Research Article

Surprising Twists in Nucleosomal DNA with Implication for Higher-order Folding

Stefjord Todolli, Robert T. Young, Abigail S. Watkins, Antonio Bu Sha, John Yager and Wilma K. Olson *

Department of Chemistry & Chemical Biology, *Center for Quantitative Biology, Rutgers, the State University of New Jersey, Piscataway, NJ 08854, USA*

Correspondence to Wilma K. Olson: *wilma.olson@rutgers.edu (W.K. Olson)*
https://doi.org/10.1016/j.jmb.2021.167121
Edited by Anna Panchenko

Abstract

While nucleosomes are dynamic entities that must undergo structural deformations to perform their functions, the general view from available high-resolution structures is a largely static one. Even though numerous examples of twist defects have been documented, the DNA wrapped around the histone core is generally thought to be overtwisted. Analysis of available high-resolution structures from the Protein Data Bank reveals a heterogeneous distribution of twist along the nucleosomal DNA, with clear patterns that are consistent with the literature, and a significant fraction of structures that are undertwisted. The subtle differences in nucleosomal DNA folding, which extend beyond twist, have implications for nucleosome disassembly and modeled higher-order structures. Simulations of oligonucleosome arrays built with undertwisted models behave very differently from those constructed from overtwisted models, in terms of compaction and inter-nucleosome contacts, introducing configurational changes equivalent to those associated with 2–3 base-pair changes in nucleosome spacing. Differences in the nucleosomal DNA pathway, which underlie the way that DNA enters and exits the nucleosome, give rise to different nucleosome-decorated minicircles and affect the topological mix of configurational states.

Introduction

The dynamic features of the nucleosome are critical to its function. Nucleosomal DNA needs to peel away from the spool of histone proteins around which it wraps in order for assorted binding factors to access the embedded genetic information (exceptions include pioneer transcription factors, which target sites on the nucleosome exterior[1]). The ease of nucleotide exposure depends upon the DNA pathway, with residues at the ends of the ~150-bp bound fragment much more likely to undergo spontaneous 'breathing' motions, i.e., unwrapping and rewrapping, than those in the middle.[2] Complete unwrapping comes only upon major perturbation of the local environment, e.g., addition of salts that compete with the highly charged histones for access to DNA[3] or imposition of external forces from invading molecular machinery[4] or direct molecular manipulations.[5,6]

Maps of histone-DNA interactions, obtained by mechanically unzipping the DNA of single nucleosomes,[7] find the contacts near the exit and entry DNA to be particularly weak, allowing for the spontaneous peeling of DNA ends consistent with assays of nucleosomal DNA digestion by restriction enzymes[2] and opened configurations detected with fluorescence resonance energy transfer measurements.[8–10] While many of the structures of nucleosomes captured in recent cryogenic electron microscopy (cryo-EM) studies bear close resemblance to high-resolution crystal structures where

DNA is tightly wrapped in a superhelical pathway around the core of histone proteins, there are examples in the cryo-EM literature of open nucleosome structures with bulges at the ends of the bound DNA or with DNA displaced from the histone core.[11] The perturbations in the histone assembly that accompany these deformations seemingly limit the unwrapping to one end of the double helix, providing a structural rationale behind the DNA breathing asymmetry detected with other techniques.[10,12]

The literature also includes observations of slower nucleosomal motions, such as sliding, where the DNA and histone core translocate with respect to one another, forcing DNA-protein contacts to break in one setting and reform in another.[13,14] There is further experimental evidence of a slow opening, or gaping, of the nucleosome under physiological conditions,[15] motions hypothesized to arise from slight distortions (e.g., hydrogen-bond breakages) of the histone core.[16] The proposed separation of DNA gyres introduces a higher-order curvature in the DNA superhelical pathway thought to account for topological changes observed in nucleosome-decorated DNA minicircles.[17] The displacement of the gyres upon gaping is expected to disrupt the extended groove structure of the nucleosome, formed by the side-by-side alignment of DNA segments separated by a turn (~80 bp) of superhelix and used by selected ligands and transcription factors to recognize spatially distant DNA residues.[18,19] The two DNA gyres run roughly parallel to one another in an intact nucleosome, save for a single, ~15 bp stretch of DNA around the central base pair, the so-called dyad. The global motions, estimated from the interatomic contacts in three of the first reported high-resolution structures,[20] suggest that nucleosomal gaping may be coupled to breathing. That is, the nucleosome appears to flex about the normal of the central base pair as the ends of the DNA wrap and unwrap from the histone core. Moreover, single-molecule studies of salt-induced nucleosome disassembly have found DNA dissociation to precede the displacement of proteins.[21] The mode of DNA deformation, however, is unclear, with either gaping or unwrapping consistent with observation.

Subtle differences in the structures of nucleosomes lead to displacements between the DNA ends large enough to change the features of nucleosome-decorated DNA chains. Slight changes in the degree of DNA wrapping around idealized nucleosome models give rise to sharp jumps in the predicted folding of nucleosome-decorated minicircles,[22] with the protein-free DNA converting between an open, contact-free loop and a crossed configuration with sequentially distant residues in close contact.[23] As described in this work, perturbations in the three-dimensional structures of nucleosomes bearing DNA fragments of the same length give rise to other modes of chromatin reorganization. The precise structure of the histone-DNA assembly determines whether a regularly spaced array of nucleosomes adopts a loose, relatively flexible form with the DNA superhelical axes running roughly parallel to the overall chain direction or a stiffer, more tightly packed state with the superhelical axes oriented more nearly perpendicular to the chain direction. Our findings suggest that the twisting of successive base pairs may be responsible, at least in part, for the observed differences in global features.

Examples of nucleosomal twist defects go back to the very first high-resolution structures of the histone-DNA assembly, where the number of base pairs in one half of the complex is one less than that in the other half.[24,25] The shorter DNA twists and stretches over roughly a turn of double helix in order to preserve the stacking of terminal base pairs against those of neighboring nucleosomes in the crystal lattice. The site of twist uptake varies with DNA sequence and in the presence of small ligands, occurring in many early structures at sites ~20 or ~45 bp on either side of the dyad and appearing to fluctuate between these and other sites in solution.[26,27] There is growing experimental evidence that these small distortions of DNA may play a role in the remodeling of nucleosomes,[28–30] e.g., helping to propagate a DNA bulge on the surface of the histone core and thereby to slide the nucleosome onto a new position on DNA.[31]

Increasing interest in how local nucleosome dynamics contribute to larger-scale features of chromatin has stimulated the development of models that incorporate the intrinsic motions of nucleosomes in simulations of longer DNA chains. For example, Norouzi and Zhurkin[32] have devised a novel adhesion potential, based on the unzipping properties of single nucleosomes,[7] to treat the dynamic response of nucleosomal arrays to external forces. Schiessel and associates have used a sequence-dependent elastic potential, based on the arrangements of successive base pairs in observed and simulated DNA structures,[33] to account for the positioning of nucleosomes along DNA[33] and the unwrapped states of nucleosomes bearing different DNA sequences.[34] Other groups have employed even more coarse-grained representations of the nucleosome, e.g., with the motions of the assembly controlled by the interactions of a reduced set of points on protein and DNA,[35] to follow the potential breathing and sliding propensities of nucleosomes.[31,36–37] Here we make use of the large and growing collection of high-resolution nucleosome structures to examine motions of the histone-DNA assembly as a whole and to incorporate this information in studies of longer molecules.

There are now enough well-determined X-ray and cryo-EM structures of nucleosomes and their complexes with other molecules to begin to decipher the overall dynamics of the histone-DNA assembly from the set of available examples rather than describe the dynamics in terms of

simplified physical models. The data gathered here reveal surprising trends in the build-up and distribution of twist along the DNA and the apparent gaping of DNA between the two halves of the nucleosome not found in individual structures. The variation in twist has a marked effect on the arrangement of successive nucleosomes and, in turn, on the configurations of simulated oligonucleosome arrays and cyclic nucleosome-decorated DNA. The gaping is most pronounced near the sites of extreme under- and overtwisting. The collective information provides useful benchmarks for atomic-level studies of nucleosomes and new ideas for exploration of the properties of chromatin.

Results

Nucleosome structural diversity

The collective information found in a large number of high-resolution nucleosome structures offers insights into macromolecular organization and dynamics not evident in the isolated structures.

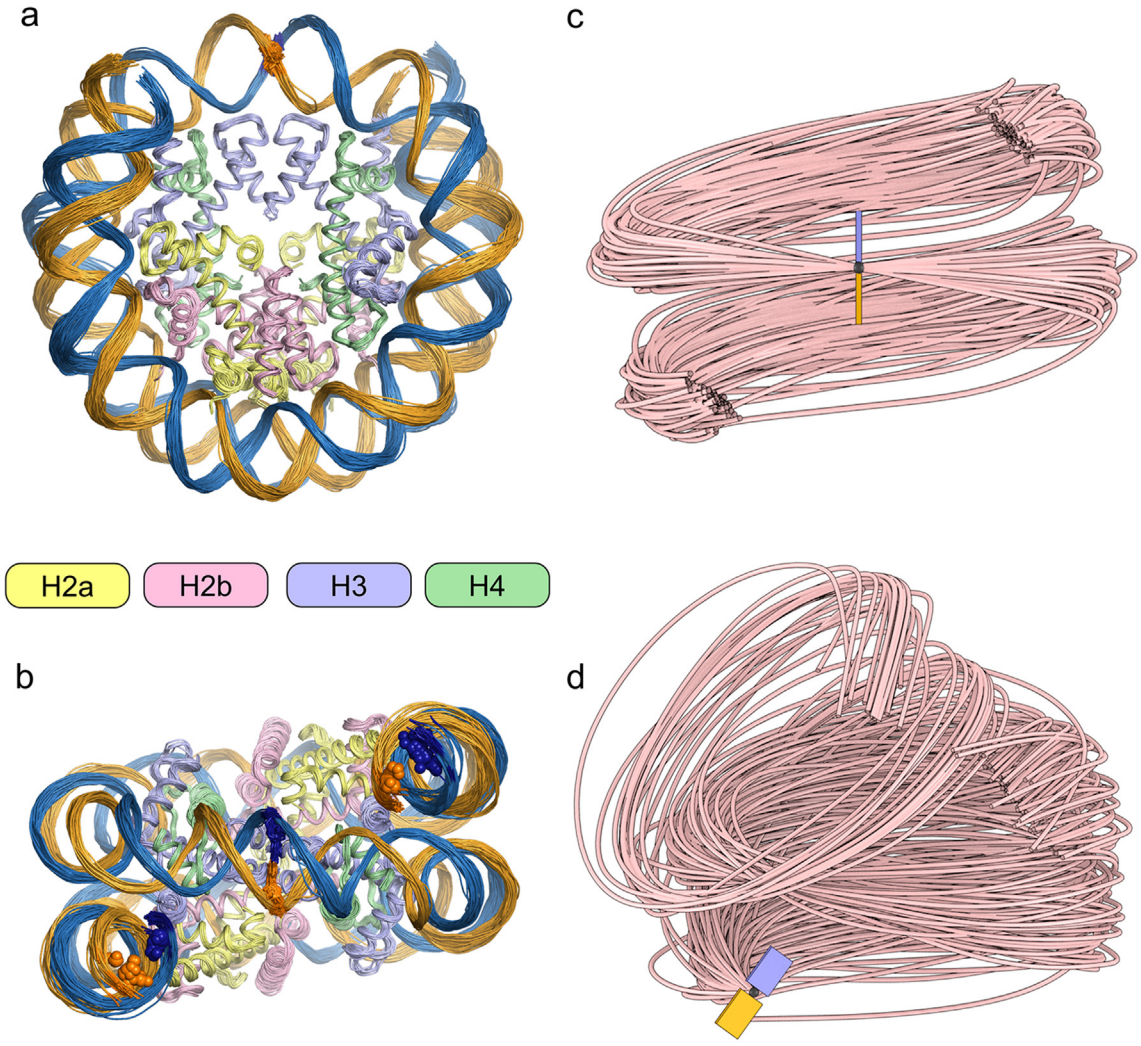

Figure 1. Molecular images illustrating the 'dynamics' of 164 high-resolution nucleosome structures from different perspectives. (a) Conventional view in the direction of the DNA superhelical axis highlighting the symmetry of the assembly and the relative deformability of DNA vs. protein. Structures are superimposed on the core histone regions[83] of the best-resolved structure (pdb file 1kx5[25]) using the PyMOL software,[84] with the DNA and histone backbones depicted by thin lines (DNA in blue/gold and histones H2A, H2B, H3, H4 in yellow, pink, blue, green, respectively) and the dyad base pairs shown as thick lines. Histone tails and DNA residues outside the central 141 bp are not shown; (b) Same depiction of nucleosomal components but viewed down the dyad axis with the C6/C8 atoms of pyrimidine/purine bases at positions ±70 shown as spheres; (c) 'Pinched' view with nucleosomes aligned on the frame of the dyad base pair (blue/gold blocks) and the centers of base pairs connected by smooth B-spline curves; (d) 'Fixed-end' view with nucleosomes aligned on the frame of base pair –70 (blue-gold blocks) and DNA pathways again depicted by B-spline curves.

Here we examine selected features in 164 nucleosome structures of 3.5-Å or better resolution available as of August 2020 in the Protein Data Bank (pdb).[38] The collection includes 146 structures determined using X-ray crystallographic methods and 18 based on cryo-EM techniques. The DNA sequence in the vast majority of structures is a variation of an α-satellite nucleosome-positioning sequence[24,39] with nearly 20% (29) bearing the Widom-601 positioning sequence.[40] The DNA chain length in the different complexes ranges from 145 to 167 bp with nearly 60% (94) of the entries containing 146 bp and almost 40% with 145 or 147 bp (29 and 34 examples, respectively). Most of the crystal structures (136) belong to the orthorhombic space group $P2_12_12_1$, with the terminal base pairs of each nucleosome stacked against the base pairs of two symmetry-related neighbors forming a pseudo-continuous DNA pathway.[41] A total of 39 structures bind additional proteins. See Figure S1 and Tables S1 and S2 for further details.

The composite structures appear quite similar when examined from the conventional top-down perspective, showing the familiar pseudo-twofold symmetry of the nucleosome about the dyad (Figure 1(a)). Here the structures are superimposed on the histone core of the best-resolved core-particle structure (pdb file 1kx5[25]) with thin lines connecting the C^{α} atoms of protein and the P atoms of DNA and thick lines highlighting the base pair on the dyad. This alignment allows for the uniform numbering and comparison of base pairs with respect to the dyad, at base-pair 0, in all considered structures (see Methods). The wider spread of lines on DNA compared to protein draws attention to the well-known enhancement in atomic mobility at the edges vs. the core of individual structures.[24] The relative spread of lines along the DNA similarly highlights the known ease of deforming the ends of the double helix compared to the center.[2,7–10] This viewpoint, however, hides subtleties in the orientation of terminal base pairs, which span a wide range of states that become apparent when the set of structures is viewed looking down the dyad axis (Figure 1(b)). The variety of orientations appears in the semicircular scatter of points depicting the purine C8 and pyrimidine C6 atoms of the terminal pairs (here located at base pairs ±70 with respect to the dyad). A related view of the same structures, superimposed in a common reference frame on the dyad with thin lines passing through the centers of successive base pairs, reveals noticeable displacements in the DNA chain ends (Figure 1(c)). The displacements become even more pronounced when the basepair pathways are superimposed at one end of the chain (here base-pair −70), with the overall structures significantly rotated with respect to one another (Figure 1(d)). Moreover, these structures fall into distinct groupings suggestive of different nucleosome states and motions of potential relevance to chromatin organization.

Nucleosomal DNA under- and overtwisting

The rotations of terminal base pairs and the large-scale reorientation of nucleosomal DNA seen in the collection of structures arise in large part from differences in the twisting of the constituent base pairs. A color-coded heat map of the build-up in twist at successive base pairs in the set of high-resolution structures reveals distinct patterns of DNA under- and overtwisting within different subsets of structures (Figure 2). Here we plot the uptake of the so-called twist of supercoiling[42,43] at every base-pair step along with the total uptake on each nucleosome relative to the average values for all nucleosomes in the dataset—an average accumulated twist of $4835_{\pm 28}^{\circ}$ over the central 140-bp steps, corresponding to an average twist per base-pair step of $\sim 34.5_{\pm 0.2}^{\circ}$ (see Table S2 for numerical values and Figure S2 for an enlarged plot of the relative total twist uptake on each nucleosome). The twist of supercoiling takes account of both the rotational and the translational contributions to the wrapping of DNA strands about one another, as opposed to the rigid-body or base-pair step parameter of the same name[44] widely used to quantify the spatial disposition of successive base pairs and to build accurate three-dimensional models. The two twists differ by a few degrees if the DNA axis, i.e., the pathway described by the centers of successive base pairs, is broken by lateral displacement of the base pairs. The differences are especially pronounced when the so-called Slide, the displacement of successive base pairs in the direction of their long-axes and a deformation of local structure known to control the superhelical pitch of nucleosomal DNA,[45] is a large positive value; see the effects of both Slide and Shift (the lateral displacement of base pairs along their short axes) on the difference in the two twists (τ_{scoil} and Ω_{step}) in all nucleosome structures in Figure S3 and the differences in the two values at individual steps in different structures in Figures S4 and S5.

The groupings of nucleosome structures, obtained by clustering the values of twist uptake over the central 141 bp of each entry (see Methods), show distinct differences at the same locations known in isolated structures to accommodate effects of DNA chain length or ligand binding,[24–27] in the vicinity of base pairs ±20 and ±50. What is new in the composite data are the various combinations of under- and overtwisting at these four sites and the net effects of these local differences on the twisting of the nucleosome as a whole, revealing sets of structures that are under- or overtwisted relative to the average (see the column of net twist uptake $\Delta Tw_{[-70,+70]}$ values between base pairs ±70 at the right of Figure 2). The undertwisted structures occur primarily in the

groupings labeled III, V, and VII. The structures in groups III and V stand out from all others in being undertwisted near base pairs ±50, with those in group III wrapping 146 bp and those in group V 147 bp. The structures in group V further stand out in terms of the relative spatial locations of chain ends, comprising most of the small, outlying set of nucleosomal pathways illustrated in Figure 1(d); see Table S2 for the coordinates of terminal base pairs in individual structures and Figure S6 for color-coded representations of twist uptake within each grouping of nucleosome structures. The undertwisted structures in group VII span a broader range of pathways with two configurations resembling those in group V and the remainder similar to those in group III. The twist uptake in group VII is neutral in the sense that roughly half of the structures are overtwisted and half are undertwisted. The overtwisted structures, however, do not show the same degree of overtwisting at individual steps as those in group VI; note the different range of color variation near base pairs ±20 and ±50 for these two groupings in Figure 2. Whereas the structures in group VI wrap 146 bp and come exclusively from X-ray data, those in group VII wrap 145–167 bp and derive from both X-ray and cryo-EM measurements. The differences in twist uptake occur in concert with subtle changes in the contact patterns of DNA with histones H2A and H2B, with the sites of close interatomic contact narrowed and shifted by 1 bp toward the ends of the underwound DNA pathways (see Figure S7).

The differences in net twist within each grouping reflect subtle differences in the local twist along the complete length of nucleosomal DNA. As evident from the levels of red/blue shading, the degree of over/undertwisting varies within the highlighted regions as well as in other parts of the DNA. For example, the ~30° difference in twist accumulated between base pairs ±70 in the R45C

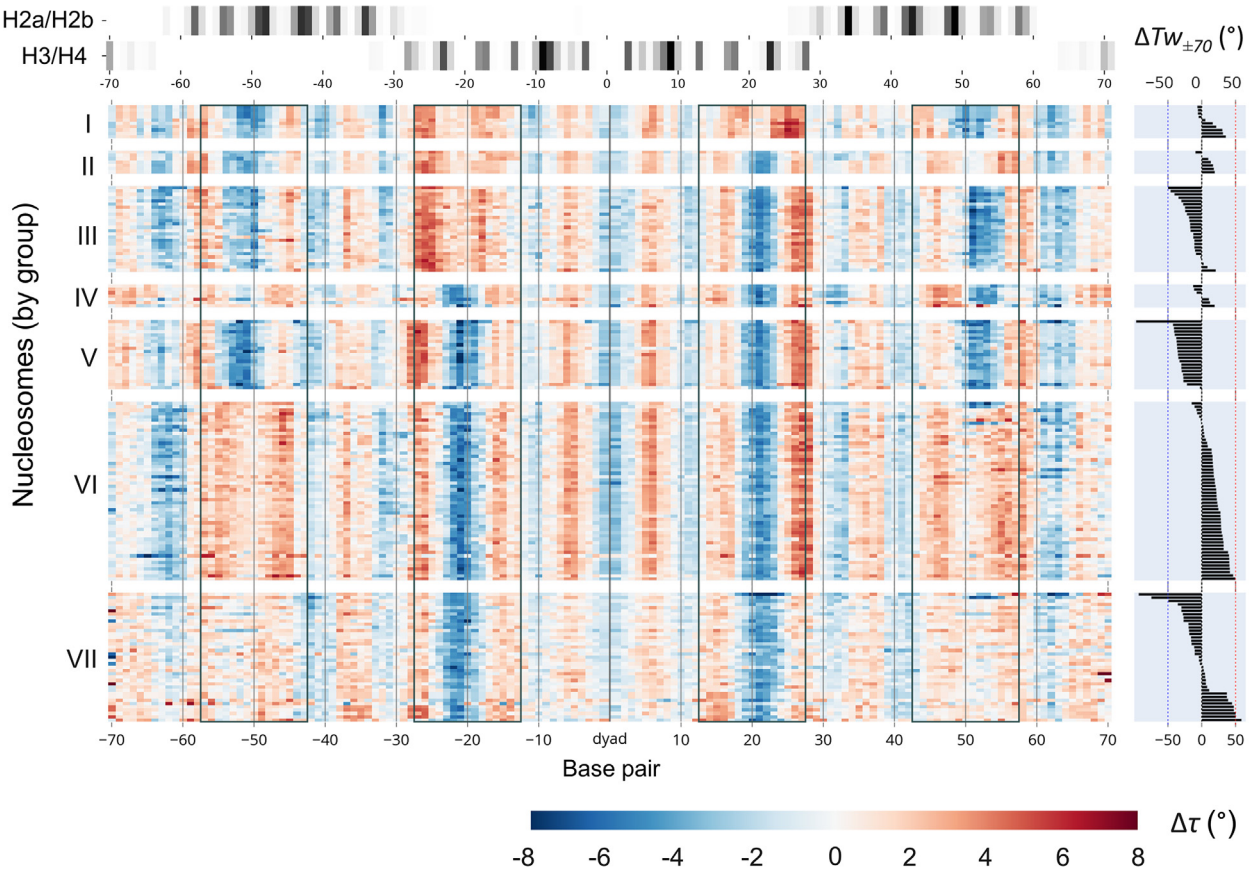

Figure 2. Color-coded heat map of the uptake of twist $\Delta\tau$, in degrees, relative to the average over all structures (~34.5°), at each base-pair step along the central 141 base pairs of the nucleosome core particles shown in Figure 1, with undertwisted steps depicted in shades of blue and overtwisted steps in shades of red. The seven panels show seven unique modes of twist build-up along the DNA (I-VII) found upon clustering the values of $\Delta\tau$ (see Methods). Gray lines highlight the ±7 bp regions around locations ±20 and ±50, which contain the largest differences in twist among the nucleosome groups. The bands across the top denote the sites and frequencies of DNA-protein contact. The vertical bar graph on the right shows the net twist uptake $\Delta Tw_{[-70,\ 70]}$, in degrees, over all steps compared to the average.

and R45H SIN mutant nucleosome structures (pdb files 1p3f and 1p3i,[46] respectively), both included in group III, reflects a jump in twist of ~12° in the vicinity of base pairs –50 and +60 in combination with smaller, primarily positive increments of twist spread over the remainder of the DNA (Table S2). There are not yet enough data to determine the precise molecular features that distinguish the undertwisted from the overtwisted nucleosomes, e.g., structures bearing additional proteins fall into both structural categories as do those bearing α-satellite sequences and the Widom 601 sequence. By contrast, structures incorporating 147 bp of DNA show a greater propensity for DNA undertwisting vs. overtwisting compared to those bearing 145 and 146 bp, which incorporate both under- and overtwisted nucleosomal DNA. As illustrated below, the wide extremes in net twist found in the complete set of structures have profound effects on other features of nucleosomal DNA as well as on the predicted properties of oligonucleosome arrays and nucleosome-bearing DNA minicircles.

Inter-gyre deformations

Knowing that the lateral displacement of successive base pairs accompanies change in DNA twist and that such displacements underlie the superhelical pitch of nucleosomal DNA,[45] we examined the distances between the centers of DNA base pairs separated by a complete turn of superhelix (here set to 78 bp; see Methods) in the different groups of nucleosomes. These values provide a measure of both the pitch and the separation of DNA gyres, and the variation in these distances among the different groupings offers an estimate of deformations that might contribute to nucleosomal gaping. The largest differences occur, as expected, between the under- and overtwisted sets of nucleosomes, noticeably in the vicinity of base pairs ±20 and ±50 and the base pairs on the opposite gyre, i.e., base pairs near ∓60 and ∓30, respectively (Figure 3). The average separation of gyres at locations ±40, directly across from the dyad, is similar in the two structural families. The opening of DNA in the general direction of the superhelical axis in currently available nucleosome structures thus differs from the hypothesized, bivalve-like gaping mechanism,[16] occurring at offside locations susceptible to large variations in twist rather than opposite the dyad. The magnitude of opening, i.e., the difference between the inter-gyre distances in certain under- vs. overtwisted nucleosomes (groups V vs. VII), is as large as 5.8 Å in the vicinity of base pairs +57 and –21. The pattern of opening remains the same if the inter-gyre distances are measured in alternate ways (compare Figures 3 vs. S8 and see Methods).

The features of the nucleosome responsible for these differences in opening become apparent when the relevant residues are highlighted in close-up views of the different structures superimposed in a common reference frame (Figure 3b). Here we draw attention to the locations of the base-pair centers (spheres) at the sites of greatest inter-gyre separation and the DNA backbones attached to the surrounding residues (thin lines). The features of the undertwisted nucleosomes are depicted in green and the overtwisted nucleosomes in gold, with the leading strand of DNA shown in darker hues. As evident from the molecular images, the wider separation distance between base pairs –47 and +31 in overwound vs. underwound nucleosomes arises primarily from movements around base-pair –47, while that between –21 and +57 reflects more pronounced movements around base-pair +57. Both of the noted sites lie in the vicinity of residues that are locally overtwisted in one grouping of nucleosome structures and locally undertwisted in the other, i.e., sites ±50 in groups V vs. VII (Figure 2). The subtle differences in the nearby DNA pathways reveal accompanying changes in the widths of the DNA major and minor grooves. The opening and closing of the DNA gyres and the movements of specific base pairs become apparent when toggling between representative under- and overtwisted nucleosome structures (illustrated in Figure S9 and Video S1 with pdb files 1kx5[25] and 5b0z[47] from groups V and VI, respectively).

Nucleosome dimer deformability

The reorientation of terminal base pairs associated with the under- and overtwisting of nucleosomal DNA also affects the configurations of successive nucleosomes within an oligomeric array. Here we report the rotation of successive nucleosomes about the axis connecting their centers in simulated arrays, with 172- or 177-bp spacing between nucleosome dyads and the protein-free DNA linker subject to elastic deformations (Figure 4). All of the nucleosomal DNA in a given array follows the same rigid 3D pathway, with the base pairs adopting the steps taken by the central 145 bp in a representative under- or overtwisted structure (here pdb file 1kx5[25] or 5b0z,[47] respectively). The rotation is measured in terms of the torsion angle ϕ described by the planes containing the cylindrical axis of each nucleosome and the center-to-center axis (see Figure 4(a)). The distributions of torsion angles reflect the equilibrium structure and room-temperature fluctuations introduced in the DNA segments linking successive nucleosomes in Monte Carlo simulations of the dimeric fragments (see Methods). The base-pair steps within these protein-free linkers are subject to bending and twisting deformations characteristic of ideal, mixed-sequence B DNA (see Methods).

The differences in twist on the nucleosome models lead to significant reorientation of successive nucleosomes (Figure 4(b)), with the

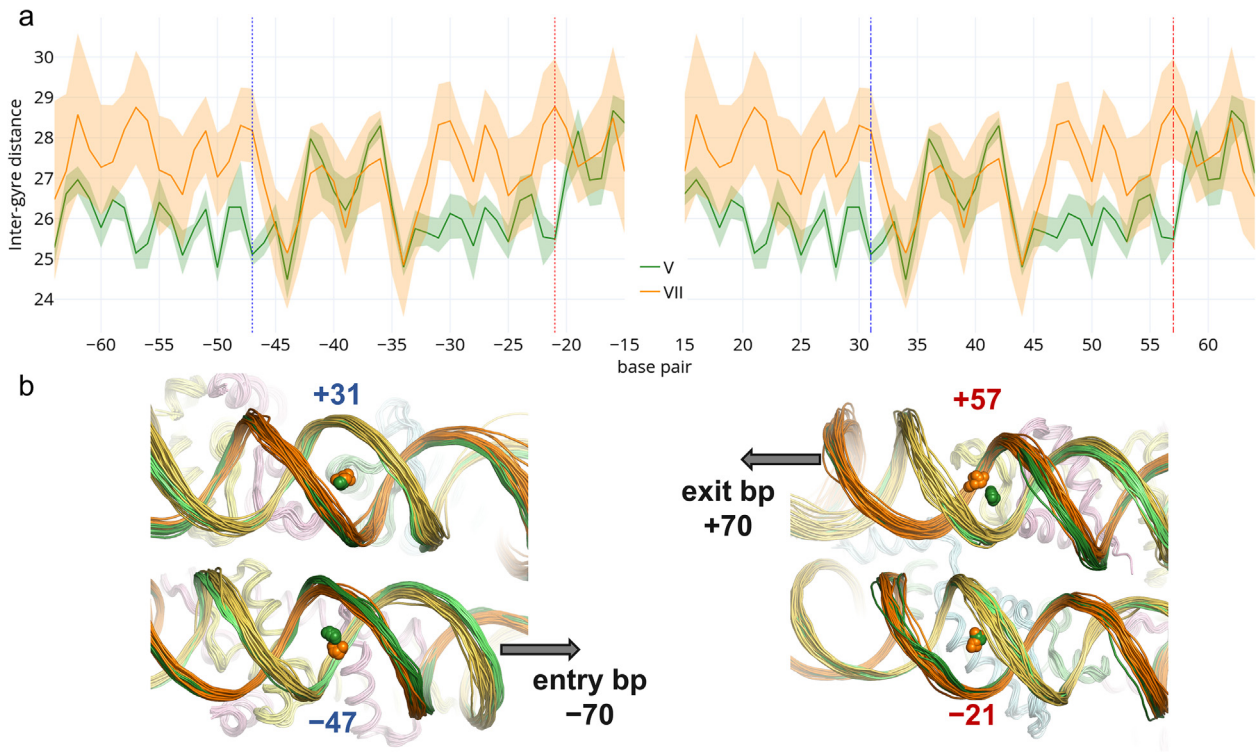

Figure 3. DNA gaping propensities, measured in terms of the distances between the centers of base pairs separated by a full (78-bp) superhelical turn, in representative under- and overtwisted nucleosome structures. (a) Average inter-gyre distances, in Ångströms, for nucleosome groups V (undertwisted) and VII (under- and overtwisted) shown respectively by green and gold centerlines. Shaded regions mark the 10th and 90th percentiles of values in each group; (b) Molecular close-ups of the DNA gyre separation in the vicinity of locations –47 and +31 (left) and –21 and +57 (right), with base-pair centers depicted by spheres and DNA backbones by thin lines. Nucleosomes are aligned as in Figure 1(a, b), and assigned the same color-coding as the distance plots. The leading strand of each structure is shown in a darker hue and the directions toward the entry and exit base pairs are noted.

values of the torsion angles ϕ in the undertwisted arrays reduced by ~70° compared to those in the overtwisted arrays regardless of the assumed spacing. The differences are expected to be even larger in models constructed with more extreme examples of underwound and overwound nucleosomal DNA. The respective twist uptake on the central 141 bp of the two nucleosome models used here is –27.4° (pdb file 1kx5[25]) and +3.7° (pdb file 5b0z[47]), values of much lesser magnitude than those of the most under- and overwound nucleosomes in the dataset (–97.3° in a cis platin-treated nucleosome (pdb file 3b6f[48]) and 58.8° in the human telomeric nucleosome (pdb file 6l9h[49]). The 5-bp change of spacing from 172 to 177 bp increases the torsion between successive nucleosomes by ~180°. The ~70° increment in ϕ found upon deformation of the nucleosome is comparable to the changes in the torsion angle brought about by a 2–3 bp increase in spacing between nucleosomes (compare the distributions of ϕ in Figure 4(a) with those associated with 1-bp changes in DNA linker length between undertwisted nucleosomes in Figure S10).

Oligonucleosome arrays

The differences in nucleosomal DNA twist and dinucleosome configuration noted above lead to distinctly different oligonucleosome properties. Monte Carlo simulated nucleosome arrays incorporating undertwisted nucleosomal pathways show different levels of compaction and interaction compared to those constructed with overtwisted pathways. For example, the computed sedimentation coefficients *s* of ensembles of 12-mer arrays bearing undertwisted (1kx5[25]) nucleosomes with 172-bp (dyad-to-dyad) spacing exceed those of identical constructs bearing overtwisted (5b0z[47]) nucleosomes (curves labeled 0 in Figure 5(a)). The lesser chain extension of the former system (Figure 5(b)) follows from the inverse dependence of *s* on average inter-nucleosome spacing (see Methods). Here we show average 3D configurations of the two arrays, with pathways between successive nucleosomes constructed from the averages of the rigid-body parameters used to specify the precise arrangement of neighboring base pairs within each simulated linker

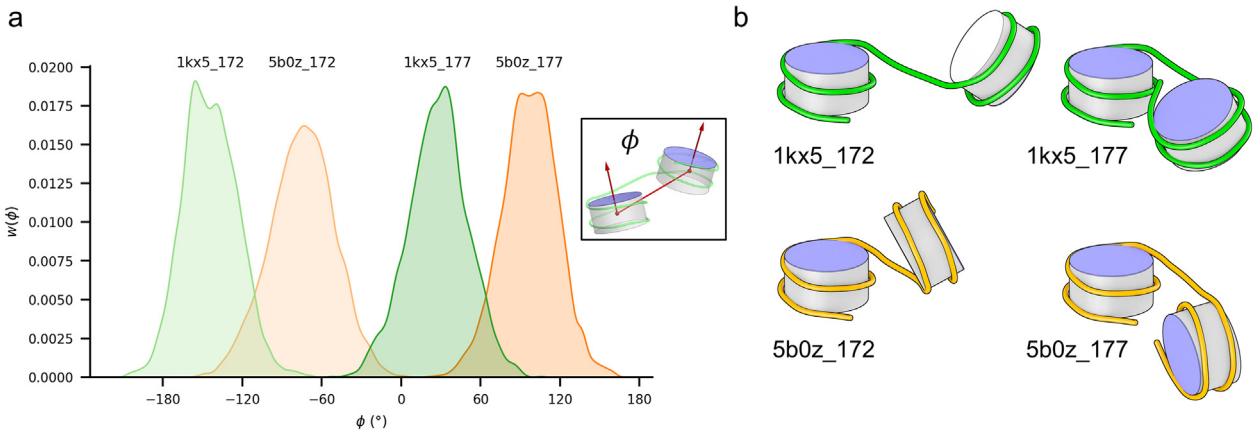

Figure 4. Influence of nucleosomal twist on the orientation of successive nucleosomes in Monte Carlo simulated 12-mer nucleosome arrays. (a) Distribution of the torsion angle ϕ, in degrees, described by the cylindrical axis of each nucleosome and the axis connecting their centers (inset). Nucleosome pairs selected from the central regions of simulated arrays with 172- or 177-bp spacing. Graphs are labeled in terms of the rigid pathway used to model the nucleosomal DNA, undertwisted (pdb file 1kx5[25]) and overtwisted (pdb file 5b0z[47]). Average values for 172-bp arrays bearing undertwisted and overtwisted nucleosomes are –144° and –75°, respectively. The corresponding values in arrays with 177-bp spacing are 28° and 97°; (b) Three-dimensional representations of successive dimers constructed from the average base-pair step parameters of the deformable protein-free linkers and the fixed nucleosome geometry in each simulated array. The histone core is shown as a light gray cylinder, with the top shown in blue to note the orientation in space. DNA is depicted as a thin tube connecting the base-pair centers. All four dimer configurations are aligned in a common frame on the first nucleosome.

configuration (see Videos S2 and S3 for sampled configurational states). The arrays comprised of overtwisted nucleosomes more closely match the reported sedimentation coefficients[50] than those with undertwisted nucleosomes. The limited overlap of the computed values of *s* with the observed data (41 S and 38 S in solutions of 150 mM NaCl and 1 mM $MgCl_2$, respectively), however, shows that the simulated arrays are less extended than those detected experimentally.

Simulated 12-mer arrays incorporating a mix of under- and overtwisted nucleosomes show intermediate behavior. For example, the incorporation of undertwisted nucleosomes compresses the extended arrays containing purely overtwisted (5b0z[47]) nucleosomes, with a shift in the predicted distribution of sedimentation coefficients toward higher values of *s* and a concomitant decrease in end-to-end displacement (Figure 5(a, b)). The modifications also tighten the arrays in terms of the frequency of interactions with nearby residues (Figure 5(c)), with a build-up in close ($\leq$110 Å nucleosome center-to-center) contacts between alternate (*i*, *i* + 2) nucleosomes and every third (*i*, *i* + 3) nucleosome in proportion to the number of undertwisted nucleosomes placed in the center of an array of overtwisted nucleosomes. The addition of overtwisted nucleosomes to an array containing purely undertwisted (1kx5[25]) nucleosomes has the opposite effect. The distribution of *s* shifts toward lower values, the chain becomes more extended on average, and the contacts between alternate (*i*, *i* + 2) and more distant (*i*, *i* + 5) nucleosomes decrease. The contacts between every third (*i*, *i* + 3) nucleosome, however, increase with the opening of the array. The average structures (Figure 5(b)) show how the mix of under- and overtwisted states might increase the accessibility of specific nucleosomes or introduce local bends in an array.

Mononucleosome DNA minicircles

The twist of DNA on the nucleosome also influences the configuration and topological mix of nucleosome-decorated DNA minicircles. Here we examine the effect of nucleosomal twist on 359-bp closed circular molecules bearing a single 141-bp nucleosome, a system that when free of proteins and is torsionally relaxed comprises 34.2 turns of a B-DNA double helix, i.e., 359 bp ÷ 10.5 bp/turn. The configuration of the protein-free DNA loop anchored by the nucleosome in each minicircle is described in terms of the angle γ formed by the base-pair normal at the mid-point of the loop with the plane containing the starting point, mid-point, and end-point of the loop (Figure 6(a)), i.e., the midpoint of the loop and the terminal base pairs of the nucleosome (see Figure S11). The rotation of the loop with respect to this plane is greatest and the elastic energy of the loop lowest when overtwisted nucleosomes are incorporated in minicircles of linking number *Lk* 33 (Figure 6(b)). Variation among the different nucleosome structures, here all 164 surveyed structures, leads to large-scale rearrangements of the topoisomer,

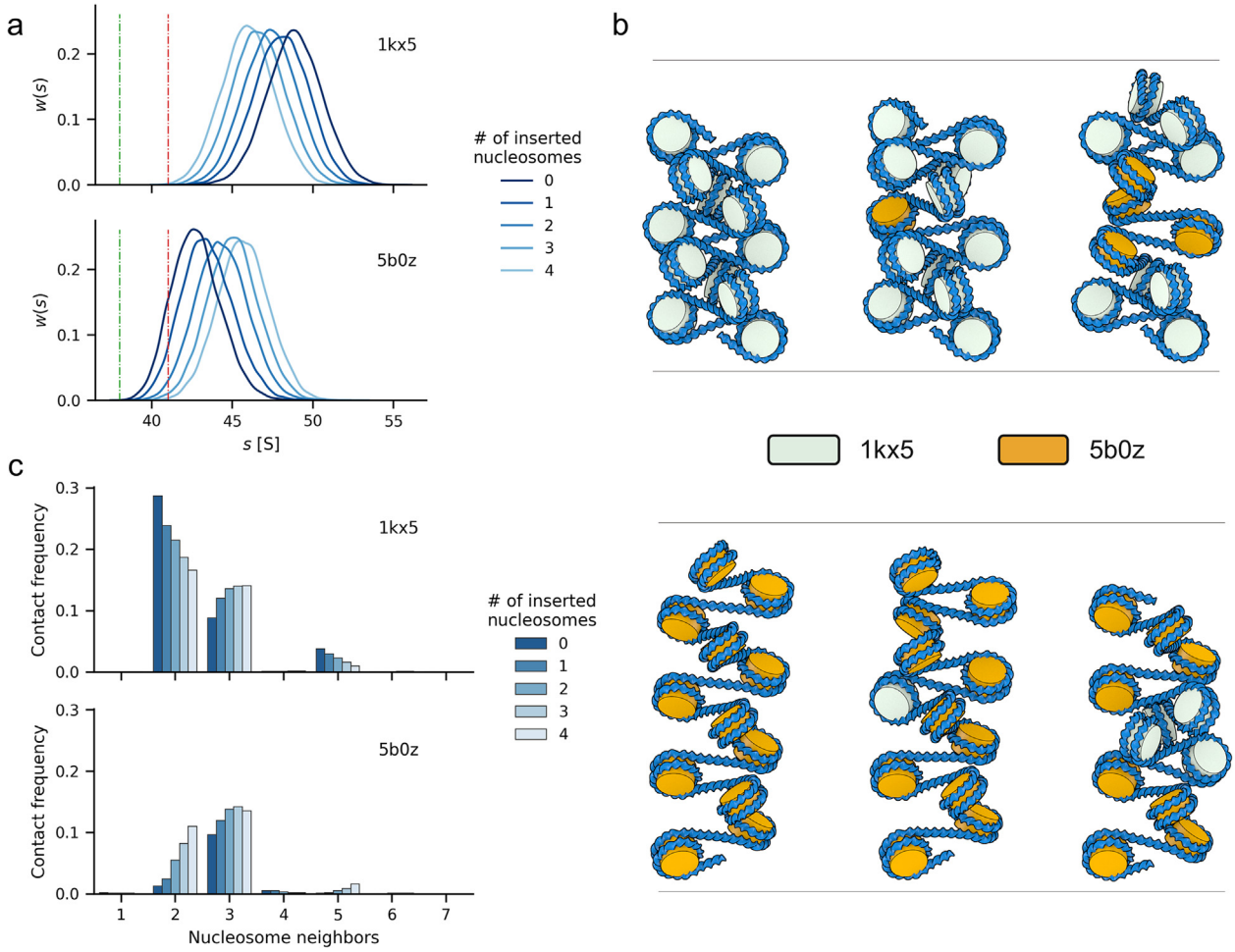

Figure 5. Effect of nucleosomal twist on the configurations of Monte Carlo simulated nucleosome arrays. (a) Distribution of the sedimentation coefficients *s* computed from simulated ensembles of 12-mer arrays with 172-bp spacing built from a mix of undertwisted (pdb file 1kx5[25]) and overtwisted (pdb file 5b0z[47]) nucleosome pathways. The color intensity indicates the number of overtwisted (top) or undertwisted (bottom) nucleosomes incorporated respectively into an array of otherwise undertwisted or overtwisted nucleosomes. Dashed vertical lines denote the values of $s_{20°,w}$ observed in ultracentrifgation studies.[50] (b) Three-dimensional configurations of nucleosome arrays constructed from the ensemble-averaged base-pair step parameters of the deformable protein-free DNA linkers, and the fixed nucleosome geometry in each simulated array. Nucleosomes are depicted as cylinders and color-coded in light green and gold for the undertwisted (1kx5) and overtwisted (5b0z) models, respectively. See the widely ranging configurations that contribute to these simplified, static images in Videos S2 and S3. (c) Frequency of intra-chain nucleosome contacts for the nucleosome arrays described in part (a). The horizontal axis denotes the nucleosome separation, where 1 indicates immediate neighbors, and so on. The contact frequency shown in the vertical axis is the measure introduced in 32,59.

including a ~40° decrease in γ and accompanying changes in the writhing number. The mix of topoisomers is greater when the nucleosomes are underwound, with more pronounced uptake of nucleosomes on minicircles of *Lk* 32 and 34 (Figure 6(c)). The differences in minicircle configuration brought about by changes in the nucleosomal twist become apparent when structures bearing pathways representative of different twist groupings are superimposed in a common reference frame (Figure 6(d)), here the central 141 bp of undertwisted nucleosomes from groups III and V (pdb files 1aoi[24] and 1kx5,[25] respectively) and an overtwisted nucleosome from group VI (pdb file 5b0z[47]) in the histone reference frame (see Methods and Table S5 for the writhing number and twist of these and all other constructs).

As anticipated from early studies of idealized nucleosome-decorated minicircles,[22,51] allowance for DNA 'breathing' changes the configurational picture. Peeling the DNA off the nucleosome lengthens and reorients the protein-free DNA loop. Here we simulate symmetric DNA breathing by removing the same number of base pairs from both ends of

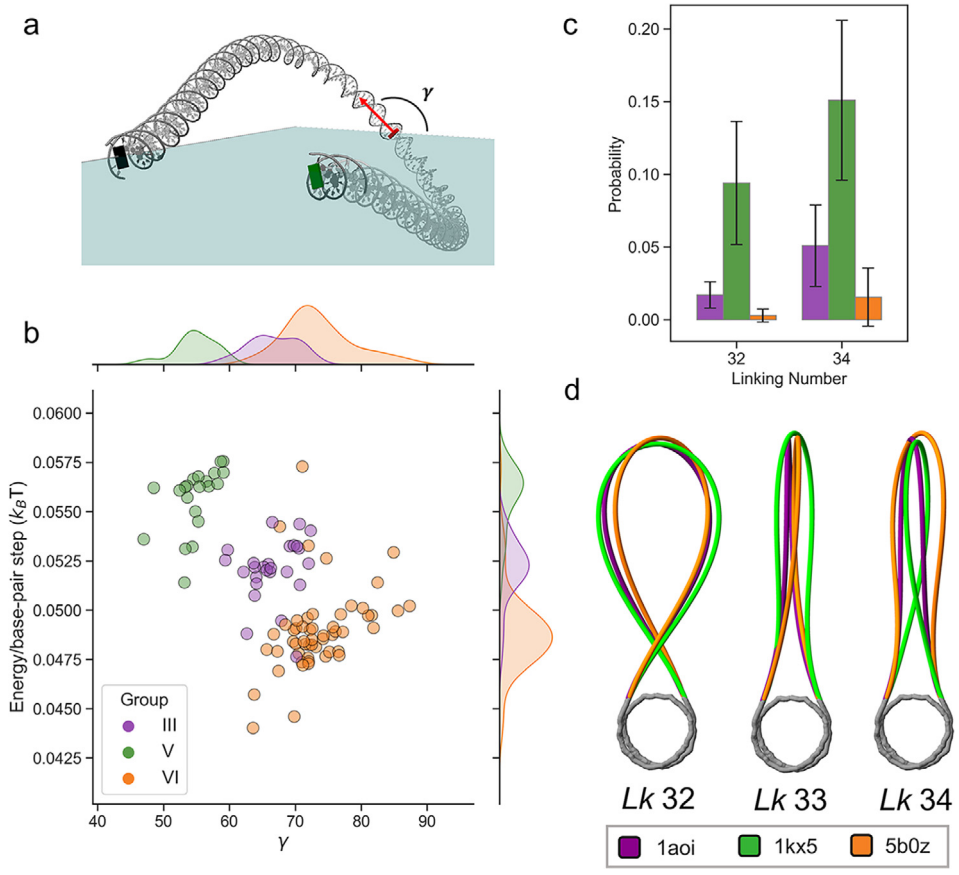

Figure 6. Effect of nucleosomal twist on the energy-optimized configurations of 359-bp DNA minicircles bearing a 141-bp rigid nucleosomal DNA fragment from different groupings of high-resolution structures. (a) Illustration of the out-of-plane rotation angle γ formed by the base-pair normal at the mid-point of the protein-free DNA loop with the plane containing the starting point, mid-point, and end-point of the loop (colored blocks); (b) Scatter plot of the optimized loop energy per base-pair step, in units of k_BT, vs. the angle γ, in degrees, for topoisomers with linking number *Lk* 33 (undertwisted groups III and V in purple and green, respectively, and overtwisted group VI in gold). Smooth curves on the edges of the scatter plot are the relative densities of individual parameters for each set of data; (c) Probabilities of occurrence of minor topoisomers (*Lk* 32 or 34) of minicircles found for the different groupings; (d) Schematics of optimized minicircles of specified linking number bearing nucleosome pathways representative of each structural grouping (pdb files 1aoi,[24] 1kx5,[25] 5b0z[47] from groups III, V, VI, respectively).

representative nucleosomal pathways on a 359-bp minicircle. The level of base-pair peeling, up to 10 bp from the ends of the nucleosome, is consistent with the adhesion energies estimated from the mechanical unzipping of single nucleosomes.[7,32] Chains bearing the overtwisted nucleosome resist changes in configuration as evidenced by the limited variation in the writhing number *Wr* of the dominant *Lk* 33 topoisomer upon peeling (Figure 7(a)), showing an increase in value only after 8 bp are removed from both ends of the superhelical pathway. The minicircles containing more undertwisted nucleosomes, by contrast, show comparable jumps in *Wr* upon removal of 4 bp from the nucleosome ends. The loop orientation angle γ shows similar sensitivity to nucleosomal twist (Figure S12), decreasing in magnitude at the sites where *Wr* shows a jump in value. Thus, localized changes in nucleosomal twist modulate both the breathing patterns and the overall fold of the modeled topoisomer. The minicircles bearing overtwisted nucleosomes adopt a wider mix of topological states than those incorporating undertwisted nucleosomes upon opening. The values of the average linking number $\langle Lk \rangle$ reveal a sizable population of the *Lk* 34 topoisomer as the DNA ends peel off the overtwisted nucleosome but limited variation upon similar peeling of the undertwisted nucleosomes (Figure 7(b)). The increase in *Lk* in the former minicircles occurs in concert with a reorientation of the DNA loop in excess of 100° and a jump in *Wr* of 0.5 or more depending upon the degree of breathing. Changes in *Lk*, of course, necessitate the breaking and rejoining of DNA strands. See sample configurations in Figure 7(c).

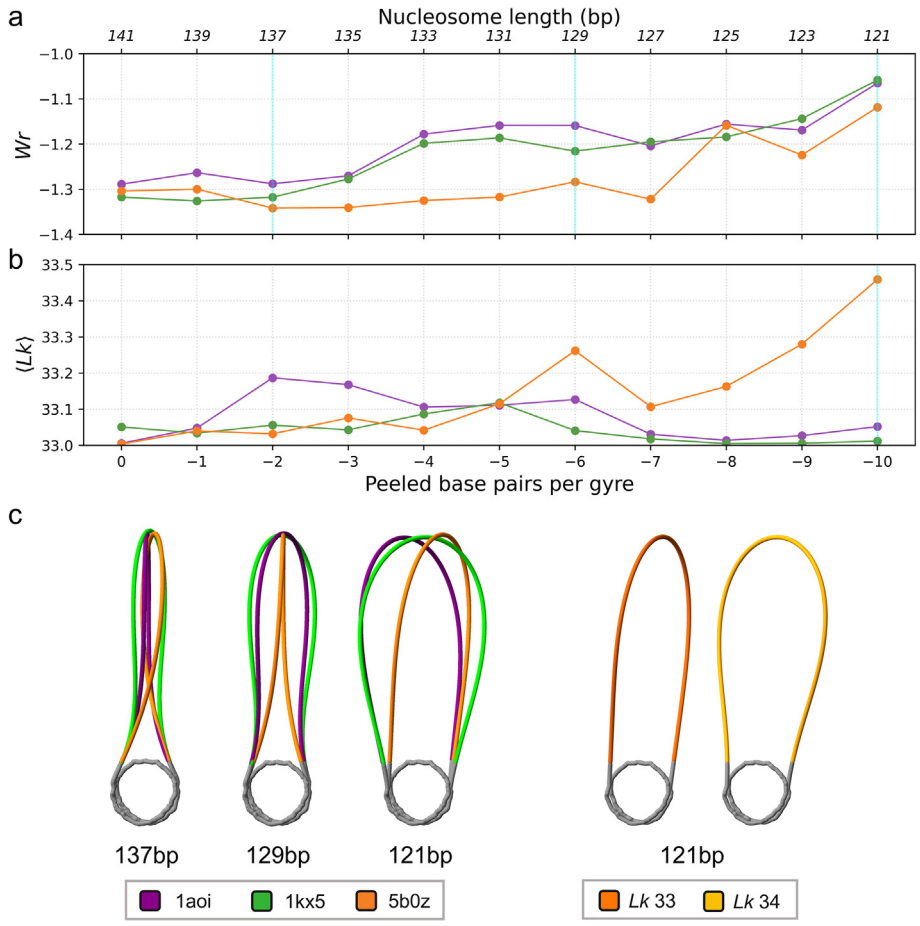

Figure 7. Effect of symmetric nucleosome 'breathing' on the energy-optimized configurations of 359-bp DNA minicircles bearing nucleosomal DNA fragments representative of three groupings of high-resolution structures. (a) Variation in the writhing number *Wr*, as a function of the degree of breathing, in minicircles of *Lk* 33; (b) Influence of the same levels of breathing on the average linking number ⟨*Lk*⟩ of all low-energy topoisomers, with breathing described in terms of either the number of DNA base pairs constrained on the nucleosome (upper label) or the number of base pairs peeled from both ends of the assembly (lower label); (c) Schematics illustrating the transitions between opened and closed states as DNA is unwrapped in *Lk* 33 topoisomers bearing different nucleosome structures (left) and the ~50:50 mix of topoisomers found when 10 bp are peeled off both ends of the overtwisted nucleosome (right). The cyan lines in (a) and (b) denote the degree of peeling associated with the structures shown in (c). See the legend to Figure 6.

Discussion

The idea that nucleosomal DNA is overtwisted goes back to early interpretations[52] of the classic DNA linking number paradox, i.e., why the level of supercoiling in the SV40 minichromosome—appro ximately one superhelical turn per bound histone-DNA assembly—is less than expected from the observed core-particle structure.[53] That is, the left-handed superhelical wrapping of DNA around the histone protein core was hypothesized to overwind the double helix, and indeed, the DNA superhelix reported in the first well-resolved nucleosome core particle structure (pdb file 1aoi[24]) exhibited such overwinding. The asymmetric uptake of torsional stress on arrays of 5S and 601 nucleosome positioning sequences[54,55] lends further support to this idea. These systems absorb positive stress more easily than negative stress in single-molecule manipulation studies.

Given these findings, it is not surprising that interest in the twisting of DNA on the nucleosome has focused more on localized perturbations within individual structures, e.g., the increase in twist needed to compensate for the lesser number of base pairs in one half of a structure compared to the other[24,25] or the incorporation of a ligand known to unwind the double helix,[26,27] than the twist of nucleosomal DNA as a whole. Moreover, the reported structural differences help to account for the constraints placed on DNA base pairs within the crystal lattice, in which the base pairs at the ends of each nucleosome stack against those of two neighbors. Interestingly, the first high-

resolution structural example of a nucleosome bearing the 601 sequence, in the complex with the RCC1 protein,[56] shows overtwisting of DNA at sites where the α-satellite DNA in the best-resolved structure[25] is less twisted. In contrast to the α-satellite DNA, the ends of the 601-RCC1 sequence are not constrained by crystal packing, raising the question of how the differences in sequence and/or crystallographic environment might give rise to the changes in DNA twist.

As shown herein, a local perspective on the nucleosome, including the top-down visualization of the protein-DNA assembly used in the structural literature (Figure 1), ignores the build-up of twist over the full length of nucleosomal DNA and any effects that the overall twist might have on global nucleosomal structure and larger-scale chromatin properties. Looking at this build-up in a large number of nucleosome structures reveals the surprising presence of globally undertwisted as well as globally overtwisted DNA (Figure 2). Indeed, the DNA pathways in many of the currently best-resolved structures are appreciably undertwisted. Moreover, many of the undertwisted DNA pathways differ from the overtwisted pathways at the very same sites, in the vicinity of base pairs ±50, as those found to distinguish the DNA in the histone assembly with RCC1[56] from that in the best-resolved core particle structure.[25] As expected from other anecdotal examples of DNA under/overwinding,[24–27] differences in DNA twist also occur in the vicinity of base pairs ±20. The set of high-resolution structures considered here includes six of the 2^4 = 16 possible combinations of DNA under- or overtwisting at positions ±20 and ±50. There are not yet enough structural examples to understand what might determine the observed build-up of twist on the nucleosomes (Figure S1). Under- and overtwisted pathways occur in approximately equal numbers on nucleosomes bearing α-satellite sequences, on those containing the Widom 601 DNA sequence, and on those incorporating other proteins. The pool of surveyed structures based on X-ray crystallographic data includes roughly equal numbers of examples with underwound and overwound DNA. The few structures determined by cryogenic electron microscopy methods, by contrast, tend to incorporate undertwisted DNA.

The gaping of nucleosomal DNA, measured in terms of the distances between the centers of DNA base pairs separated by a complete turn of superhelix, shows greatest variation in the vicinity of the base pairs (±20 and ±50) subject to the largest variations in twist (Figure 3). The largest variations, found between under- and overtwisted nucleosomes, are much smaller on average than the distances extracted from single-molecule fluorescence resonance energy transfer measurements at the same locations,[15] i.e., 3.2 ±0.9 Å within the high-resolution structures vs. 5–10 Å between dyes tethered to base pairs ±50 and ±28. Molecular images show that the measured gaping between under- and overtwisted structures arises primarily from movements near base pairs –47 and +57, sites where DNA comes in close contact with proteins H2A and H2B (Figure S7) and potentially involved in histone H2A·H2B dimer exchange[57] and nucleosome disassembly.[9,21] The inter-gyre separation at these and other sites is consistently greater in nucleosomes bearing overwound than underwound DNA. The separation of gyres at locations ±40, directly across from the dyad, is similar in all nucleosome structures, suggestive of an off-center opening as opposed to the hypothesized bivalve-like gaping of the nucleosome.[16]

The differences in twist between under- vs. overwound nucleosomal DNA lead to sizable reorientation and displacement of the ends of these structures, even when constrained by crystal packing. The terminal base pairs of one nucleosome do not necessarily stack in a uniform manner against the ends of another. The DNA twist at the 'nick' between neighboring nucleosomes (i.e., the double-stranded break between terminal base pairs along the pseudo-continuous DNA pathway connecting one nucleosome to the next) is thus highly variable (see Table S2 and Figure S13). Moreover, the differences in rotation/displacement of base pairs at the ends of undertwisted vs. overtwisted pathways change the orientation of successive nucleosomes and the underlying configurations of nucleosome-decorated DNA chains (Figure 4), along the same lines as the rearrangement of successive nucleosomes brought about by changes in nucleosome spacing[58–60] and expected from the rotation of base pairs along the DNA double-helical structure. The reorientation of nucleosomes associated with changes in twist in simulated nucleosome arrays is comparable to that associated with 2–3 bp changes in nucleosome spacing, depending upon the precise models used to describe the nucleosomal DNA. The differences in dinucleosome configuration lead, in turn, to distinctly different oligonucleosome properties (Figure 5). Simulated nucleosome arrays incorporating undertwisted nucleosomal pathways show different levels of compaction and interaction compared to those constructed with overtwisted pathways. Arrays incorporating a mix of under- and overtwisted nucleosomes show how the mix might increase the accessibility of specific nucleosomes or introduce local bends in an array. The degree of large-scale deformation depends, in turn, on the dynamic equilibrium between twist-defect states in individual nucleosomes, crudely approximated in the present work by two distinctly different DNA pathways localized at selected sites along the modeled arrays.

Finally, the twist of DNA on the nucleosome also influences the predicted configurations and

topological mix of nucleosome-decorated DNA minicircles. The choice of nucleosomal pathway introduces subtle perturbations in the folding of the protein-free loop anchored by the nucleosome in a modeled 359-bp closed circular molecule (Figure 6). The variation in twist among these structures introduces large-scale changes in the overall folding and topological mix of minicircles, again suggestive of how nucleosome dynamics may affect the larger-scale features of chromatin. Allowance for DNA 'breathing' changes the configurational picture, with a chain bearing a representative overtwisted nucleosomal pathway more resistant than one with an undertwisted pathway to spatial rearrangements and better able to accommodate a wider range of topological states as DNA peels off the two ends of the structure (Figure 7). The capability to take account of realistic features of the nucleosome makes it possible to examine how subtle changes in local structure, such as localized under- and overtwisting of DNA, may contribute to large-scale changes in chromatin architecture.

Materials and Methods

Nucleosome data and annotation

The deformations of nucleosomes reported herein are based on the configurations of 164 core particle structures of 3.5 Å or better resolution stored in the Protein Data Bank (pdb)[38] as of August, 2020. The examples were found from a search for all files containing the character string "nucleosom" and the coordinates of eight or more polypeptide chains and two or more DNA chains. In view of the discrepancies among the various files, we have implemented an automated scheme to annotate DNA residues in a consistent manner and to assign standard reference frames to each core-particle structure. We identify the dyad base pair through comparison with the currently best-resolved nucleosome structure (pdb file 1kx5[25])— first by least-squares fitting of the histone C^{α} atoms in each structure (residues 17–96, 35–119, 45–131, and 31–92 in the folded domains of histones H2A, H2B, H3, and H4, respectively) to those in the globular core of the 1kx5 reference (Table S4) and then determining which DNA base pair of the fitted structure lies closest to the axis passing through the dyad of the 1kx5 reference. Nucleotides can then be numbered consistently, with the dyad corresponding to base-pair zero and the residues of greatest negative and positive value located respectively at the 5′- and 3′-ends of each DNA strand. The leading strand is taken as the first DNA chain listed in the pdb file. We focus attention here on the central 141 bp of each structure given the variability in both the length of DNA and the conformations of terminal residues in the archived structures.

Nucleosomal twist and gaping

The twist of nucleosomal DNA is described in terms of the twist of supercoiling, a quantity that when added to the writhing number of a covalently closed DNA molecule yields integral values of the linking number[42,43] (see Table S5), and the gaping in terms of the distances between the centers of DNA base pairs on neighboring gyres of each nucleosome. The writhing number is a standard measure of the global folding of DNA and the total twist quantifies the rotation of the DNA strands with respect to the line segments connecting successive base pairs. The linking number is a topological invariant, which remains constant as long as the DNA strands stay intact.[61,62] The twist of supercoiling takes account of the contributions of base-pair translations of a chiral nature as well as the relative orientations of successive base-pair frames. The large, lateral displacements (so-called Slide) known to accompany the sharp histone-induced bending of nucleosomal DNA[45] fall into this category and make a notable contribution to the total twist, 4° or more for every such move.[42] See Figure S3.

Values of the twist of supercoiling are based on a treatment of DNA as a discrete, or segmented, ribbon.[43] The centerline of the ribbon is defined by the origins of coordinate frames fitted on successive base pairs, and the edge by the tips of unit vectors along the axes that point toward the leading strand. The uptake of twist along this ribbon is expressed relative to the average uptake over all nucleosomes in the dataset, i.e., $\Delta Tw = Tw - \langle Tw \rangle$, where Tw is the sum of the twist of supercoiling τ at each base-pair step and $\langle Tw \rangle$ the corresponding average. While the twist uptake can be calculated for any range of base pairs, the reported data describe the symmetrical region between base-pairs ±70. Similar values are obtained if the ribbon is defined by the locations of the pyrimidine C6 and purine C8 atoms on the paired bases, i.e., using the midpoints of the vectors connecting C6 to C8 as the centerline of the ribbon and the locations of the respective C6/C8 atoms in the leading strand as the edge. The twist uptake tends to be much larger if measured in terms of the base-pair step parameter of the same name (see Table S3 for comparative values of the accumulated twist of supercoiling determined using the base-pair reference frame and the C6/C8 positions of each base pair and the accumulated values of the step parameter twist in each nucleosome).

The base-pair frames needed to evaluate the twist of supercoiling are identified with the 3DNA software[63] and checked for base-pair identity using the annotations of base interactions provided by DSSR.[64] Discrepancies between the observed vs. expected identities of paired bases are corrected by appropriate transformations of atomic coordinates. For example, the pyrimidines found in the conformationally unlikely *syn* glycosyl form and

associated with a complementary purine in a non-canonical parallel association in some structures are flipped by a 180° rotation about the C1′-N1 glycosyl linkage into an *anti* form with appropriate Watson-Crick geometry (e.g., cytosines at position –70 in pdb files 2nqb, 3c1b, 3kwq, etc.). See Table S6 for a complete listing of all 'corrected' glycosyl orientations. Structures found to contain flipped-out bases and other unusual base content (e.g., noncanonical pairs that cannot be converted to Watson-Crick arrangements by simple torsional changes) are excluded from the dataset.

The nucleosomal DNA pathways are divided into distinct groups based on deviations of the twist $\Delta\tau$ from the average over all base-pair steps. The values of $\Delta\tau$ over the central 140-bp steps of each structure are first smoothed by calculating a rolling average with a window of five at each position. These data are then organized using the HDBSCAN clustering algorithm,[65] an extension of the DBSCAN algorithm,[66,67] which clusters the data into groups with a preselected minimum number of members but without preconceived notions of the number of such groups or the need to normalize the data. Here using the default minimum group size of five that is built into the Python library implementation, the algorithm identifies a set of seven nucleosome clusters and assigns each structure either as a member of a cluster or as noise/unassigned. The distinct patterns of over- and under-twisting in the vicinity of base pairs ±20 and ±50 are then used as the basis of a second round of clustering, implemented with K-means,[68] to place all of the data, including the structures initially identified as noise, into seven groups. The latter clustering is based on the smoothed $\Delta\tau$ values within ±7 bp of base pairs ±50 and ±20, yielding the desired number of groups with 7–54 members.

The distance between DNA gyres, or gaping, is measured in terms of the distances between the centers of DNA base pairs separated by a complete superhelical turn. The superposition of nucleosome structures on the globular histone core of pdb file 1kx5, described above, allows for incorporation of the superhelical axis of the reference structure in each entry, i.e., the line that minimizes the sum of residuals of the distance from that line to each base-pair center.[45] The gaping is then determined in two different ways, first by identifying the base-pair centers from the two gyres that come in closest contact when viewed down the superhelical axis and then by choosing the base pairs with the most commonly determined superhelical spacing in all structures. The former base pairs are those with the shortest separation distances in the plane perpendicular to the superhelical axis and the latter are base pairs separated by the 78-bp average. The magnitude of gaping is given in both cases by the relative displacement of the selected points along the DNA superhelical axis.

Nucleosome arrays

Simulations of oligonucleosome arrays have been carried out along lines previously described,[58,69,70] with the nucleosomes treated as rigid bodies locked in the 3D arrangements found in different core-particle structures and the intervening linker DNA modeled as a series of base-pair steps subject to bending and twisting deformations consistent with the solution properties of ideal, mixed-sequence DNA.[71] The potential governing the changes in linker configuration ignores the base sequence-dependent structural and deformational features of DNA—incorporating a bending constant consistent with the persistence length of mixed-sequence DNA (~500 Å), limiting the variation in twist to values compatible with the topological and binding properties of DNA minicircles,[72,73] and treating the DNA as a naturally straight, inextensible molecule subject to isotropic bending and independent fluctuations in twist. The changes in configuration are measured relative to a naturally straight B-DNA helix with 10.5 bp per turn and an axial displacement of 3.4 Å/bp. with a bending deformation of 4.84° or a change in twist of 4.09° raising the energy per base-pair step by 0.5 k_BT.[71] This simple model helps to decipher the effects of the different nucleosomes on the global properties of the arrays.

Excluded volume, i.e., molecular overlap, is detected with software from a rigid-body simulator (OpenDE; www.ode.org), with spheres of 10 Å radius encircling the base pairs of linker DNA and cylinders (84 Å diameter × 35 Å height) around the nucleosomes. Also considered are electrostatic interactions between representative points on the negatively charged nucleotides, the positively charged N-terminal histone tails (placed at the sites where H2A, H2B, H3, H4 exit the 1kx5 histone core), and the 26 charge clusters found in previous analysis of the 1kx5 globular core.[58] The protein charge sites are embedded in the symmetrized 1kx5 reference frame (see locations in Table S7) for approximate placement on other nucleosomes.

Configurations of nucleosome-decorated DNA are generated through Markov-Monte Carlo sampling of the rigid-body parameters between successive base pairs along the protein-free DNA linkers, i.e., the angles (Tilt, Roll, Twist) describing the orientation of coordinate frames embedded in successive base-pair planes and the components of the displacement vector (Shift, Slide, Rise) joining successive base-pair centers.[44,63] Average chain configurations are built from the average values of the rigid-body parameters used to describe the orientations and displacements of successive base pairs along the flexible DNA linkers and the specific values at each step along the intervening rigid nucleosomes. The histone core is treated as a side group of the nucleosomal DNA, i.e., the protein

atoms are expressed in the reference frame of one of the base pairs in the nucleosome assembly. Different mixes of nucleosomal pathways are used as crude models of the experimentally known dynamic equilibrium of twist-defect states.[26,27]

The sedimentation coefficient s of an array of N nucleosomes is estimated with the formula of Hansen et al.,[74]

$$s = \left(1 + \frac{R}{N}\sum_i \sum_{j>i} \frac{1}{r_{ij}}\right) s_1, \quad (1)$$

where s_1 is the sedimentation coefficient and R the radius of a single nucleosome (taken to be 11 S and 56 Å, respectively), and r_{ij} is the distance between the origins of the superhelical reference frames of nucleosomes i and j.

The frequency of inter-nucleosome interactions is described in terms of the contact score introduced by Norouzi and Zhurkin,[32,59] where nucleosomes i and $i + n$ are considered to be in close contact if the center-to-center distance is less than 110 Å (the radii of two nucleosomes).

Nucleosome-bearing DNA minicircles

Looped DNA pathway. The models of nucleosome-bearing DNA minicircles consist of two regions, a fragment fixed along the DNA pathway observed in an arbitrary high-resolution nucleosome structure and a free connecting loop. The starting configuration of the protein-free DNA loop is described by a Bézier curve:

$$\mathbf{r}_{loop}(u) = \sum_{j=0}^{n} B_j^n(u)\ \mathbf{p}_j. \quad (2)$$

The $\mathbf{p}_j$ in this expression are controlling points that determine the contour of the curve, while the $B_j^n(u)$ are coefficients defined in terms of the parameter u, $0 \le u \le 1$, and the number of controlling points, $n + 1$[75]:

$$B_j^n(u) = \frac{n!}{j!(n-j)!}\, u^j (1-u)^{n-j}. \quad (3)$$

The coordinates $\mathbf{r}_{loop}$ and the tangents $\mathbf{r}'_{loop}$ at the boundaries of the loop, where u = 0 and 1, are directly related to the controlling points.

$$\begin{aligned} \mathbf{r}_{loop}(0) &= \mathbf{p}_0 \\ \mathbf{r}'_{loop}(0) &= n(\mathbf{p}_1 - \mathbf{p}_0) \\ \mathbf{r}_{loop}(1) &= \mathbf{p}_n \\ \mathbf{r}'_{loop}(1) &= n(\mathbf{p}_n - \mathbf{p}_{n-1}) \end{aligned}. \quad (4)$$

This property of the curve makes it possible to connect the protein-free loop smoothly to the nucleosomal DNA with an appropriate choice of controlling points $\mathbf{p}_0$, $\mathbf{p}_1$, $\mathbf{p}_{n-1}$, $\mathbf{p}_n$.

In practice, we choose $n = 3$ (i.e., four controlling points), generating a set of smooth, twice differentiable curves. Then, from Eq. (4) it follows that points $\mathbf{p}_0$ and $\mathbf{p}_3$ coincide with the termini of the nucleosomal fragment, while points $\mathbf{p}_0$ and $\mathbf{p}_1$ determine the tangent at the start of the loop and points $\mathbf{p}_2$ and $\mathbf{p}_3$ the tangent at the end of the loop. The locations of $\mathbf{p}_1$ and $\mathbf{p}_2$ are varied to fix the contour length at a specified value, i.e., DNA chain length. The constraints on nucleosomal end conditions in combination with the size of the minicircles treated herein limit the spatial range of $\mathbf{p}_1$ and $\mathbf{p}_2$, thereby preventing the introduction of inflection points in the Bézier curve and making it possible to construct base-pair-level models and characterize DNA topology along the lines described below.

We take advantage of the approximation of Gravesen,[76]

$$L = \frac{2L_c + (n-1)L_p}{n+1}, \quad (5)$$

to find the desired contour length L of the Bézier curve. Here L_c is the distance between the first and last controlling points, L_p the sum of the distances between consecutive controlling points (total polygon length), and $n + 1$ the number of controlling points of the desired curve. Given the fixed locations of $\mathbf{p}_0$ and $\mathbf{p}_3$ and the constraints on the directions in which the two intermediate points may move, the positions of $\mathbf{p}_1$ and $\mathbf{p}_2$ that satisfy the desired value of L for an order 3 curve are estimated with an iterative procedure, increasing the polygon length in successive steps,

$$\begin{aligned} \mathbf{p}_1^f &= \left(1 + \frac{\Delta L_p}{L_p^i}\right)(\mathbf{p}_1^i - \mathbf{p}_0) \\ \mathbf{p}_2^f &= \left(1 + \frac{\Delta L_p}{L_p^i}\right)(\mathbf{p}_3^i - \mathbf{p}_3) \end{aligned}, \quad (6)$$

and using Eq. (5) to estimate L. The $\mathbf{p}_1^i$, $\mathbf{p}_2^i$, and L_p^i refer respectively to the initial controlling points and polygon length at each step and the ΔL_p to the imposed change in polygon length. The Bézier curve, once determined, is then divided into equidistant segments close to the standard 3.4-Å B-DNA spacing. These points serve as the origins of the base pairs that are placed along the curve.

The ends of the loop are chosen to coincide with the origins of the coordinate frames on arbitrary base pairs within a given high-resolution structure and the tangents to coincide with the normals of the selected residues, i.e., with the vector $\mathbf{p}_3$–$\mathbf{p}_2$ running parallel to the normal of the first nucleosomal base pair and the vector $\mathbf{p}_1$–$\mathbf{p}_0$ parallel to that of the last nucleosomal pair. The locations and axes of the base-pair frames in a given structure are extracted from the pdb file using the 3DNA software[63] and corrected, as needed, as described above.

Looped base pairs. Base pairs are placed on the constructed Bézier curves by making use of the local [$\mathbf{t}(s)$,$\mathbf{n}(s)$,$\mathbf{b}(s)$] Frenet-Serret trihedron, which can be directly determined from the equation and derivatives of the curve at a desired position s.

The long and short axes of the base pairs, respectively termed here $\mathbf{d}_1(s)$ and $\mathbf{d}_2(s)$, lie in the $\mathbf{n}(s)$-$\mathbf{b}(s)$ plane and the base-pair normal $\mathbf{d}_3(s)$ coincides with the local tangent vector. The [$\mathbf{d}_1(s)$, $\mathbf{d}_2(s)$,$\mathbf{d}_3(s)$] base-pair frame is related to the [$\mathbf{t}(s)$, $\mathbf{n}(s)$,$\mathbf{b}(s)$] frame through $\omega(s)$, the angle through which the base pair rotates upon translation to position s along the curve,

$$\begin{aligned}\mathbf{d}_1(s) &= \cos\omega(s)\mathbf{n}(s) - \sin\omega(s)\mathbf{b}(s)\\ \mathbf{d}_2(s) &= \sin\omega(s)\mathbf{n}(s) + \cos\omega(s)\mathbf{b}(s)\,.\\ \mathbf{d}_3(s) &= \mathbf{t}(s)\end{aligned} \tag{7}$$

The [$\mathbf{d}_1(s)$,$\mathbf{d}_2(s)$,$\mathbf{d}_3(s)$] frames are used, in turn, to generate a complete atomic description of the chain. Each residue of the DNA loop is constructed in a manner analogous that used to build a standard fiber diffraction model,[77] with the coordinates $\mathbf{x}_{j,q}$ of the j^{th} atom in the q^{th} residue expressed in the reference frame of the Bézier curve by:

$$\begin{aligned}\boldsymbol{x}_{j,q} = \boldsymbol{r}(s_q) &+ \left[x_j\cos\omega(s) - y_j\sin\omega(s)\right]\boldsymbol{n}(s_q)\\ &+ \left[x_j\sin\omega(s) + y_j\cos\omega(s)\right]\boldsymbol{b}(s_q) + z_j\boldsymbol{t}(s_q).\end{aligned} \tag{8}$$

Here the $\boldsymbol{r}(s_q)$ are the coordinates of the evenly spaced points along the Bézier curve and the (x_j, y_j, z_j) are the coordinates of the jth base-pair atom in the standard Watson-Crick reference frame.[78] Atoms of the sugar-phosphate backbone from the canonical B-DNA fiber diffraction model[77] are appended to the base-pair frame at this stage of model building and treated as a rigid unit upon optimization (see below). Should an inflection point arise in the Bézier curve, the severe under- and overtwisting brought about by the accompanying discontinuity in the Frenet-Serret frame can be corrected with an energy optimization procedure, whereby the angle of twist is perturbed at consecutive base-pair steps until a minimum-energy state is achieved with all values of twist close to the average.

Topoisomer construction. The linking number Lk of the nucleosome-decorated minicircle depends upon the choice of $\omega(s)$ used in the construction of base-pair frames. The value of Lk is determined, along with the values of the writhing number and the total twist of supercoiling, from the generated frames using methods previously developed for the analysis of the topological properties of DNA expressed as a collection of discrete atoms.[43,51] The increment in $\omega(s)$ needed to obtain a desired value of the linking number for the composite nucleosome-Bézier construct depends upon the difference between the linking number Lk^i of the initially generated pathway and the desired value of Lk and the number of base-pair steps N_{loop} in the free DNA loop, i.e., $\Delta\omega(s) = (Lk - Lk^i)\ 2\pi/N_{\text{loop}}$. The imposed value of Lk has no effect on the pathway connecting the centers of base pairs in the different toposiomers but rather on the twisting of base pairs along the protein-free DNA loop. The global fold of the most torsionally stressed topoisomers, however, may change upon optimization of the total energy (cf. seq.). The torsional stress in the generated models is proportional to the difference between Lk and the relaxed value Lk_0 given by the integer nearest to N/10.5, where 10.5 is the assumed helical repeat of DNA and N the total number of base pairs in the nucleosome-decorated minicircle.

Loop optimization. The configuration of the protein-free DNA loop is optimized using a procedure that takes account of the local elasticity of DNA and can be applied to chain fragments in which the first and last base pairs are spatially constrained.[79] The deformations of successive base pairs are expressed in terms of the six aforementioned rigid-body parameters[44,63] and guided by the same potential[71] as that used in the simulation of nucleosome arrays. The configuration of the loop as a whole is monitored by a second set of variables that keep track of the vectorial displacements of successive base pairs in a global reference frame. The introduction of the latter quantities makes it possible to take direct account of the spatial constraints imposed on the DNA and to use unconstrained numerical optimization methods. The approach differs from the optimization of DNA elastic energy used in our earlier work,[80] which requires explicit specification of the forces and moments acting on the constrained base pairs (including an educated first guess of these values), or methods used by others[81] that take the boundary conditions into account through Lagrange multipliers. A Debye-Hückel term is used to prevent the self contact of DNA residues separated by 11 bp or more—taking the charge on each phosphate group to be –0.24 esu in accordance with the predictions of counterion condensation theory,[82] assuming the dielectric medium to be a 100 mM aqueous monovalent salt solution, and placing the charges of complementary residues on the base-pair center. The wide range of optimized structures reflects the choice of nucleosome model, here constrained to the DNA pathways in all 164 surveyed structures.

CRediT authorship contribution statement

Stefjord Todolli: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Visualization, Writing - review & editing. **Robert T. Young:** Software, Methodology, Validation, Investigation, Formal analysis, Data curation, Visualization, Writing - review & editing. **Abigail S. Watkins:** Software. **Antonio Bu Sha:** Formal analysis. **John Yager:** Formal analysis. **Wilma K. Olson:** Conceptualization, Supervision, Resources, Validation, Funding acquisition, Project administration, Writing - original draft, Writing - review & editing.

Acknowledgement

This work was generously supported by the U.S. Public Health Service under research grant GM34809 and the Rutgers Center for Discrete Mathematics and Theoretical Computer Science (REU award to ASW from National Science Foundation grant CCF-1852215).

Conflict of Interest Statement

The authors declare no conflicts of interest.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2021.167121.

Received 10 November 2020;
Accepted 21 June 2021;
Available online 28 June 2021

Keywords:
nucleosomal twist uptake;
nucleosome gaping;
oligonucleosome array;
DNA minicircle;
Monte Carlo DNA simulation;
undertwisted nucleosome

References

1. Zaret, K.S., (2020). Pioneer transcription factors initiating gene network changes. *Annu. Rev. Genet.*, **54**, 367–385.
2. Polach, K.J., Widom, J., (1995). Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. *J. Mol. Biol.*, **254**, 130–149.
3. Chen, Y., Tokuda, J.M., Topping, T., Meisburger, S.P., Pabit, S.A., Gloss, L.M., et al., (2017). Asymmetric unwrapping of nucleosomal DNA propagates asymmetric opening and dissociation of the histone core. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 334–339.
4. Mohapatra, S., Lin, C.-T., Feng, X.A., Basu, A., Ha, T., (2020). Single-molecule analysis and engineering of DNA motors. *Chem. Rev.*, **120**, 36–78.
5. Cui, Y., Bustamante, C., (2000). Pulling a single chromatin fiber reveals the forces that maintain its higher-order structure. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 127–132.
6. Brower-Toland, B.D., Smith, C.L., Yeh, R.C., Lis, J.T., Peterson, C.L., Wang, M.D., (2002). Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 1960–1965.
7. Hall, M.A., Shundrovsky, A., Bai, L., Fulbright, R.M., Lis, J.T., Wang, M.D., (2009). High-resolution dynamic mapping of histone-DNA interactions in a nucleosome. *Nat. Struct. Mol. Biol.*, **16**, 124–129.
8. Li, G., Widom, J., (2004). Nucleosomes facilitate their own invasion. *Nat. Struct. Mol. Biol.*, **11**, 763–769.
9. Gansen, A., Valeri, A., Hauger, F., Felekyan, S., Kalinin, S., Tóth, K., et al., (2009). Nucleosome disassembly intermediates characterized by single-molecule FRET. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 15308–15313.
10. Ngo, T.T.M., Zhang, Q., Zhou, R., Yodh, J.G., Ha, T., (2015). Asymmetric unwrapping of nucleosomes under tension directed by DNA local flexibility. *Cell*, **160**, 1135–1144.
11. Bilokapic, S., Strauss, M., Halic, M., (2018). Histone octamer rearranges to adapt to DNA unwrapping. *Nat. Struct. Mol. Biol.*, **25**, 101–108.
12. Mauney, A.W., Tokuda, J.M., Gloss, L.M., Gonzalez, O., Pollack, L., (2018). Local DNA sequence controls asymmetry of DNA unwrapping from nucleosome core particles. *Biophys. J.*, **115**, 773–781.
13. Miyagi, A., Ando, T., Lyubchenko, Y.L., (2011). Dynamics of nucleosomes assessed with time-lapse high-speed atomic force microscopy. *Biochemistry*, **50**, 7901–7908.
14. Blosser, T.R., Yang, J.G., Stone, M.D., Narlikar, G.J., Zhuang, X., (2009). Dynamics of nucleosome remodelling by individual ACF complexes. *Nature*, **462**, 1022–1027.
15. Ngo, T.T.M., Ha, T., (2015). Nucleosomes undergo slow spontaneous gaping. *Nucleic Acids Res.*, **43**, 3964–3971.
16. Mozziconacci, J., Victor, J.-M., (2003). Nucleosome gaping supports a functional structure for the 30nm chromatin fiber. *J. Struct. Biol.*, **143**, 72–76.
17. Sivolob, A., Lavelle, C., Prunell, A., (2003). Sequence-dependent nucleosome structural and dynamic polymorphism. Potential involvement of histone H2B N-terminal tail proximal domain. *J. Mol. Biol.*, **326**, 49–63.
18. Edayathumangalam, R.S., Weyermann, P., Gottesfeld, J.M., Dervan, P.B., Luger, K., (2004). Molecular recognition of the nucleosomal "supergroove". *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 6864–6869.
19. Zhu, F., Farnung, L., Kaasinen, E., Sahu, B., Yin, Y., Wei, B., et al., (2018). The interaction landscape between transcription factors and the nucleosome. *Nature*, **562**, 76–81.
20. Ramaswamy, A., Bahar, I., Ioshikhes, I., (2005). Structural dynamics of nucleosome core particle: comparison with nucleosomes containing histone variants. *Proteins*, **58**, 683–696.
21. Lee, J., Lee, T.-H., (2017). Single-molecule investigations on histone H2A–H2B dynamics in the nucleosome. *Biochemistry*, **56**, 977–985.
22. Zhang, P., Tobias, I., Olson, W.K., (1994). Computer simulation of protein-induced structural changes in closed circular DNA. *J. Mol. Biol.*, **242**, 271–290.
23. Tobias, I., Coleman, B., Olson, W.K., (1994). Dependence of DNA tertiary structure on end conditions: theory and implications for topological transitions. *J. Chem. Phys.*, **101**, 10990–10996.
24. Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., Richmond, T.J., (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, **389**, 251–260.
25. Davey, C.A., Sargent, D.F., Luger, K., Mäder, A.W., Richmond, T.J., (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.*, **319**, 1097–1113.
26. Suto, R.K., Edayathumangalam, R.S., White, C.L., Melander, C., Gottesfeld, J.M., Dervan, P.B., et al., (2003). Crystal structures of nucleosome core particles in

complex with minor groove DNA-binding ligands. *J. Mol. Biol.*, **326**, 371–380.
27. Edayathumangalam, R.S., Weyermann, P., Dervan, P.B., Gottesfeld, J.M., Luger, K., (2005). Nucleosomes in solution exist as a mixture of twist-defect states. *J. Mol. Biol.*, **345**, 103–114.
28. Winger, J., Nodelman, I.M., Levendosky, R.F., Bowman, G.D., (2018). A twist defect mechanism for ATP-dependent translocation of nucleosomal DNA. *eLife*, **7**, e34100
29. Li, M., Xia, X., Tian, Y., Jia, Q., Liu, X., Lu, Y., et al., (2019). Mechanism of DNA translocation underlying chromatin remodelling by Snf2. *Nature*, **567**, 409–413.
30. Farnung, L., Ochmann, M., Cramer, P., (2020). Nucleosome-CHD4 chromatin remodeler structure maps human disease mutations. *eLife*, **9**, e56178
31. Brandani, G.B., Niina, T., Tan, C., Takada, S., (2018). DNA sliding in nucleosomes via twist defect propagation revealed by molecular simulations. *Nucleic Acids Res.*, **46**, 2788–2801.
32. Norouzi, D., Zhurkin, V.B., (2018). Dynamics of chromatin fibers: comparison of Monte Carlo simulations with force spectroscopy. *Biophys. J.*, **115**, 1644–1655.
33. Eslami-Mossallam, B., Schram, R.D., Tompitak, M., van Noort, J., Schiessel, H., (2016). Multiplexing genetic and nucleosome positioning codes: a computational approach. *PLoS ONE*, **11**, e0156905
34. van Deelen, K., Schiessel, H., de Bruin, L., (2020). Ensembles of breathing nucleosomes: a computational study. *Biophys J.*, **118**, 2297–2308.
35. Freeman, G.S., Lequieu, J.P., Hinckley, D.M., Whitmer, J. K., de Pablo, J.J., (2014). DNA shape dominates sequence affinity in nucleosome formation. *Phys. Rev. Lett.*, **113**, 168101
36. Kenzaki, H., Takada, S., (2015). Partial unwrapping and histone tail dynamics in nucleosome revealed by coarse-grained molecular simulations. *PLoS Comput. Biol.*, **11**, e1004443
37. Lequieu, J., Córdoba, A., Schwartz, D.C., de Pablo, J.J., (2016). Tension-dependent free energies of nucleosome unwrapping. *ACS Cent. Sci.*, **2**, 660–666.
38. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., et al., (2000). The Protein Data Bank. *Nucleic Acids. Res.*, **28**, 235–242.
39. Yang, T.P., Hansen, S.K., Oishi, K.K., Ryder, O.A., Hamkalo, B.A., (1982). Characterization of a cloned repetitive DNA sequence concentrated on the human X chromosome. *Proc. Natl. Acad. Sci. U.S.A.*, **79**, 6593–6597.
40. Lowary, P.T., 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.
41. Harp, J.M., Hanson, B.L., Timm, D.E., Bunick, G.J., (2000). Asymmetries in the nucleosome core particle at 2.5 Ångstrom resolution. *Acta Crystallogr. D*, **56**, 1513–1534.
42. Britton, L., Olson, W.K., Tobias, I., (2009). Two perspectives on the twist of DNA. *J. Chem. Phys.*, **131**, 245101
43. Clauvelin, N., Tobias, I., Olson, W.K., (2012). Characterization of the geometry and topology of DNA pictured as a discrete collection of atoms. *J. Chem. Theor. Comp.*, **8**, 1092–1107.
44. Dickerson, R.E., Bansal, M., Calladine, C.R., Diekmann, S., Hunter, W.N., Kennard, O., et al., (1989). Definitions and nomenclature of nucleic acid structure parameters. *Nucleic Acids Res.*, **17**, 1797–1803.
45. Tolstorukov, M.Y., Colasanti, A.V., McCandlish, D., Olson, W.K., Zhurkin, V.B., (2007). A novel roll-and-slide mechanism of DNA folding in chromatin: implications for nucleosome positioning. *J. Mol. Biol.*, **371**, 725–738.
46. Muthurajan, U.M., Bao, Y., Forsberg, L.J., Edayathumangalam, R.S., Dyer, P.N., White, C.L., et al., (2004). Crystal structures of histone Sin mutant nucleosomes reveal altered protein-DNA interactions. *EMBO J.*, **23**, 260–271.
47. Suzuki, Y., Horikoshi, N., Kato, D., Kurumizaka, H., (2016). Crystal structure of the nucleosome containing histone H3 with crotonylated lysine 122. *Biochem. Biophys. Res. Commun.*, **469**, 483–489.
48. Wu, B., Dröge, P., Davey, C.A., (2008). Site selectivity of platinum anticancer therapeutics. *Nat. Chem. Biol.*, **4**, 110–112.
49. Soman, A., Liew, C.W., Teo, H.L., Berezhnoy, N.V., Olieric, V., Korolev, N., et al., (2020). The human telomeric nucleosome displays distinct structural and dynamic properties. *Nucleic Acids Res.*, **48**, 5383–5396.
50. Correll, S.J., Schubert, M.H., Grigoryev, S.A., (2012). Short nucleosome repeats impose rotational modulations on chromatin fibre folding. *EMBO J.*, **31**, 2416–2426.
51. Swigon, D., Coleman, B.D., Tobias, I., (1998). The elastic rod model for DNA and its application to the tertiary structure of DNA minicircles in mononucleosomes. *Biophys. J.*, **74**, 2515–2530.
52. Finch, J.T., Lutter, L.C., Rhodes, D., Brown, R.S., Rushton, B., Levitt, M., et al., (1977). Structure of nucleosome core particles of chromatin. *Nature*, **269**, 29–36.
53. Germond, J.E., Hirt, B., Oudet, P., Gross-Bellard, M., Chambon, P., (1975). Folding of the DNA double helix in chromatin-like structures from simian virus 40. *Proc. Natl. Acad. Sci. U.S.A.*, **72**, 1843–1847.
54. Bancaud, A., Conde e Silva, N., Barbi, M., Wagner, G., Allemand, J.-F., Mozziconacci, J., et al., (2006). Structural plasticity of single chromatin fibers revealed by torsional manipulation. *Nat. Struct. Mol. Biol.*, **13**, 444–450.
55. Kaczmarczyk, A., Meng, H., Ordu, O., van Noort, J., Dekker, N.H., (2020). Chromatin fibers stabilize nucleosomes under torsional stress. *Nat. Commun.*, **11**, 126.
56. Makde, R.D., England, J.R., Yennawar, H.P., Tan, S., (2010). Structure of RCC1 chromatin factor bound to the nucleosome core particle. *Nature*, **47**, 562–566.
57. Bowman, G.D., Deindl, S., (2019). Remodeling the genome with DNA twists. *Science*, **366**, 35–36.
58. Nizovtseva, E.V., Clauvelin, N., Todolli, S., Polikanov, Y.S., Kulaeva, O.I., Wengrzynek, S., et al., (2017). Nucleosome-free DNA regions differentially affect distant communication in chromatin. *Nucleic Acids Res.*, **45**, 3059–3067.
59. Bass, M.V., Nikitina, T., Norouzi, D., Zhurkin, V.B., Grigoryev, S.A., (2019). Nucleosome spacing periodically modulates nucleosome chain folding and DNA topology in circular nucleosome arrays. *J. Biol. Chem.*, **294**, 4233–4246.
60. Takizawa, Y., Ho, C.-H., Tachiwana, H., Matsunami, H., Kobayashi, W., Suzuki, M., et al., (2020). Cryo-EM structures of centromeric trinucleosomes containing a central CENP-A nucleosome. *Structure*, **28**, 44–53.
61. Călugăranu, G., (1961). Sur les classes d'isotopie des noeuds tridimensionnels et leurs invariants. *Czech. Math. J.*, **11**, 588–625.
62. White, J.H., (1969). Self-linking and the Gauss integral in higher dimensions. *Am. J. Math.*, **91**, 693–728.

63. Lu, X.-J., Olson, W.K., (2003). 3DNA: a software package for the analysis, rebuilding, and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.*, **31**, 5108–5121.
64. Lu, X.-J., Bussemaker, H.J., Olson, W.K., (2015). DSSR: an integrated software tool for dissecting the spatial structure of RNA. *Nucleic Acids Res.*, **43**, e142
65. McInnes, L., Healy, J., Astels, S., (2017). hdbscan: Hierarchical density based clustering. *J. Open Source Softw.*, **2**, 205.
66. Ester, M., Kriegel, H.-P., Sander, J., Xu, X., (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In: Simoudis, E., Han, J., Fayyad, U. (Eds.), *KDD'96: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, The AAAI Press, Menlo Park, California, pp. 226–231.
67. Schubert, E., Sander, J., Ester, M., Kriegel, H.-P., Xu, X., (2017). DBSCAN revisited, revisited: why and how you should (still) use DBSCAN. *ACM Trans. Database Syst.*, **42** (19)
68. McQueen, J.B., (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, pp. 281–297.
69. Kulaeva, O.I., Zheng, G., Polikanov, Y.S., Colasanti, A.V., Clauvelin, N., Mukhopadhyay, S., et al., (2012). Internucleosomal interactions mediated by histone tails allow distant communication in chromatin. *J. Biol. Chem.*, **287**, 20248–20257.
70. Clauvelin, N., Lo, P., Kulaeva, O.I., Nizovtseva, E.V., Dias-Montes, J., Zola, J., et al., (2015). Nucleosome positioning and composition modulate *in silico* chromatin flexibility. *J. Phys.: Condens. Matter*, **27**, 064112
71. Czapla, L., Swigon, D., Olson, W.K., (2006). Sequence-dependent effects in the cyclization of short DNA. *J. Chem. Theor. Comp.*, **2**, 685–695.
72. Horowitz, D.S., Wang, J.C., (1984). Torsional rigidity of DNA and length dependence of the free energy of DNA supercoiling. *J. Mol. Biol.*, **173**, 75–91.
73. Heath, P.J., Clendenning, J.B., Fujimoto, B.S., Schurr, J.M., (1996). Effect of bending strain on the torsion elastic constant of DNA. *J. Mol. Biol.*, **260**, 718–730.
74. Hansen, J.C., Ausio, J., Stanik, V.H., van Holde, K.E., (1989). Homogeneous reconstituted oligonucleosomes, evidence for salt-dependent folding in the absence of histone H1. *Biochemistry*, **28**, 9129–9136.
75. Mortenson, M.E., (1997). Geometric Modeling. John Wiley and Sons, Inc., New York, pp. 81–112 (Chapter 4).
76. Gravesen, J., (1997). Adaptive subdivision and the length and energy of Bézier curves. *Comp. Geom.*, **8**, 13–31.
77. Chandrasekaran, R., Arnott, S., (1989). The structures of DNA and RNA helices in oriented fibers. In: Saenger, W. (Ed.), *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology, Group VII/1b, Nucleic Acids*, Springer-Verlag, Berlin, pp. 31–170.
78. Olson, W.K., Bansal, M., Burley, S.K., Dickerson, R.E., Gerstein, M., Harvey, S.C., et al., (2001). A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.*, **313**, 229–237.
79. Clauvelin, N., Olson, W.K., (2021). Synergy between protein positioning and DNA elasticity: energy minimization of protein-decorated DNA minicircles. *J. Phys. Chem. B*, **125**, 2277–2287.
80. Coleman, B.D., Olson, W.K., Swigon, D., (2003). Theory of sequence-dependent DNA elasticity. *J. Chem. Phys.*, **118**, 7127–7140.
81. Zhang, Y.L., Crothers, D.M., (2003). Statistical mechanics of sequence-dependent circular DNA and its application for DNA cyclization. *Biophys. J.*, **84**, 136–153.
82. Manning, G.S., (1978). The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.*, **11**, 179–246.
83. Ujvári, A., Hsieh, F.-K., Luse, S.W., Studitsky, V.M., Luse, D.S., (2008). Histone N-terminal tails interfere with nucleosome traversal by RNA polymerase II. *J. Biol. Chem.*, **283**, 32236–32243.
84. Schrodinger_LLC, The PyMOL Molecular Graphics System, Version 1.8, 2015.