. Notably, these two modes alternate semiperiodically, with the first mode always followed by the second (and vice versa). The free energy along this path (Fig. 4*B*) indicates that the regions corresponding to this first mode are associated with large free energy barriers (e.g., $\xi = 0.05, 0.4, 0.8$), whereas regions corresponding to the second mode have a flat free energy surface (e.g., $\xi = 0.2, 0.6$). This mechanism is consistent with the loop propagation model described above, where DNA translocation is achieved independently of DNA rotation.

In contrast, the minimum energy path for the TTAGGG repeat exhibits a strong coupling between DNA translocation, S_T, and DNA rotation, S_R (Fig. 4*C*). In fact, the relationship between S_T and S_R observed here corresponds exactly to the 10 bp pitch of DNA: One complete rotation of DNA leads to a translocation of 10 bp. This translocation–rotation coupling is similar to the twist diffusion model, in which DNA repositions via a corkscrew-like motion, where DNA translocation is accompanied by DNA rotation to maintain minor-groove contacts with the histone protein. The energy surface corresponding to this mechanism is rough, with barriers $\approx 2k_B T$, but largely uniform, with no global translocational positioning preferences across the 20 bp region considered here (Fig. 4*D*).

In previous work, Schiessel et al. estimated that the energy barriers of DNA loop formation are $\approx 20k_B T$ (26), whereas other work by Kulíc and Schiessel estimate the energy barriers of twist defects to be 1 to 2 $k_B T$ for some DNA sequences and $8 - 10k_B T$ for others (27). Other work by Tolstorukov et al. has estimated barriers that are much higher at $\approx 50 - 100k_B T$ (56). The energy barriers reported in our work are comparable to the predictions of refs. 26, 27 and are somewhat lower than those of ref. 56. We note that these past models rely heavily on

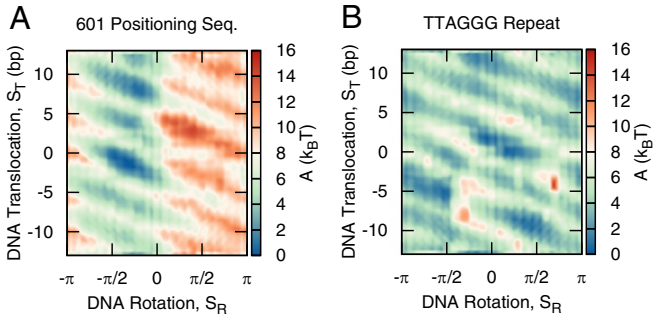

Fig. 3. Free energy surface of DNA repositioning for (*A*) 601 positioning sequence and (*B*) TTAGGG repeat. The 601 sequence demonstrates strong rotational, S_R, and translocational, S_T, positioning preferences, whereas the TTAGGG repeat does not.

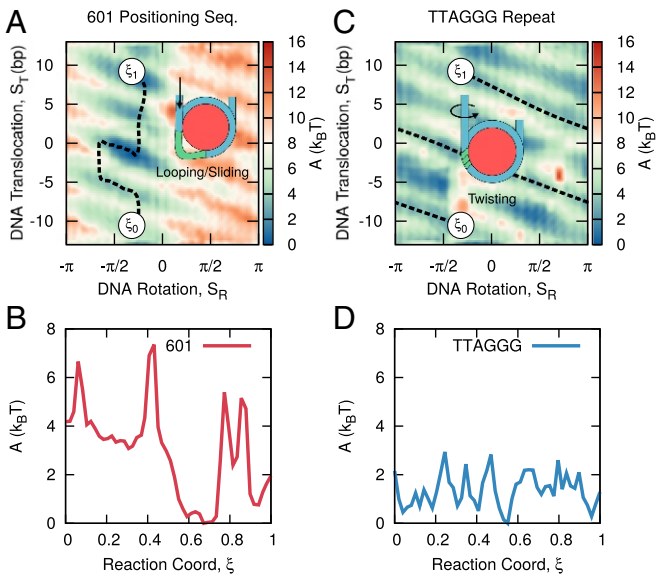

Fig. 4. Minimum free energy path corresponding to 20 bp of DNA translocation for (*A* and *B*) 601 positioning sequence and (*C* and *D*) TTAGGG repeat. (*A*) The 601 sequence exhibits a minimum energy path similar to loop propagation, whereas (*B*) the TTAGGG repeat exhibits a path characteristic of twist diffusion. (*C*) The corresponding energy barriers along this path are large for the 601 positioning sequence and (*D*) relatively small for the TTAGGG repeat.

nucleosome crystal structures to make assumptions about the locations of contacts between the DNA and histones. Our model makes no such assumptions about the locations of these DNA–histone contacts, and it incorporates them through explicit molecular interactions between the DNA and histone proteins. This feature permits our model to account for the effects of thermal fluctuations and the concomitant configurational sampling when determining the free energies of nucleosomal DNA. Additionally, our model simultaneously explains the free energies of sequence-dependent nucleosome formation (47), nucleosome unwrapping (48), and in this work, nucleosome repositioning (Fig. S2). Though previous models have examined these phenomena individually, a comprehensive view of the nucleosome requires that a wide range of experimental measurements are matched and explained simultaneously by a single model, as in our approach here.

We should also note that the minimum energy paths observed in Fig. 4 are not strictly equivalent to the "loop diffusion" and "twist defect" mechanisms described in the Introduction. Since the order parameters S_R and S_T were defined to give the global translocational and rotational position of nucleosomal DNA, they do not resolve the detailed energies corresponding to the motion of a single twist or loop defect. Despite this, the sequence-dependent repositioning mechanism in Fig. 4 contains striking similarities to the loop diffusion and twist defect mechanisms proposed elsewhere in the literature.

The minimum free energy paths observed in Fig. 4 can also be used to explain the mean-squared displacement measurements shown in Fig. 2. The single dynamic regime observed for the TTAGGG repeat (Fig. 2*B*) is found to correspond to the uniform corkscrew-like motion of repositioning, which, from the minimum energy path, appears to be the dominant repositioning mode during DNA motion. The two dynamic regimes observed for the 601 sequence (Fig. 2*A*), however, correspond to the two modes of repositioning present in the minimum energy path. We can now interpret the slow repositioning mode ($\sim t^{0.57}$), as corresponding to the formation and propagation of DNA loops, where DNA is translated quickly. The fast repositioning mode ($\sim t^{0.40}$) then corresponds to the small events that proceed the formation of the subsequent DNA loop.

Taken together, the dynamic and thermodynamic (i.e., free energy) evidence indicates that the 601 sequence and the TTAGGG repeat do indeed reposition through different processes, which are reminiscent of the loop propagation or twist defect mechanism, respectively. Note, however, that the 601 sequence and the TTAGGG repeat represent the two extremes of the DNA sequence affinity for the nucleosome; the free energy difference between the two corresponding nucleosomes is ≈ 20 kJ/mol (9, 54). It is therefore unclear whether this result can be generalized to other sequences, especially naturally occurring genomic sequences with intermediate affinities for the nucleosome.

To investigate whether a truly sequence-dependent mechanism of repositioning indeed exists, we generated a small library of sequences across the range of binding free energies for the histone core. Binding free energies are given by $\Delta\Delta G$, with smaller values of $\Delta\Delta G$ corresponding to stronger DNA–histone binding. These sequences range from strongly bound sequences based on the 601 positioning sequence (9, 12) (c1/601, c2, c3) to intermediately bound sequences (57) (TG, TRGC, TG-T) to weakly bound sequences based on other telomeric repeat sequences (TTAGGG in mamals, TGTGTGGG in *Saccharomyces cerevisiae*, TTGGGG in Tetrahymena) (54, 58). For each of these sequences, we simulated 100 independent and unbiased realizations of nucleosome repositioning by first initializing the system at $S_T \approx 0$ and then performing molecular dynamics until $S_T = \pm 10$. The mechanism of repositioning (i.e., looping or twisting) as well as the repositioning time were examined for each realization. By performing this analysis over an ensemble of trajectories, we were able to generate a probabilistic picture of the different repositioning mechanisms and time scales that are dominant for different DNA sequences.

The mechanism of repositioning for our library of DNA sequences is illustrated in Fig. 5*A*. Consistent with our earlier results, it is strongly dependent on the binding free energy, $\Delta\Delta G$, of the DNA sequence. Strongly bound sequences (low $\Delta\Delta G$) are dominated by the loop propagation repositioning mechanism, with a looping probability, $P_{\text{loop}} \approx 1$. In contrast, weakly bound sequences, associated with high values of $\Delta\Delta G$, are dominated by DNA twisting. However, even though twisting dominates in these sequences, $P_{\text{loop}} = 20 - 40\%$, indicating that DNA looping still occurs. Perhaps the most notable feature of Fig. 5*A* is that the transition between these

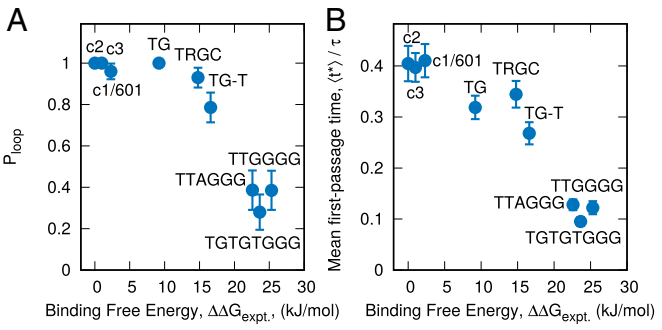

Fig. 5. Sequence-dependent nucleosome repositioning. (*A*) Probability of repositioning through a loop-like mechanism, P_{loop}, for different DNA sequences. Stronger binding sequences (low $\Delta\Delta G$) exhibit looping, whereas weaker binding sequences exhibit a twisting mechanism. (*B*) The mean first-passage time, $\langle \tau^* \rangle$, of a 10-bp DNA translocation for a different DNA sequences. Despite the different repositioning mechanisms exhibited by these different sequences, $\langle \tau^* \rangle$ displays a simple, near-linear decay with $\Delta\Delta G$.

two repositioning mechanisms is smooth, with intermediately binding sequences exhibiting both the looping and twisting mechanism. Importantly, this transition occurs over the range $\Delta\Delta G = 10 - 20$ kJ/mol, which corresponds to almost all naturally occurring DNA sequences. In naturally occurring chromatin, our results therefore indicate that both twisting and looping mechanisms are present simultaneously.

We are not aware of other evidence suggesting that both mechanisms exist simultaneously for naturally occurring DNA sequences, and if correct, this result has important implications for the study of chromatin. Experimentally, it suggests that single-molecule experiments that rely on the 601 sequences may indeed be missing important physics that characterizes naturally occurring DNA sequences. More broadly, however, the simultaneous presence of both looping and twisting could represent an important mechanism for the coupling of DNA sequence and chromatin dynamics, with different combinations influencing the 3D structure of chromatin in both complementary and competing ways.

The time scale of repositioning as a function of DNA sequence can also be quantified from this ensemble of trajectories(Fig. 5*B*) by the mean first-passage time, $\langle \tau^* \rangle$, corresponding to a transition from $S_T \approx 0$ to $S_T = \pm 10$. That time scale is found to be approximately linear with respect to $\Delta\Delta G$, even though the repositioning mechanism changes significantly over this range; $\langle \tau^* \rangle$ appears to be largely independent of mechanism and is instead a simple function of the binding strength of the DNA to the histone surface.

A comment on the sheer magnitude of information contained within this ensemble of trajectories is in order. It consists of 100 independent simulations for 9 different DNA sequences, each of which encompasses at least 10 bp of DNA translocation. These trajectories consist of 350τ of cumulative simulation time, whose spatial, near-Angstrom level resolution approaches that of fully atomistic representations. These trajectories are provided through our website (see *Materials and Methods*), and their analysis might yield additional insights into the structure and dynamics of the nucleosome with a spatial and temporal resolution that complements that of single-molecule experiments. In the section that follows, we further analyze this ensemble of trajectories to highlight additional details of pertaining to the loop propagation and twist diffusion repositioning events.

DNA Loops Are Distributed Unevenly on the Histone Surface. Most theories suggest that, during loop propagation, DNA loops are first introduced in the outer wrap of the nucleosome, where the DNA is known to transiently disassociate from the histone core (59). Such theories suggest that these DNA loops then propagate deeper into the nucleosome, through the dyad, and then exit the nucleosome on the other side, thereby leading to a net DNA translocation (26). In this view, one might expect DNA loops to be, on average, evenly distributed across the histone surface, particularly since loops located at different regions in the nucleosome have comparable energies. However, by analyzing our ensemble of trajectories, we demonstrate that this traditional view of loop propagation is incomplete and that DNA loops are distributed unevenly on the histone surface.

To quantify the position of loops in these trajectories, we first define a variable, θ, that measures the location of loops on the histone surface (Fig. 6*A*; *Materials and Methods*). We then construct a histogram of the loop locations extracted from our trajectories, as shown in Fig. 6*B*. In contrast to the predictions of the loop diffusion model, we observe that loops are distributed in a highly uneven manner along the histone surface, with different regions on the histone either enhancing or inhibiting the formation of loops. Loops are rarely found at the histone dyad ($\theta = 0$) and are strongly favored at $\theta \approx \pm\pi/2$, a location ± 20 bases from the dyad frequently referred to as SHL$\pm$2. These

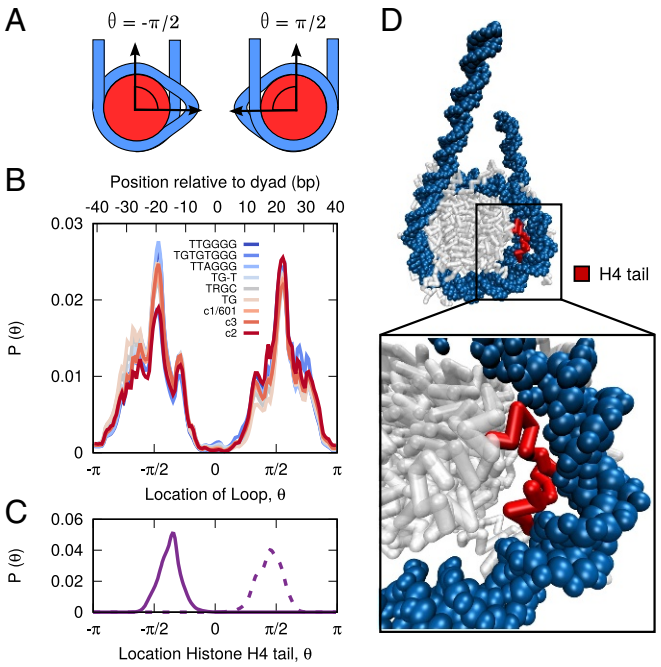

Fig. 6. Distribution of DNA loops on histone surface. (*A*) Diagram describing the relationship between θ and loops on the histone surface. (*B*) Distribution of loop positions for different DNA sequences. Loops are distributed unevenly on the histone surface, in a manner largely independent of DNA sequence. (*C*) Location of histone H4 tail. The H4 tail is colocalized with the location of DNA loops. (*D*) Molecular snapshot highlighting the role of the H4 tail in stabilizing DNA loops at $\theta \approx \pm\pi/2$, a position ± 20 bp from the histone dyad.

loops are observed to be ≈ 10 bp in size, and their locations are found to be insensitive to DNA sequence, with each of our nine DNA sequences exhibiting a similar histogram. This result was unexpected, since these different DNA sequences were observed to reposition by dramatically different mechanisms (cf Fig. 5*A*). This lead us to propose that the location of loops is dictated by a feature of the histone and not the DNA sequence.

One explanation for the observed distribution of loops could be the variable strength of different DNA–histone binding sites to DNA. Single-molecule experiments (40) have shown that certain regions of nucleosomal DNA are more strongly bound to the histone proteins than others (40). For the 601 sequence, these experiments identified three regions of strong DNA–histone binding, with the strongest binding observed at the histone dyad and two other regions at ± 40 bp from the dyad. These three regions correspond exactly with regions in our calculations where looping is absent (Fig. 6*B*), implying that regions of strong DNA–histone binding suppress the formation of DNA loops. Additionally, the lack of sequence dependence in our results suggests that the strong binding regions observed in ref. 40 might be a general feature of nucleosome structure and not simply a specific feature of the 601 sequence used in the experiments.

The observed distribution of loops could also be related to the ability of histone tails to enhance the formation of loops at certain regions in the nucleosome. The location in the nucleosome where we observe the highest probability of loop formation ($\theta = \pm\pi/2$, SHL$\pm$2) is also associated with the protrusion of the tail of histone H4 from the histone surface (1, 51). Our molecular trajectories show that the tail of histone H4 is strongly colocalized with the DNA loops, with the highest probability of being located at $\theta \approx \pm\pi/2$ (Fig. 6*C*). Beyond mere colocalization, however, our simulations suggest that the H4 tails are in fact critical for stabilizing the DNA loops present at $\theta = \pm\pi/2$. When

molecular configurations that exhibit loops are visualized, the H4 tail is found to be bent backward toward the histone core and to be in contact with the in-facing side of the looped DNA (Fig. 6*D*). By adopting that orientation, the H4 tail stabilizes the formation of a loop by restoring the DNA–histone contacts that were originally disrupted. Through this process, the energy of loop formation at SHL±2 is greatly reduced (and hence the high probability of finding loops at that position).

The H4 histone tail, and the DNA region at SHL±2, is rapidly becoming a nexus for nucleosome repositioning phenomena. For many years, SHL±2 has been known to exhibit increased structural variability in crystal structures (1, 60, 61) and is now associated with weak DNA–histone contacts (40). Furthermore, SHL±2 is the site of DNA translocation inside the nucleosome (62) and is a known site where SWI/SNF and ISWI remodelers associate (63). Additionally, the H4 tail is now considered an integral component of chromatin remodeling, with H4 both facilitating remodeling at SHL±2 (42) and playing a role in the mechanism by which ACF senses linker DNA length (41). Our results add to this chorus of results by identifying SHL±2 as a location of prominent DNA looping, through a mechanism that is dependent on the H4 tail. Building on these previous studies, our results suggest a potential for cross-talk between the mechanism of repositioning, the H4 tail, and chromatin remodeler action.

The uneven distribution of loops also has important implications for the mechanism of DNA repositioning. Our observation that loops are rarely found at certain locations along the histone (i.e., $\theta = 0, \pm\pi$) suggests that DNA loops do not slowly propagate across the histone surface as in the loop propagation model but instead rapidly localize to SHL±2. In this view, the H4 tail at SHL±2 could function as a step-wise molecular ratchet, where a DNA loop is (1) pulled into the nucleosome and stored at SHL±2, (2) pulled across the dyad by the H4 tail on the other side at SHL∓2), and finally (3) released from the nucleosome, thus leading to net nucleosome motion. Our simulations present evidence for this type of H4 tail-mediated repositioning event.

Nucleosome Mobility Can Be Enhanced Through an Applied Torque. We now turn our attention to twist-diffusion repositioning. Since the results above indicate that DNA translocation and rotation are coupled in a corkscrew-like motion, chromatin remodelers could be expected to reposition nucleosomes through a "drill"-like mechanism. That is, a chromatin remodeler that applies a purely rotational torque to the DNA molecule could in fact facilitate translocational nucleosomal motion. It is also of interest to determine how DNA sequence, especially a sequence's preference for looping or twisting, might influence this motion.

To examine this possibility, we applied a rotational torque to the DNA molecule (see *Materials and Methods*) and observed that a DNA can indeed be repositioned through a drill-like mechanism (Movie S1). This effect was then quantified by applying a range of rotational torques to the DNA molecule and measuring the effect on repositioning dynamics. The results are shown in Fig. 7. For both the TTAGGG repeat and the 601 sequences, repositioning times (as measured by $\langle\tau^*\rangle$) are significantly reduced when torque is applied. Small torques, of ≈ 8 pN nm, are sufficient to reduce the repositioning rate by more than 50%. For larger torques, the decay times follow a simple exponential form (dotted lines). It is of interest to note that when these results are normalized by the value corresponding to zero torque, τ_0^*, both sequences collapse onto a single line (Fig. 7*B*). Since the TTAGGG repeat repositions through twisting, we had originally expected that its motion would be significantly enhanced when torque is applied. However, since the 601 sequence repositions through looping, an applied torque was not expected to have a large effect on the rate of repositioning. The fact that both sequences are affected equally by applied torque suggests that chromatin remodelers might in fact be able to overlook or bypass many sequence effects. By simply applying a torque, remodelers can override the sequence-dependent repositioning mechanisms (i.e., looping or twisting) and force DNA to reposition via a corkscrew-like motion. Further, even if the torques applied by these remodelers are small, the dynamics of nucleosomal DNA can be altered considerably.

Conclusion

In this work, we present a coarse-grained molecular model that can simultaneously reproduce the free energies of sequence-dependent nucleosome formation, nucleosome unwrapping, and nucleosome repositioning. With this model, we provide in silico evidence that both the loop propagation and twist diffusion mechanisms can occur within the nucleosome. Our results indicate that such mechanisms depend on DNA sequence and that the binding free energy of a given DNA sequence is an excellent predictor of which is the dominant repositioning mechanism. The free energy is also correlated with the characteristic time scale corresponding to a particular repositioning event. It is found that for most naturally occurring DNA sequences, which exhibit a moderate binding energy, both looping and twisting repositioning mechanisms coexist. Importantly, a number of previously unknown features have been identified within the nucleosome as it repositions. These include an asymmetric distribution of DNA loops, a strong influence of the H4 tail in their formation, and the dominating effect of torques in the mobilization of nucleosomes.

One of the central findings of this work is that DNA sequence can lead to a wide range of mechanisms and dynamics of nucleosome positioning. However, we have also shown that certain features are not influenced by DNA sequence, including the locations of DNA loops, and the effect of torque on nucleosome mobility. Building on these findings, it will be important to pursue experimental studies that go beyond the 601 positioning sequence and toward naturally occurring, moderately binding DNA sequences. For cases where it is not possible to perform single-molecule experiments with genomic DNA sequences, the 3SPN-AICG model presented here can serve as a complementary tool to predict situations where sequence dependence might be important.

Materials and Methods

The model of the nucleosome is identical that that used previously (48), where DNA is represented by the 3SPN.2C model (63) and the histone

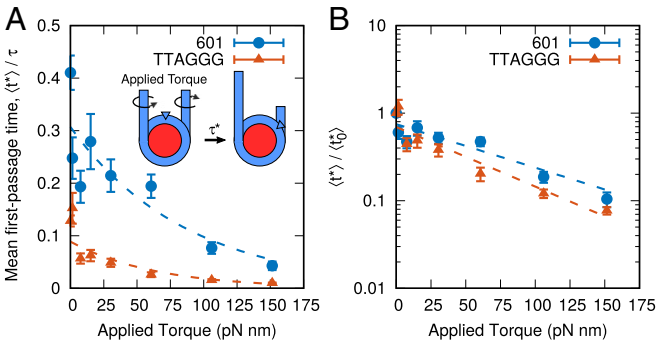

Fig. 7. Applied torque can mobilize nucleosomes. (*A*) Effect of applied torque on mean first-passage times, $\langle\tau^*\rangle$, for 601 sequence and TTAGGG repeat. Small amounts of torque dramatically decrease $\langle\tau^*\rangle$ with a dependence well fit by an exponential decay (dotted lines). (*B*) Upon renormalizing by the zero-torque value $\langle\tau_0^*\rangle$, both sequences exhibit a similar dependence on applied torque. This suggests a mechanism by which chromatin remodelers override sequence-positioning preferences.

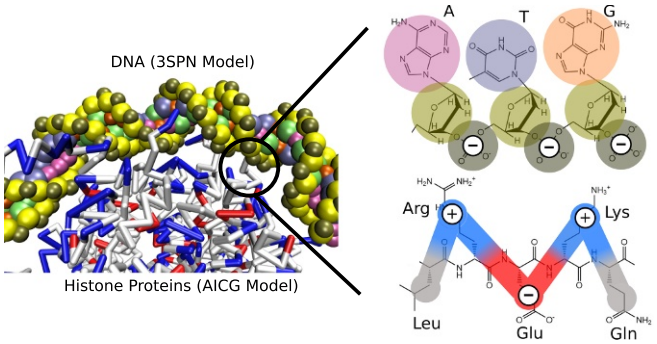

Fig. 8. Description of coarse-grained nucleosome model. DNA is represented by the 3SPN.2C model (63), which represents each nucleotide by three sites located at the center-of-mass of the phosphate (brown), sugar (yellow), and base (pink, purple, orange, green). The histone proteins are represented by the AICG model (50), where each amino acid is represented by a single site, located at the center-of-mass of the amino acid side chain. Interactions between the DNA and histones are represented by Coulombic interactions at the level of Debye–Hückel theory. Phosphate sites of DNA are given a charge of −1, whereas protein sites are given a charge of −1, 0, +1 (red, white, and blue, respectively) corresponding to the charge of that amino acid under physiological pH. Nucleosome configurations and DNA–histone contacts in the 3SPN-AICG model arise naturally from a balance between these Coulombic interactions, with no bias toward the observed nucleosome crystal structure (51).

proteins by the AICG model (50) (Fig. 8). The 3SPN.2C model is the latest version of the 3SPN model (64–66), where DNA is represented by three force sites, located at the center of mass of the phosphate, sugar, and base. The 3SPN.2 model has been parameterized to reproduce the melting of double-stranded to single-stranded DNA and correctly predicts the effects of DNA sequence and salt. Additionally, 3SPN.2C can reproduce the sequence-dependent curvature and sequence-dependent flexibility of DNA (63). The AICG used for the histone proteins represents each amino acid by a single site located at the center of mass of the side chain (50). Interactions between 3SPN.2C and AICG models are represented only by electrostatic and excluded volume effects as described previously (48). Electrostatic forces are introduced at the level of Debye–Hückel theory. All simulations were performed in the canonical ensemble using a Langevin thermostat and 150 mM ionic strength. The coarse-grained topology of the 3SPN-AICG model permits the simulation of very long time scales characteristic of nucleosome repositioning.

A Langevin thermostat represents the solvent implicitly, and the absolute magnitudes of the time scales predicted by this model are less informative than the relative ones. As a consequence, it is conventional to normalize the times in an implicit solvent simulation relative to some other time scale of interest, τ. In this work, we define τ as the characteristic time scale of spontaneous nucleosomal DNA unwrapping as measured experimentally (59, 67). The details of the definition of τ and how it was computed are given in Fig. S3.

We also note that the use of Debye–Hückel theory to represent the DNA–histone interactions in the 3SPN-AICG model neglects multivalent ions and water-mediated hydrogen bonds that have been suggested to mediate DNA–histone contacts (51). For this reason, the 3SPN-AICG model can only capture this complex electrostatic environment to the first order and may smooth out certain aspects of the nucleosome repositioning energy surface. Higher order approximations are available for the 3SPN model (68) but significantly limit the length and time scales that can be examined.

As an initial configuration to examine nucleosome repositioning, we use the 223 bp configuration used previously (48) based on the 1KX5 crystal structure (51) and a proposed structure of exiting nucleosomal DNA (69). We note that this configuration was only used as the initial configuration, and no information from either structure was directly encoded into the nucleosome model. For the c1/601, c2, c3, TG, TRGC, and TG-T sequences used, the 223 bp sequence was generated by taking the defined 147 bp sequence and periodically appending 38 bp on each side. Specifically, if X_i denotes the identity of the ith base pair of a 147 bp DNA sequence, then the resulting 223 bp sequence would be $X_{109}, X_{110}, \ldots, X_{147}, X_1, \ldots, X_{147}, X_1, \ldots, X_{38}$.

Order Parameters. To characterize the movement of DNA around the histone proteins, we explicitly track the rotational and translocational position of the DNA relative to the protein dyad. To define the translocational position of the DNA molecule, we define an order parameter, S_T, defined as:

$$S_T = \frac{1}{\lambda}\left\langle \pm \arccos\left(\frac{\boldsymbol{P}\cdot\boldsymbol{P}_0}{\|\boldsymbol{P}\|\|\boldsymbol{P}_0\|}\right)\right\rangle,$$

where $\boldsymbol{P}$ is a vector from the center of a base step to the center of the protein, $\boldsymbol{P}_0$ is the value of $\boldsymbol{P}$ from the 1KX5 nucleosomal crystal structure (51), and the angle brackets denote an average over base steps at the −15, −5, +5, and +15 positions relative to the dyad. The positive sign is chosen if $(\boldsymbol{P}\times\boldsymbol{P}_0)\cdot\hat{\boldsymbol{f}} \leq 0$ (negative if > 0), where $\hat{\boldsymbol{f}}$ is a vector that points along the center of the nucleosomal DNA superhelix. Therefore, positive S_T corresponds to "forward" translocation of DNA (toward 5′ end), whereas negative S_T corresponds to "backward" motion (toward 3′ end) as shown in Fig. 1*B*. Lastly, λ represents a conversion factor from radians to base pairs of DNA translocation and is defined as $\lambda = 0.08$ rad/bp. λ is calculated by dividing the circumference of the histone proteins (1), $r \approx 42$Å, by 2π and the distance between adjacent base pairs, 3.3Å.

The rotational position of DNA is defined by an order parameter, S_R, given by:

$$S_R = \left\langle \pm \arccos\left(\frac{\boldsymbol{P}\cdot\boldsymbol{B}}{\|\boldsymbol{P}\|\|\boldsymbol{B}\|}\right)\right\rangle,$$

where $\boldsymbol{B}$ is a vector from the center of a given base step on the sense strand to its complementary base step on the antisense strand and all other terms are the same as defined for S_T. The positive sign is chosen if $(\boldsymbol{P}\times\boldsymbol{B})\cdot\boldsymbol{D} \leq 0$ (negative if > 0), where $\boldsymbol{D}$ is a vector in the 5′ to 3′ direction along the sense strand. Notably, when $S_R = -\pi/2$, the minor groove is oriented toward the protein core, and when $S_R = \pi/2$, it is oriented away from it (see Fig. 1*B*). The order parameter S_R has been used previously to quantify the positioning preferences of different DNA sequences to the histone proteins (47).

To apply a torque to the DNA molecule, we simply applied a constant force along the S_R order parameter. Since S_R is defined as the average over four different DNA base steps, this applied torque is divided evenly among the base steps at −15, −5, +5, and +15 positions relative to the central base pair. By definition this force results in rotation of the DNA molecule around the histone proteins and can be converted to a torque using the diameter of the DNA.

Free Energy Methods. Two-dimensional free energy surfaces (Fig. 3) were obtained along S_T and S_R using 2D umbrella sampling and WHAM (70, 71). Errors in these free energies were assessed by reconstructing three independent free energy surfaces and computing the SD. The majority of the free energy is found to be accurate to within $\pm 1 k_B T$. The largest errors ($3k_BT$) correspond to large values of the free energy ($12k_BT$), leading to an acceptable relative error of $\approx 25\%$ (Fig. S4).

To determine probable reaction paths along these free energy surface (Fig. 4), we use the string method (72, 73), as implemented in the SSAGES package (miccomcodes.com). Rather than randomly guessing initial paths, we used a hybrid approach where the finite-temperature string method (73) was used at a high temperature to generate an ensemble of possible paths, which were then energy minimized using the zero-temperature string method (72). This hybrid approach was found to yield a much wider range of possible paths than the zero-temperature or finite-temperature string methods alone. Since the free energy surfaces used in the string method were obtained from a complex molecular system, they are expectedly rough, and many probable paths contained similar energies. Accordingly, the paths reported in Fig. 4 represent a representative minimum energy path chosen from an ensemble of possible paths predicted using the finite-temperature string.

Quantification of DNA Loop Locations. To quantify the locations of DNA loops on the histone surface, it was necessary to first define an orthonormal basis to represent the rotational position of the histone proteins. This orthonormal basis consisted of three vectors, $\hat{\boldsymbol{f}}$, $\hat{\boldsymbol{v}}$, $\hat{\boldsymbol{u}}$, where $\hat{\boldsymbol{u}}$ points from the center of mass of the histone to the dyad, $\hat{\boldsymbol{f}}$ (as defined above) points along the nucleosomal DNA superhelix, and $\hat{\boldsymbol{u}}\times\hat{\boldsymbol{f}} = \hat{\boldsymbol{v}}$. After this basis was defined, the center of the DNA helical axis was calculated for each base pair using the Kahn method (74) as described previously (66). A vector was then constructed from the center of mass of the protein to the helical axis site of the ith base pair with orientation $\hat{\boldsymbol{w}}_i'$ and magnitude $\ell_i = |\boldsymbol{w}_i'|$. The vector rejection of $\hat{\boldsymbol{w}}_i'$ and $\hat{\boldsymbol{f}}$, $\hat{\boldsymbol{w}}_i = \hat{\boldsymbol{w}}_i' - \hat{\boldsymbol{w}}_i'\cdot\hat{\boldsymbol{f}}$ is computed, where $\hat{\boldsymbol{w}}_i$ is the projection of $\hat{\boldsymbol{w}}_i'$ onto the plane containing $\hat{\boldsymbol{v}}$ and $\hat{\boldsymbol{u}}$, perpendicular to $\hat{\boldsymbol{f}}$. The location of each base pair is then defined by the angle θ_i, where $\theta_i = \arccos(\hat{\boldsymbol{w}}_i\cdot\hat{\boldsymbol{u}})$

(see Fig. 6*A*). This process characterizes each base pair of DNA by two values: ℓ_i the distance of that base pair to the histone center of mass, and θ_i the location of the base pair relative to the dyad. Note that since DNA is dynamically repositioning, θ_i and ℓ_i are not constant during a simulation. We therefore found it more useful to compute the average distances as a function of theta, $\langle \ell(\theta) \rangle$.

Next, to define whether a loop was present, we first construct a normalization curve, $\langle \bar{\ell}(\theta) \rangle$, which characterizes the average distance of a base pair at a given θ in the absence of loops. To ensure that loops did not form in this calculation, $\langle \bar{\ell}(\theta) \rangle$ was calculated at a very low salt for a strongly bound DNA sequence. $\langle \bar{\ell}(\theta) \rangle$ is then used to normalize $\langle \ell(\theta) \rangle$ to obtain $\Delta\ell(\theta)$, where $\Delta\ell(\theta) = \langle \ell(\theta) \rangle - \langle \bar{\ell}(\theta) \rangle$. Using this metric, in the absence of DNA loops, when DNA is in close contact with the histone surface, $\Delta\ell \approx 0$ for all θ. However, when DNA loops form and DNA–histone contacts are disrupted, $\Delta\ell > 0$. We define loops as base pairs where $\Delta\ell(\theta) > \delta$, where δ is some predefined threshold. We choose $\delta = 8$Å, corresponding to the Debye length at 150 mM, the length at which the DNA–histone attraction has significantly decayed. Small variations in the value of δ had little effect on our results.

All simulation codes and results presented in this work are freely available from our website (datahub.uchicago.edu and miccomcodes.org).

ACKNOWLEDGMENTS. The authors thank Andrés Córdoba, Gordon S. Freeman, and Daniel M. Hinckley for helpful discussions; Joshua Moller for the codes used to calculate the histone orientation; and the Midway computing cluster at the University of Chicago and the University of Wisconsin–Madison Center for High Throughput Computing for computational resources. The study of charge-driven complexation in polymeric materials, including DNA, is supported by the US Department of Commerce Award 70NANB14H012, National Institute of Standards and Technology, as part of the Center for Hierarchical Materials Design (CHiMaD). The development of a fast GPU code for molecular dynamics simulations (DASH) and the development of advanced sampling codes (SAGESS) required for free energy calculations was supported by the US Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division through the Midwest Integrated Center for Computational Materials (MICCoM and miccomcodes.org). D.C.S. acknowledges support from the National Human Genome Research Institute (NIH R01-HG-000225).

1. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome resolution core particle at 2.8 A. *Nature* 389:251–260.
2. Knezetic JA, Luse DS (1986) The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. *Cell* 45:95–104.
3. Lorch Y, LaPointe JW, Kornberg RD (1987) Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell* 49: 203–210.
4. Archer TK, Cordingley MG, Wolford RG, Hager GL (1991) Transcription factor access is mediated by accurately positioned nucleosomes on the mouse mammary tumor virus promoter. *Mol Cell Biol* 11:688–698.
5. Valouev A, et al. (2011) Determinants of nucleosome organization in primary human cells. *Nature* 474:516–520.
6. Schones DE, et al. (2008) Dynamic regulation of nucleosome positioning in the human genome. *Cell* 132:887–898.
7. Hendrich B, Bickmore W (2001) Human diseases with underlying defects in chromatin structure and modification. *Hum Mol Genet* 10:2233–2242.
8. Bhaumik SR, Smith E, Shilatifard A (2007) Covalent modifications of histones during development and disease pathogenesis. *Nat Struct Mol Biol* 14:1008–1016.
9. Lowary PT, Widom J (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol* 276: 19–42.
10. Thastrom A, et al. (1999) Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J Mol Biol* 288:213–229.
11. Struhl K, Segal E (2013) Determinants of nucleosome positioning. *Nat Struct Mol Biol* 20:267–273.
12. Segal E, et al. (2006) A genomic code for nucleosome positioning. *Nature* 442:772–778.
13. Kaplan N, et al. (2009) The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* 458:362–366.
14. Zhang Y, et al. (2009) Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. *Nat Struct Mol Biol* 16:847–852.
15. Stein A, Takasuka TE, Collings CK (2010) Are nucleosome positions in vivo primarily determined by histone-DNA sequence preferences? *Nucleic Acids Res* 38:709–719.
16. Mueller-Planitz F, Klinker H, Becker PB (2013) Nucleosome sliding mechanisms: New twists in a looped history. *Nat Struct Mol Biol* 20:1026–1032.
17. Wippo CJ, et al. (2011) The RSC chromatin remodelling enzyme has a unique role in directing the accurate positioning of nucleosomes. *EMBO J* 30:1277–1288.
18. Fazzio TG, Tsukiyama T (2003) Chromatin remodeling in vivo: Evidence for a nucleosome sliding mechanism. *Mol Cell* 12:1333–1340.
19. Whitehouse I, Rando OJ, Delrow J, Tsukiyama T (2007) Chromatin remodelling at promoters suppresses antisense transcription. *Nature* 450:1031–1035.
20. Ito T, Bulger M, Pazin MJ, Kobayashi R, Kadonaga JT (1997) ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor. *Cell* 90:145–155.
21. Yang JG, Madrid TS, Sevastopoulos E, Narlikar GJ (2006) The chromatin-remodeling enzyme ACF is an ATP-dependent DNA length sensor that regulates nucleosome spacing. *Nat Struct Mol Biol* 13:1078–1083.
22. Racki LR, et al. (2009) The chromatin remodeller ACF acts as a dimeric motor to space nucleosomes. *Nature* 462:1016–1021.
23. Pennings S, Meersseman G, Bradbury EM (1991) Mobility of positioned nucleosomes on 5 S rDNA. *J Mol Biol* 220:101–110.
24. Meersseman G, Pennings S, Bradbury EM (1992) Mobile nucleosomes–A general behavior. *EMBO J* 11:2951–2959.
25. Flaus A, Richmond TJ (1998) Positioning and stability of nucleosomes on MMTV 3'LTR sequences. *J Mol Biol* 275:427–441.
26. Schiessel H, Widom J, Bruinsma R, Gelbart W (2001) Polymer reptation and nucleosome repositioning. *Phys Rev Lett* 86:4414–4417.
27. Kulić IM, Schiessel H (2003) Nucleosome repositioning via loop formation. *Biophys J* 84:3197–3211.
28. Lorch Y, Davis B, Kornberg RD (2005) Chromatin remodeling by DNA bending, not twisting. *Proc Nat Acad Sci USA* 102:1329–1332.
29. Ranjith P, Yan J, Marko JF (2007) Nucleosome hopping and sliding kinetics determined from dynamics of single chromatin fibers in Xenopus egg extracts. *Proc Nat Acad Sci USA* 104:13649–13654.
30. Strohner R, et al. (2005) A 'loop recapture' mechanism for ACF-dependent nucleosome remodeling. *Nat Struct Mol Biol* 12:683–690.
31. Pasi M, Lavery R (2016) Structure and dynamics of DNA loops on nucleosomes studied with atomistic, microsecond-scale molecular dynamics. *Nucleic Acids Res* 44:5450–5456.
32. Kulić I, Schiessel H (2003) Chromatin dynamics: Nucleosomes go mobile through twist defects. *Phys Rev Lett* 91:148103.
33. Richmond TJ, Davey CA (2003) The structure of DNA in the nucleosome core. *Nature* 423:145–150.
34. Suto RK, et al. (2003) Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. *J Mol Biol* 326:371–380.
35. Gottesfeld JM, Belitsky JM, Melander C, Dervan PB, Luger K (2002) Blocking transcription through a nucleosome with synthetic DNA ligands. *J Mol Biol* 321:249–263.
36. Blosser TR, Yang JG, Stone MD, Narlikar GJ, Zhuang X (2009) Dynamics of nucleosome remodelling by individual ACF complexes. *Nature* 462:1022–1027.
37. Deindl S, et al. (2013) ISWI remodelers slide nucleosomes with coordinated multi-base-pair entry steps and single-base-pair exit steps. *Cell* 152:442–452.
38. Eslami-Mossallam B, Schiessel H, van Noort J (2015) Nucleosome dynamics: Sequence matters. *Adv Colloid Interface Sci* 232:101–113.
39. Pisano S, et al. (2007) Telomeric nucleosomes are intrinsically mobile. *J Mol Biol* 369:1153–1162.
40. Hall MA, et al. (2009) High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. *Nat Struct Mol Biol* 16:124–129.
41. Hwang WL, Deindl S, Harada BT, Zhuang X (2014) Histone H4 tail mediates allosteric regulation of nucleosome remodelling by linker DNA. *Nature* 512:213–217.
42. Mueller-Planitz F, Klinker H, Ludwigsen J, Becker PB (2013) The ATPase domain of ISWI is an autonomous nucleosome remodeling machine. *Nat Struct Mol Biol* 20:82–89.
43. Klinker H, et al. (2014) ISWI remodelling of physiological chromatin fibres acetylated at lysine 16 of histone H4. *PLoS One* 9:e88411.
44. Stockdale C, Flaus A, Ferreira H, Owen-Hughes T (2006) Analysis of nucleosome repositioning by yeast ISWI and Chd1 chromatin remodeling complexes. *J Biol Chem* 281:16279–16288.
45. Tan S, Davey CA (2011) Nucleosome structural studies. *Curr Opin Struc Biol* 21:128–136.
46. Fathizadeh A, Berdy Besya A, Reza Ejtehadi M, Schiessel H (2013) Rigid-body molecular dynamics of DNA inside a nucleosome. *Eur Phys J E Soft Matter* 36:21.
47. Freeman GS, Lequieu JP, Hinckley DM, Whitmer JK, de Pablo JJ (2014) DNA shape dominates sequence affinity in nucleosome formation. *Phys Rev Lett* 113:168101.
48. Lequieu J, Córdoba A, Schwartz DC, de Pablo JJ (2016) Tension-dependent free energies of nucleosome unwrapping. *ACS Cent Sci* 2:660–666.
49. Freeman GS, Hinckley DM, Lequieu JP, Whitmer JK, de Pablo JJ (2014) Coarse-grained modeling of DNA curvature. *J Chem Phys* 141:165103.
50. Li W, Wolynes PG, Takada S (2011) Frustration, specific sequence dependence, and nonlinearity in large-amplitude fluctuations of allosteric proteins. *Proc Nat Acad Sci USA* 108:3504–3509.
51. Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. *J Mol Biol* 319:1097–1113.
52. Moyzis RK, et al. (1988) A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. *Proc Nat Acad Sci USA* 85:6622–6626.
53. Morin GB (1989) The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell* 59:521–529.
54. Filesi I, Cacchione S, De Santis P, Rossetti L, Savino M (2000) The main role of the sequence-dependent DNA elasticity in determining the free energy of nucleosome formation on telomeric DNAs. *Biophys Chem* 83:223–237.
55. Morozov AV, et al. (2009) Using DNA mechanics to predict in vitro nucleosome positions and formation energies. *Nucleic Acids Res* 37:4707–4722.

56. Tolstorukov MY, Colasanti AV, McCandlish DM, Olson WK, Zhurkin VB (2007) A novel roll-and-slide mechanism of DNA folding in chromatin: Implications for nucleosome positioning. *J Mol Biol* 371:725–738.
57. Shrader TE, Crothers DM (1989) Artificial nucleosome positioning sequences. *Proc Nat Acad Sci USA* 86:7481–7482.
58. Cacchione S, Cerone MA, Savino M (1997) In vitro low propensity to form nucleosomes of four telomeric sequences. *FEBS Lett* 400:37–41.
59. Li G, Levitus M, Bustamante C, Widom J (2005) Rapid spontaneous accessibility of nucleosomal DNA. *Nat Struct Mol Biol* 12:46–53.
60. Edayathumangalam RS, Weyermann P, Dervan PB, Gottesfeld JM, Luger K (2005) Nucleosomes in solution exist as a mixture of twist-defect states. *J Mol Biol* 345:103–114.
61. Bowman GD (2010) Mechanisms of ATP-dependent nucleosome sliding. *Curr Opin Struc Biol* 20:73–81.
62. Zofall M, Persinger J, Kassabov SR, Bartholomew B (2006) Chromatin remodeling by ISW2 and SWI/SNF requires DNA translocation inside the nucleosome. *Nat Struct Mol Biol* 13:339–346.
63. Dechassa ML, et al. (2012) Disparity in the DNA translocase domains of SWI/SNF and ISW2. *Nucleic Acids Res* 40:4412–4421.
64. Knotts TA, Rathore N, Schwartz DC, de Pablo JJ (2007) A coarse grain model for DNA. *J Chem Phys*126:084901.
65. Sambriski EJ, Schwartz DC, de Pablo JJ (2009) A mesoscale model of DNA and its renaturation. *Biophys J* 96:1675–1690.
66. Hinckley DM, Freeman GS, Whitmer JK, de Pablo JJ (2013) An experimentally-informed coarse-grained 3-site-per-nucleotide model of DNA: Structure, thermodynamics, and dynamics of hybridization. *J Chem Phys* 139:144903.
67. Tims HS, Gurunathan K, Levitus M, Widom J (2011) Dynamics of nucleosome invasion by DNA binding proteins. *J Mol Biol* 411:430–448.
68. Hinckley DM, De Pablo JJ (2015) Coarse-grained ions for nucleic acid modeling. *J Chem Theor Comput* 11:5436–5446.
69. Hussain S, et al. (2010) Single-base resolution mapping of H1-nucleosome interactions and 3D organization of the nucleosome. *Proc Nat Acad Sci USA* 107:9620–9625.
70. Kästner J (2011) Umbrella sampling. *Wiley Interdiscip Rev Comput Mol Sci* 1:932–942.
71. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA (1995) Multidimensional free-energy calculations using the weighted histogram analysis method. *J Comput Chem* 16:1339–1350.
72. E W, Ren W, Vanden-Eijnden E (2007) Simplified and improved string method for computing the minimum energy paths in barrier-crossing events. *J Chem Phys* 126:164103.
73. Vanden-Eijnden E, Venturoli M (2009) Revisiting the finite temperature string method for the calculation of reaction tubes and free energies. *J Chem Phys* 130:194103.
74. Kahn PC (1989) Defining the axis of a helix. *Comput Chem* 13:185–189.
75. Mihardja S, Spakowitz AJ, Zhang Y, Bustamante C (2006) Effect of force on mononucleosomal dynamics. *Proc Nat Acad Sci USA* 103:15871–15876.
76. Flaus A, Luger K, Tan S, Richmond TJ (1996) Mapping nucleosome position at single base-pair resolution by using site-directed hydroxyl radicals. *Proc Nat Acad Sci USA* 93:1370–1375.