

Supporting Information

Collepardo-Guevara and Schlick 10.1073/pnas.1315872111

SI Text

Modeling Description

Our mesoscale oligonucleosome model integrates the following different coarse-grained descriptions for the nucleosome, histone tails, linker DNA, linker histones (LHs), and the physiological environment (see an illustration of the model components and the repeating oligonucleosome motif in Fig. S1):

- **Nucleosome:** The protein histone core (without the 10 protruding histone tails) and the DNA around it are modeled as a single rigid irregular body. The nucleosome surface is defined by 300 Debye–Hückel effective charges uniformly distributed on it. We use the discrete surface charge optimization (DiSCO) algorithm (1) to optimize the values of the charges so that they reproduce [typically within <10% error (2) the full atom electric field at distances >5 Å around the nucleosome core. In addition, each charge is assigned an effective excluded volume to account for the excluded volume of the nucleosome core. The 10 histone tails protruding out of each core are modeled as flexible chains of beads. Each histone tail bead comprises 5 aa and is positioned on the C_{β} atom of the center amino acid. Each tail bead is assigned a charge that equals the sum of the charges of its 5 aa multiplied by a scaling factor to account for salt-dependent screening (3). The first bead of each tail is rigidly attached to the nucleosome core at locations consistent with the 1KX5 PDB crystal structure (4). The stretching and bending flexibility constants of each tail interbead segment are modeled by harmonic potentials with parameters developed to mimic their atomistic flexibilities (5).
- **Linker DNA:** The DNA connecting consecutive nucleosomes is treated as a chain of spherical beads with a salt-dependent charge that mimics the electrostatic potential of linear DNA using the procedure of Stigter (6), and an excluded volume that prevents overlap with other chromatin components. The mechanical properties of the linker DNA chains (interbead stretching, bending, and torsion) are described with a modified worm-like chain model (7–9). The equilibrium DNA interbead segment is 3 nm (~ 9 bp), thus to model the nucleosome-repeat lengths (NRLs) of 173, 182, 191, 200, 209, 218, and 227 bp, we use two to seven DNA beads per linker (three to eight segments), respectively.
- **LHs:** These proteins are modeled based on the structure of rat H1d LH as predicted by Bharath et al. (10, 11) using fold recognition and molecular modeling. We use three beads, rigidly attached to one another, to model the C-terminal domain (two beads) and the globular domain (one bead) of LH, and neglect the relatively uncharged N-terminal domain. Each LH bead is assigned a Debye–Hückel charge, also optimized with DiSCO, and an excluded volume. The chain of LHs beads is placed on the dyad axis of each nucleosome.
- **Solvent and ions:** The water around the oligonucleosome is treated implicitly as a continuum. The screening of electrostatic interactions due to the presence of monovalent ions in solution (0.15 M NaCl) is treated using a Debye–Hückel potential (electrostatic screening length of 1.27 nm^{-1}) (5).

The oligonucleosome chain is formed by attaching to each nucleosome core an entering and an exiting linker DNA at an angle of 108° , which corresponds to the 147 DNA base pairs tightly wound ~ 1.7 times around the core. The first core only has an exiting linker DNA. Each of these linker DNA beads and nucleosomes has a local coordinate system assigned so that both DNA

and nucleosome beads are able to twist around their axis. The different chromatin components interact through electrostatic (Debye–Hückel) and excluded volume (Lennard-Jones) potentials. To sample the ensemble of oligonucleosome conformations at constant temperature, we use Monte Carlo (MC) simulations with four different moves, as described below.

Modeling Validation and Limitations

As reviewed recently (12), several innovative coarse-grained models to treat oligonucleosome systems have emerged in the past years. With the enhanced computer power and the development of novel algorithms for sampling phase space, these techniques have allowed us to study the time and length scales of the chromatin system beyond what would be feasible with traditional all-atom simulations. However, despite continuous efforts to use ever-more rigorous parameterizations, coarse-graining entails multiple approximations. Our mesoscale model can simulate moderate oligonucleosome systems (up to ~ 48 -mer fibers) taking into account important chemical and physical information, such as the charged and contoured nucleosome surface, the flexibility of histone tails, and the presence of LHs. Extensive experimental validation of our model supports its accuracy and the relevance of its predictions. Agreement with experiments includes (i) salt-dependent compaction measurements (sedimentation coefficients and packing ratios) of 12-mer chicken erythrocyte chromatin over a broad range of monovalent salt concentrations, with/without magnesium ions and with/without LHs (5); (ii) the diffusion and salt-dependent behavior of mononucleosomes, dinucleosomes, and trinucleosomes (13); (iii) the salt-dependent extension of histone tails (13); (iv) the irregular zigzag topologies of chromatin fibers (3, 14) and their enhanced compaction upon LH binding (5, 14); (v) linker crossing orientations (5); (vi) internucleosome interaction patterns consistent with cross-linking and electron microscopy experiments (14, 15); and (vii) the force extension behavior of 173- and 209-bp fibers (16, 17).

Our model does not include specific protein–protein interactions and desolvation effects, but these are expected to be relatively weak compared with the strong electrostatic interactions among chromatin components. The Debye–Hückel treatment omits ion–ion correlation effects. Although charge correlation effects are important for modeling systems with highly charged surfaces and multivalent counterions, an accurate modeling of these effects is unfeasible for the large oligonucleosome systems we study here, as it would require explicit treatment of ions and solvent. Overall, the model makes reasonable approximations to allow adequate sampling of large fiber systems and address global structural and thermodynamic questions. Detailed all-atom molecular dynamics studies are only possible for much smaller systems.

MC Algorithm

We sample oligonucleosome conformations through the following MC moves:

- **Pivot, translation, and rotation chain moves:** The global pivot move is implemented by randomly choosing one linker DNA bead or nucleosome core and a random axis passing through the chosen component. The shorter part of the oligonucleosome is rotated around this axis by an angle chosen from a uniform distribution. The local translation and rotation moves also randomly select an oligonucleosome chain component (linker DNA bead or core) and rotate or translate only that component by a randomly selected axis passing through it. All three MC moves are accepted or rejected based on the Metropolis

criterion (18). The angles and distances for the rotations and translations are sampled from uniform distributions with ranges adjusted so that the acceptance ratios remain close to 0.30 in each condition.

- Tail regrowth move: The tail regrowth move is implemented to sample histone tail conformations based on the configurational bias MC method (19, 20). It randomly selects a histone tail chain and regrows it bead by bead using the Rosenbluth scheme (21). To prevent histone tail beads from penetrating the nucleosome core, the volume enclosed within the nucleosome surface is discretized, and any trial configurations that place the beads within this volume are rejected automatically.

The pivot, translation, rotation, and tail regrowth moves are attempted with probabilities of 0.2, 0.1, 0.1, and 0.6, respectively.

Mesoscale Model Chromatin Energy

The total potential energy of our oligonucleosomes is a sum of the twisting, stretching, and bending energy of the linker DNA, stretching, and intramolecular bending of histone tails, total electrostatic energy, and excluded volume contributions:

$$E = E_T + E_S + E_B + E_{IS} + E_{IB} + E_C + E_V. \quad [S1]$$

The first term in Eq. S1 is the twisting energy of the DNA:

$$E_T = \frac{s}{2l_0} \sum_{i=1}^{N-1} (\alpha_i + \gamma_i - \phi_{N_s})^2, \quad [S2]$$

where s is the twisting rigidity of DNA, l_0 is the equilibrium separation distance between beads of relaxed DNA, N is the number of beads in the oligonucleosome chain, ϕ_{N_s} is the twist deviation penalty term per segment, and the sum $\alpha_i + \gamma_i$ gives the linker DNA twist at each bead location. The values of the twist penalty term per segment are determined by the NRL and are obtained as described in ref. 14.

The next two terms denote the stretching,

$$E_S = \frac{h}{2} \sum_{i=1}^{N-1} (l_i - l_0)^2, \quad [S3]$$

and bending energy of the linker DNA,

$$E_B = \frac{g}{2} \left[\sum_{i=1}^N (\beta_i)^2 + \sum_{i \in I_C} (\beta_i^+)^2 \right]. \quad [S4]$$

Here h and g are the stretching and bending rigidities of DNA, $l_i = |\mathbf{r}_{i+1} - \mathbf{r}_i|$ is the separation between consecutive DNA beads, I_C denotes a nucleosome particle within the oligonucleosome chain, and β_i and β_i^+ are bending angles. Further details on the geometric description of the oligonucleosome chain are provided in refs. 2, 5, and 14.

The fourth term, E_{IS} , represents the total stretching energy of the histone tails:

$$E_{IS} = \sum_{i \in I_C} \sum_{j=1}^{N_T} \sum_{k=1}^{N_{bj}-1} \frac{k_{b_{jk}}}{2} (l_{ijk} - l_{jk0})^2 + \frac{h_{tc}}{2} \sum_{i \in I_C} \sum_{j=1}^{N_T} |\mathbf{t}_{ij} - \mathbf{t}_{ij0}|^2. \quad [S5]$$

Here $N_T = 10N_C$ is the total number of histone tails, N_{bj} is the number of beads in the j th tail, and $k_{b_{jk}}$ is the stretching constant of the bond between the k th and $k+1$ th beads of the j th histone tail. l_{ijk} and l_{jk0} represent the distance between tail beads k and $k+1$, and their equilibrium separation distance, respectively. In the second term, h_{tc} is the stretching bond constant of the spring attaching the histone tail to the nucleosome core, \mathbf{t}_{ij} is the position

vector of the first tail bead in the coordinate system of its parent nucleosome, and \mathbf{t}_{ij0} is the position vector of the attachment site to the nucleosome core.

The fifth term, E_{IB} , represents the intramolecular bending contribution to the histone tail energies:

$$E_{IB} = \sum_{i \in I_C} \sum_{j=1}^{N_T} \sum_{k=1}^{N_{bj}-2} \frac{k_{\theta_{jk}}}{2} (\theta_{ijk} - \theta_{jk0})^2, \quad [S6]$$

where θ_{ijk} and θ_{jk0} represent the angle between three consecutive tail beads (k , $k+1$, and $k+2$) and their equilibrium angle, respectively, and $k_{\theta_{jk}}$ is the corresponding bending force constant.

The total electrostatic interaction energy of the oligonucleosome is given by E_C . These include nine types of Debye–Hückel interactions: nucleosome–nucleosome, nucleosome–linker DNA, nucleosome–tail, linker DNA–tail, linker DNA–linker DNA, linker DNA–nonparent LH, tail–tail, tail–LH, and LH–LH:

$$E_C = \sum_i \sum_{j \neq i} \frac{q_i q_j}{4\pi\epsilon\epsilon_0 r_{ij}} \exp(-\kappa r_{ij}), \quad [S7]$$

where q_i and q_j are the effective charges separated by a distance r_{ij} in a medium with a dielectric constant of ϵ and a salt-concentration dependent inverse Debye length of κ , ϵ_0 is the electric permittivity of vacuum. Within an individual linker DNA chain or histone tail chain, the beads belonging to the same chain do not interact electrostatically with each other as their interactions are already accounted for through the intramolecular force field (harmonic spring). Finally, linker DNA beads, LH beads, and histone tail beads directly attached to the nucleosome do not interact electrostatically with their parental nucleosomal pseudocharges. This is required to ensure that the attachment tail and linker DNA beads remain as close as possible to their equilibrium locations.

The last term, E_V , represents the total excluded volume interaction energy of the oligonucleosome. The excluded volume interactions are modeled using the Lennard-Jones potential and the total energy is given by

$$E_V = \sum_i \sum_{j \neq i} k_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right], \quad [S8]$$

where σ_{ij} is the effective diameter of the two interacting beads and k_{ij} is an energy parameter that controls the steepness of the excluded volume potential. These parameters are taken from relevant models of the components as described and given in refs. 5 and 14.

Analysis Measurements

Internucleosome Interactions. The matrices of internucleosome interactions $I'(i, j)$ describe the fraction of MC iterations that cores i and j are in contact with one another. Each matrix element is defined as

$$I'(i, j) = \text{mean}[\delta_{i,j}(M)], \quad [S9]$$

where M is the MC configurational frame, and the mean is calculated over converged MC frames used for statistical analysis, where

$$\delta_{i,j}(M) = \begin{cases} 1 & \text{if cores } i \text{ and } j \text{ are in contact at the MC frame } M, \\ 0 & \text{otherwise.} \end{cases} \quad [S10]$$

At a given MC step M , we consider nucleosomes i and j to be in contact if the shortest distance between the tail beads directly attached to i and the tail beads or core charges of core j is smaller than the tail–tail (σ_{tt})- or tail–core (σ_{tc})-excluded volume distance, respectively (3). In our computations we use this cutoff

value of 1.8 nm. Fig. S3 shows a typical 2D map $[I'(i, j)]$ of the frequency of histone tail mediated interactions for a zigzag fiber.

These matrices can be projected into normalized 1D maps

$$I(k) = \frac{\sum_{i=1}^{N_C} I'(i, i \pm k)}{\sum_{j=1}^{N_C} I(j)} \quad [\text{S11}]$$

that depict the fraction of configurations that nucleosomes separated by k linker DNAs interact with each other (i.e., are closer than their van der Waals radii) through all contacts involving a histone tail and any other component (i.e., linker DNA, nucleosome surface, or histone tail). These maps reveal the pattern of internucleosome interactions (dominant, moderate, weak) in a chromatin fiber, providing key insights into structural organization.

To characterize near-neighbor contacts, we use these internucleosome interaction patterns, $I(k)$. For an ideal solenoid fiber, $I(k)$ would show dominant $i \pm 1$ interactions, whereas ideal zigzags would display dominant $i \pm 2$ interactions.

The more diverse far-nucleosome interactions are assessed by the percentage of fibers in the ensemble that exhibit one or more interactions between neighbors separated by nine or more linker DNAs, a cutoff value much larger than the number of nucleosomes per turn. For perfectly straight and regular zigzag or solenoid fibers, the occurrence of long-range far-neighbor interactions is small (<10%; see Fig. 3B).

Tail Interactions. To calculate the interactions of tails with different nucleosome components we follow a similar procedure to that described above. Namely, we measure the fraction of the time that tails of a specific kind t ($t = \text{H2A1, H2A2, H2B, H3, and H4}$) in a chromatin chain are in contact with a specific component c of the chromatin chain ($c = \text{its parent nucleosome, a nonparental nucleosome, parent DNA linkers, or nonparental DNA linkers}$) by constructing 2D matrices with the following elements:

$$T'(t, c) = \text{mean} \left[\frac{1}{N_C N} \sum_{i \in I_c} \sum_{j=1}^N \delta_{i,j}^{t,c}(M) \right], \quad [\text{S12}]$$

where the mean is calculated using the converged MC configurations and

$$\delta_{i,j}^{t,c}(M) = \begin{cases} 1 & \text{if } j \text{ is a } c\text{-type component in contact with a tail} \\ & \text{of kind } t \text{ of nucleosome } i \text{ at frame } M, \\ 0 & \text{otherwise.} \end{cases} \quad [\text{S13}]$$

For a given frame M , we consider a specific t -kind tail of core i to be either free or in contact with only one of the N -chromatin components of the oligonucleosome chain. The t tail is in contact with a component of type c if the shortest distance between its beads and the beads or core charges of c is smaller than the shortest distance to any other type of component and also smaller than the relevant tail component-excluded volume distance. The resulting normalized patterns of interactions provide crucial information into the frequency by which different tails mediate chromatin interactions.

Bending Angles. The local bending angle between consecutive nucleosomes is defined as in ref. 5 as the angle formed between the vector exiting one nucleosome and the vector entering the next nucleosome. The former connects the centers of the first two linker DNA beads and the latter those of the last two linker DNA beads.

Calculation of Sedimentation Coefficients. As done before (14), we calculate the sedimentation coefficient of a given oligonucleosome

array conformation from the intercore distances. The sedimentation coefficient $S_{20,w}$ is approximated from S_{N_C} , where

$$\frac{S_{N_C}}{S_1} = 1 + \frac{R_1}{N_C} \sum_i \sum_j \frac{1}{R_{ij}}. \quad [\text{S14}]$$

Here, S_{N_C} represents $S_{20,w}$ for a rigid structure consisting of N_C nucleosomes of radius R_1 , R_{ij} is the distance between the centers of two nucleosomes, and S_1 is $S_{20,w}$ for a mononucleosome. We use $R_1 = 5.5$ nm and $S_1 = 11.1$ Svedberg ($1 \text{ S} = 10^{-13} \text{ s}$), as done previously (14).

Calculation of the Fiber-Packing Ratio and Fiber Axis Curvature. To calculate the fiber-packing ratio (number of nucleosomes per 11 nm of fiber length), we compute the length of the fiber axis passing through a chromatin fiber core. We define the fiber axis as a 3D parametric curve $\mathbf{r}^{\text{ax}}(i) = (r_1^{\text{ax}}(i), r_2^{\text{ax}}(i), r_3^{\text{ax}}(i))$, where $r_j^{\text{ax}}(i)$ ($j = 1, 2, \text{ and } 3$) are three functions that return the center positions of the i th nucleosome (r_{i1} , r_{i2} , and r_{i3}) in the x , y , or z direction respectively. We approximate these functions with polynomials of the form

$$r_j^{\text{ax}}(i) \approx P_j(i) = p_{1,j}i^2 + p_{2,j}i + p_{3,j}, \quad [\text{S15}]$$

by fitting the datasets $[r_{ij}]$ by a least-squares procedure. Due to the highly nonlinear fiber axis curves of bent ladder arrays, we use sixth-order order polynomials to approximate their fiber axes. We determine the coefficients of the polynomial $P_j(i)$ by minimizing the sum of the squares of the residuals l_j

$$l_j = \sum_{i=1}^{N_C} (r_{ij} - P_j(i))^2, \quad [\text{S16}]$$

which account for the differences between a proposed polynomial fit and the observed nucleosome positions. After determining the polynomial coefficients, we use Eq. S15 to produce N_C points per spatial dimension and compute the fiber length L_{fiber} as follows:

$$L_{\text{fiber}} = \sum_{i=1}^{(N_C-1)/2} |\mathbf{r}^{\text{ax}}(2i-1) - \mathbf{r}^{\text{ax}}(2i+1)|, \quad [\text{S17}]$$

where the distances are between every two consecutive nucleosome centers. The packing ratio (number of cores per 11 nm) is then calculated as the number of cores multiplied by 11 nm/ L_{fiber} . From the fiber axis, we define the local fiber radius for a given nucleosome core to be the perpendicular distance between a nucleosome core center and its closest linear fiber axis segment plus the nucleosome radius ($R_{\text{core}} = 5.5$ nm). We then average over all local fiber radii in a given fiber to obtain the fiber radius at each simulation frame. Finally, we repeat this procedure for each simulation frame and average the value to obtain a mean fiber radius. The fiber width, D_{fiber} , is twice that value. Additionally, from the parametric definition of the fiber axis, we identify the mean curvature of the chromatin fiber at each simulation frame as

$$\kappa_{\text{fiber}} = \frac{1}{N_C} \sum_{i=1}^{N_C} \frac{\mathbf{r}^{\text{ax}}(i) \times \mathbf{r}^{\text{ax}}(i)}{|\mathbf{r}^{\text{ax}}(i)|}, \quad [\text{S18}]$$

where $\mathbf{r}^{\text{ax}}(i) \approx (2p_{1,1}i + p_{2,1}, 2p_{1,2}i + p_{2,2}, 2p_{1,3}i + p_{2,3})$, and $\mathbf{r}^{\text{ax}}(i) \approx 2(p_{1,1}, p_{1,2}, p_{1,3})$.

Role of Histone Tails

The tails of all four histones contribute to chromatin condensation (22). Intrafiber and interfiber interactions are established

between the positively charged and flexible histone tails of one nucleosome with the charged and contoured surfaces of a neighboring core (tail–nucleosome), the linker DNAs entering/exiting their parent core (parent linker DNA), and the linker DNAs joined to other cores (nonparent linker DNA).

For the different nonuniform NRL fibers, Fig. S5 shows the fraction of configurations for which the different histone tails are in contact (as defined above) with a specific chromatin component (its parent nucleosome, a nonparental nucleosome, parent DNA linkers, or nonparental DNA linkers). We compare these patterns of interactions to those in uniform NRL fibers (14).

Bent ladders have a low frequency of tail–nucleosome interactions, especially with H4 tails, which reflects their ladder-like organization with less frequent face-to-face contacts. For fibers with medium-to-long NRLs, tail–nucleosome interactions are mostly mediated by the H3 tail due to its length, and the H4 tail due to its position on the nucleosome surface. The intensity of H4 tail–nucleosome interactions is larger for canonical and polymorphic fibers with medium average NRLs (between 195.5 and 209 bp), whereas that of H3 tails is stronger for canonical fibers with average NRLs smaller than 191 bp.

The H4 and H3 tails are known to have a strong impact in chromatin compaction (5, 14), and to be involved in both intrafiber and interfiber contacts with the DNA and other nucleosome surfaces (23–26). In agreement with these observed roles, interactions with nonparental DNA are also mostly mediated by H3 and H4 tails. The intensity of these interactions is strong for

the 191- to 218- and 191- to 227-bp fibers, which are the most compact fibers in our study. In addition, more intense interactions of the five tails with nonparental DNAs are noted in direct connection to our observation of increased interfiber interactions in bent fiber forms. That is, in interdigitated structures (i.e., predominant in 173- to 209-bp, 173- to 227-bp, and polymorphic systems) tail–nonparent DNA interactions are stronger compared with noninterdigitated uniform NRL systems. The only exception are the H3 interactions in bent ladders; although such arrays are interdigitated, DNA stems cannot form with LH, making it easier for the H3 tails to strongly interact with the more immediate parent DNAs.

In fact, due to their proximity with the nucleosome dyad, the H3 and H2A₂ (C termini of the H2A) tails mediate the majority of parent linker DNA interactions for all fibers. Tail–parent DNA interactions screen the electrostatic repulsion among entering and exiting DNAs, facilitating DNA stem formation and fiber condensation. The intensity of these interactions becomes stronger as the average NRL increases and the patterns for uniform and nonuniform NRL fibers lie very close.

Although H2A₁ (N termini of the H2A) and H2B tails spend most of the time interacting with their parent cores, the H2B tail–nucleosome interactions are maximum for fibers with the highest occurrence of far-neighbor contacts (fibers 9, 11, and 13). This suggests that H2B are involved in lateral interfiber interactions due to their ideal positions for interdigitation on the nucleosome periphery and their length.

- Zhang Q, Beard DA, Schlick T (2003) Constructing irregular surfaces to enclose macromolecular complexes for mesoscale modeling using the discrete surface charge optimization (DISCO) algorithm. *J Comput Chem* 24(16):2063–2074.
- Beard DA, Schlick T (2001) Computational modeling predicts the structure and dynamics of chromatin fiber. *Structure* 9(2):105–114.
- Arya G, Schlick T (2006) Role of histone tails in chromatin folding revealed by a mesoscopic oligonucleosome model. *Proc Natl Acad Sci USA* 103(44):16236–16241.
- Davey CA, Sargent DF, Luger K, Mäder AW, Richmond TJ (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J Mol Biol* 319(5):1097–1113.
- Arya G, Schlick T (2009) A tale of tails: How histone tails mediate chromatin compaction in different salt and linker histone environments. *J Phys Chem A* 113(16):4045–4059.
- Stigter D (1977) Interactions of highly charged colloidal cylinders with applications to double-stranded. *Biopolymers* 16(7):1435–1448.
- Allison S, Austin R, Hogan M (1989) Bending and twisting dynamics of short linear DNAs. Analysis of the triplet anisotropy decay of a 209 base pair fragment by Brownian simulation. *J Chem Phys* 90(7):3843–3854.
- Heath PJ, Gebe JA, Allison SA, Schurr JM (1996) Comparison of analytical theory with Brownian dynamics simulations for small linear and circular DNAs. *Macromolecules* 29(10):3583–3596.
- Jian H, Schlick T, Vologodskii A (1998) Internal motion of supercoiled DNA: Brownian dynamics simulations of site juxtaposition. *J Mol Biol* 284(2):287–296.
- Bharath MM, Chandra NR, Rao MR (2002) Prediction of an HMGB-box fold in the C-terminal domain of histone H1: Insights into its role in DNA condensation. *Proteins* 49(1):71–81.
- Bharath MM, Chandra NR, Rao MR (2003) Molecular modeling of the chromatosome particle. *Nucleic Acids Res* 31(14):4264–4274.
- Korolev N, Fan Y, Lyubartsev AP, Nordenskiöld L (2012) Modelling chromatin structure and dynamics: Status and prospects. *Curr Opin Struct Biol* 22(2):151–159.
- Arya G, Zhang Q, Schlick T (2006) Flexible histone tails in a new mesoscopic oligonucleosome model. *Biophys J* 91(1):133–150.
- Perišić O, Collepardo-Guevara R, Schlick T (2010) Modeling studies of chromatin fiber structure as a function of DNA linker length. *J Mol Biol* 403(5):777–802.
- Grigoryev SA, Arya G, Correll S, Woodcock CL, Schlick T (2009) Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. *Proc Natl Acad Sci USA* 106(32):13317–13322.
- Collepardo-Guevara R, Schlick T (2011) The effect of linker histone's nucleosome binding affinity on chromatin unfolding mechanisms. *Biophys J* 101(7):1670–1680.
- Collepardo-Guevara R, Schlick T (2012) Crucial role of dynamic linker histone binding and divalent ions for DNA accessibility and gene regulation revealed by mesoscale modeling of oligonucleosomes. *Nucleic Acids Res* 40(18):8803–8817.
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equations of State Calculations by Fast Computing Machines. *J Chem Phys* 21(6):1087–1092.
- Frenkel D, Mooij GCA, Smit B (1992) Novel scheme to study structural and thermal properties of continuously deformable molecules. *J Phys Condens Matter* 4(12):3053–3076.
- de Pablo JJ, Laso M, Suter UW (1992) Simulation of polyethylene above and below the melting point. *J Chem Phys* 96(3):2395–2403.
- Rosenbluth MN, Rosenbluth AW (1955) Monte carlo calculation of the average extension of molecular chains. *J Chem Phys* 23(2):356–359.
- Gordon F, Luger K, Hansen JC (2005) The core histone N-terminal tail domains function independently and additively during salt-dependent oligomerization of nucleosomal arrays. *J Biol Chem* 280(40):33701–33706.
- Dorigo B, Schalch T, Bystrycky K, Richmond TJ (2003) Chromatin fiber folding: Requirement for the histone H4 N-terminal tail. *J Mol Biol* 327(1):85–96.
- Zheng C, Lu X, Hansen JC, Hayes JJ (2005) Salt-dependent intra- and internucleosomal interactions of the H3 tail domain in a model oligonucleosomal array. *J Biol Chem* 280(39):33552–33557.
- Kan PY, Lu X, Hansen JC, Hayes JJ (2007) The H3 tail domain participates in multiple interactions during folding and self-association of nucleosome arrays. *Mol Cell Biol* 27(6):2084–2091.
- Kan PY, Caterino TL, Hayes JJ (2009) The H4 tail domain participates in intra- and internucleosome interactions with protein and DNA during folding and oligomerization of nucleosome arrays. *Mol Cell Biol* 29(2):538–546.

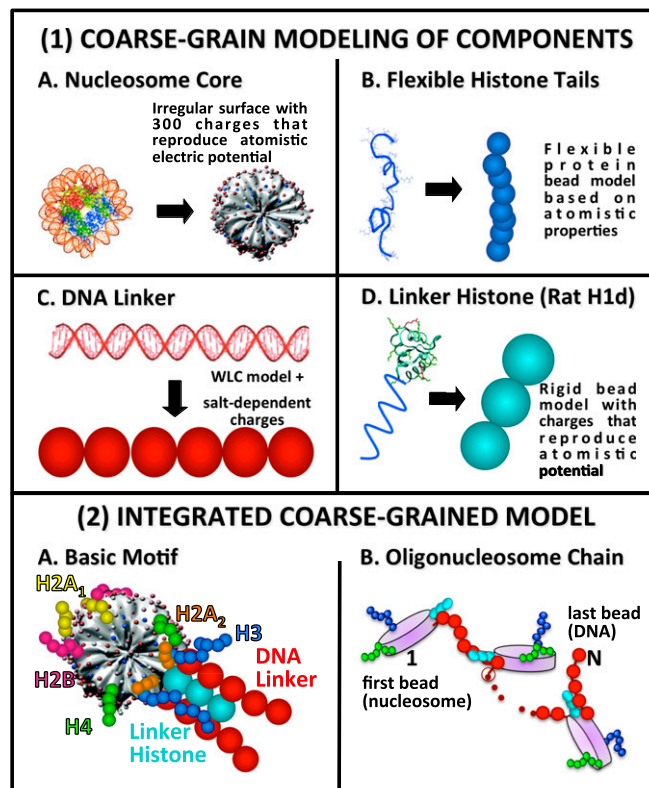


Fig. S1. Summary of the mesoscale modeling. (1) Brief description of the separate coarse-grained treatments used for A. The nucleosome core (with DNA wrapped around) representation results in a reduction from $\sim 25,000$ atoms to 300 charges, (B) the histone tails (which consist of $\sim 4,000$ atoms) are interpreted with 50 tail beads, (C) the linker DNA description models ~ 800 atoms per DNA twist with one bead per twist, and (D) the globular and C-terminal domains of LH (which consists of $\sim 2,500$ atoms) are modeled with three beads. (2) Representation of the integrated coarse-grained model including A. Basic coarse-grained repeating motif, and (B) assembly of the oligonucleosome chain where the nucleosome (in purple, with selected tails shown in blue and green and the LH in turquoise) and linker DNA beads (in red) are numbered/indexed in the direction of the oligonucleosome chain starting from $i = 1$ for the first nucleosome core to N for the last linker DNA bead.

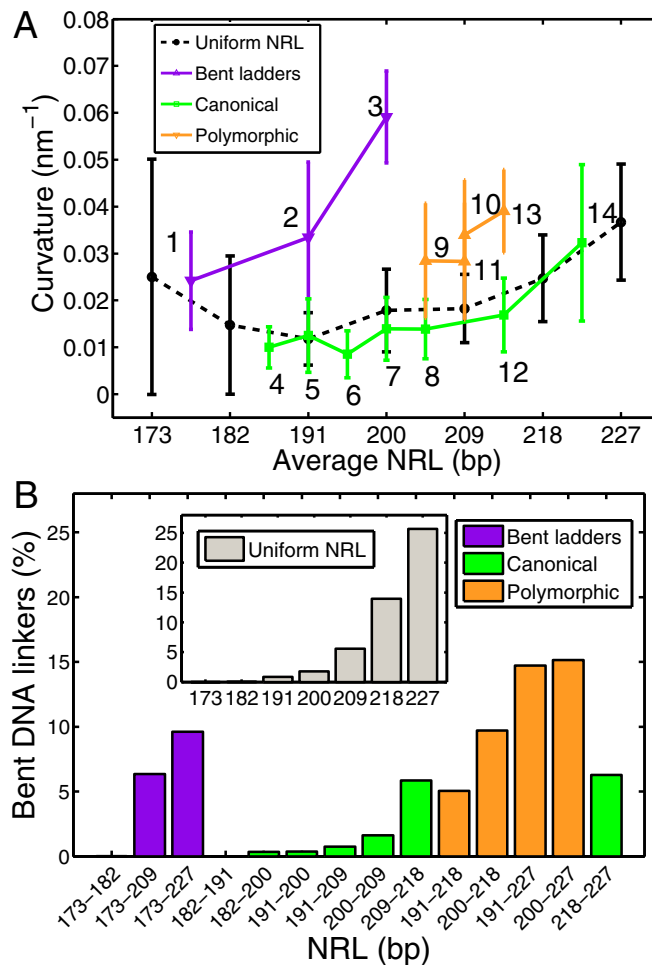


Fig. S2. Fiber axis bending and DNA linker bending. (A) Curvature (ensemble average and SD) of the fiber axis versus the average NRL across the fiber; and (B) fraction of DNA linkers in the ensemble with a bending angle $>60^\circ$; we use this threshold as it lies between the expected average bending angles for ideal zigzag (0°) and solenoid (120°) fibers. As detailed in Table 1, data for nonuniform NRL fibers have been separated into three types according to their structural behavior: bent ladder (purple), canonical (green), and polymorphic (orange). The black dashed curves (A) and gray bars (B, *Inset*) show the results obtained with our previous uniform NRL simulations (14). The numbers next to the nonuniform NRL data points (A) indicate the fiber number as given in Table 1.

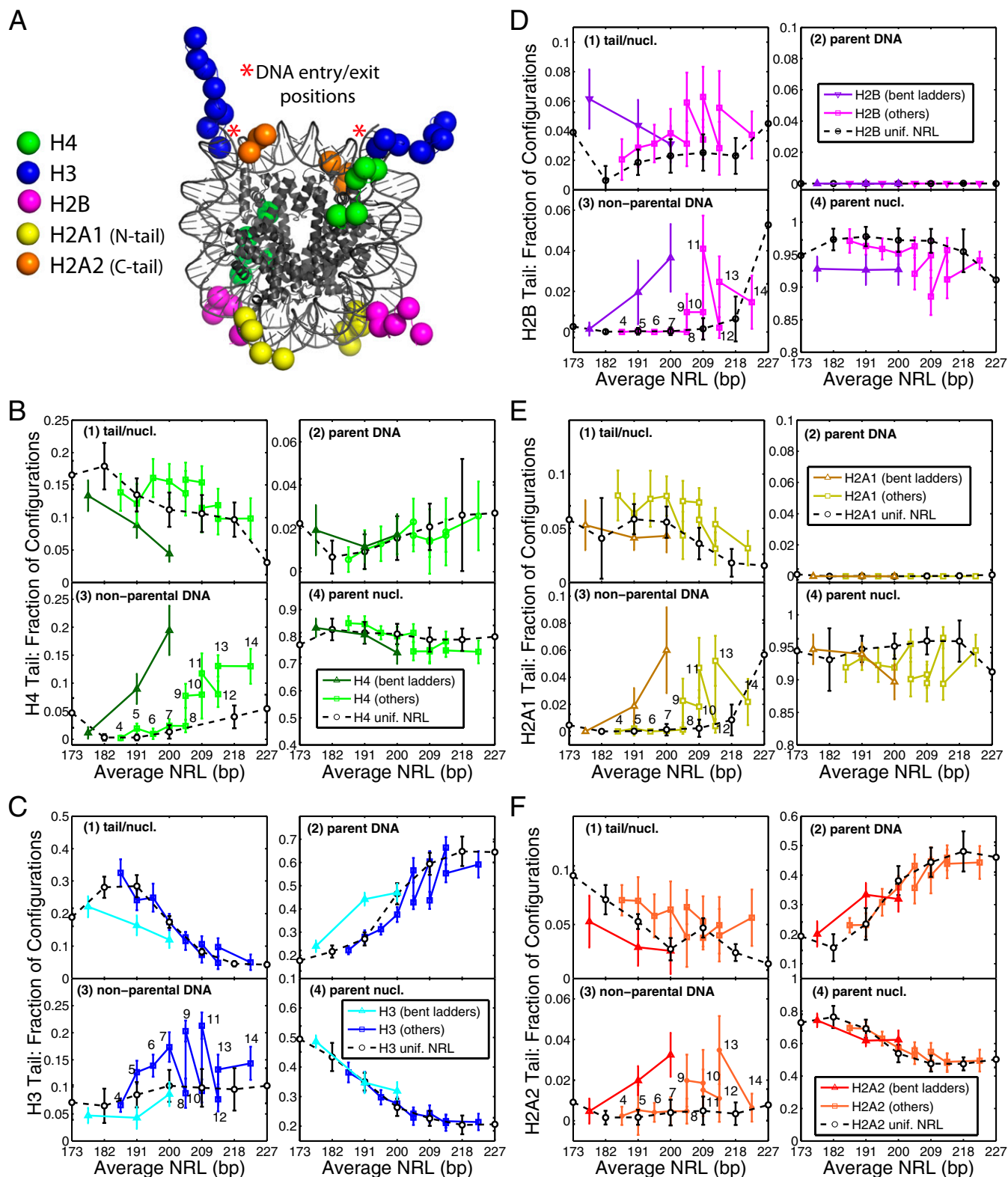
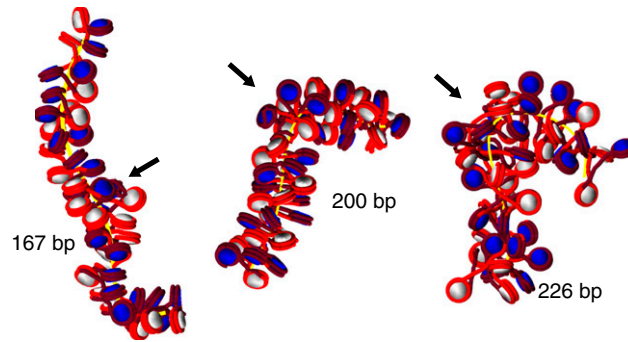


Fig. S6. (A) Illustration of the histone tails within the nucleosome particle and (B–F) frequency analyses (ensemble averages and SD) of different tail interactions in nonuniform NRL fibers [triangles represent bent ladders and squares represent “others” (canonical and polymorphic)] versus uniform NRL fibers (black circles and dashed lines). B–F are for H4, H3, H2B, H2A₁ (N-terminal), and H2A₂ (C-terminal) tails, respectively. In B–F the curves measure the interactions of tails (1) between nucleosomes, (2) with parent linker DNA, (3) with nonparent DNA linkers, and (4) with parent nucleosome. H2A₁ and H2A₂ denote N termini and C termini, respectively, of the H2A tails.

A Nonuniform NRL from single nucleosome removal



B Nonuniform NRL from double nucleosome removal

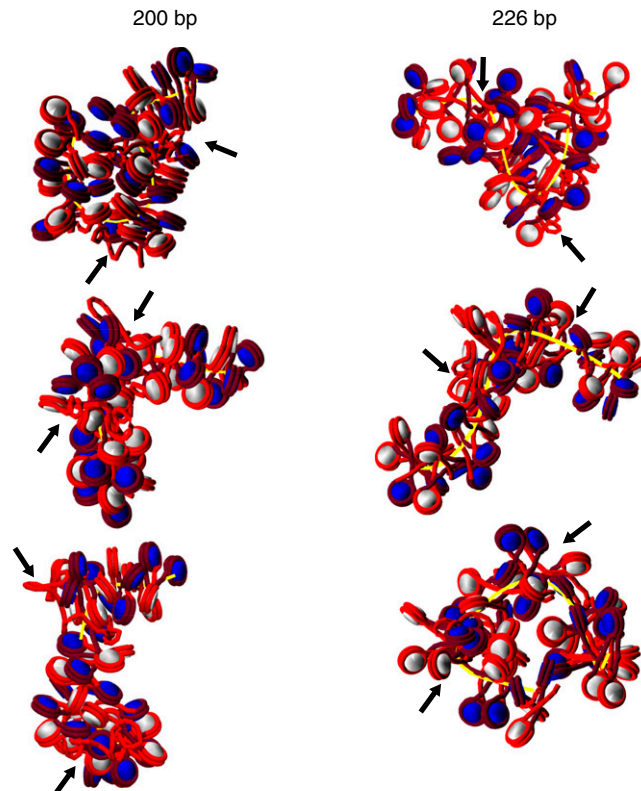


Fig. S7. Representative snapshots obtained after removing one or two nucleosomes from 48-core oligonucleosomes with uniform NRLs. (A) Removal of one nucleosome at position 24 (center nucleosome) in 167-, 200-, and 226-bp fibers causes higher bending of the fiber axis as the NRL increases. (B) Removal of two nucleosomes at positions 16 and 32 in 200- and 226-bp fibers causes a wide range of chromatin fiber forms. Sites where nucleosomes have been removed are indicated by arrows. After nucleosome removal, the length of the linker DNA connecting the two immediate neighbors of the removed core increases by 147 bp; thus, the resulting chromatin fibers have nonuniform NRLs with a large NRL variation.

