Sensitive effect of linker histone binding mode and subtype on chromatin condensation

Ognjen Perišić[1], Stephanie Portillo-Ledesma[1], Tamar Schlick[1,2,3,*]

[1]Department of Chemistry, New York University, 1001 Silver, 100 Washington Square East, New York, New York 10003, United States

[2]Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012, United States

[3]New York University ECNU - Center for Computational Chemistry at NYU Shanghai, 3663 North Zhongshan Road, Shanghai, 200062, China

*To whom correspondence should be addressed: schlick@nyu.edu

Supplementary material

March 19, 2019

Chromatin mesoscale model

DNA: The linker DNA is treated as a modified worm-like chain of n_b discrete spherical beads, (1, 2) with parameters developed using Stigter's procedure (3). Each DNA bead connects two DNA segments, each representing approximately 9 base pairs, with an inter-bead equilibrium length (l_0) of 3 nm (1). Each bead is assigned an excluded volume through the Lennard–Jones potential to prevent a possible overlap between DNA beads and other components of the chromatin array. This approach significantly reduces the number of degrees of freedom (from around 800 atoms to approximately 1 bead per DNA twist). The dynamics of DNA chains are governed by the internal force field comprising of stretching, bending, and twisting energy terms as described in Ref. (4). Within the oligonucleosome chain, each linker DNA bead and nucleosome is allowed to twist about the DNA axis. This is implemented by assigning local coordinate systems to all DNA linker beads and nucleosome cores. The coordinate system of each chain component i is specified by three orthonormal unit vectors $\{a_i, b_i, c_i\}$, where $c_i = a_i \times b_i$. For each nucleosome core i, three additional coordinate systems are defined to describe the DNA bending and twisting at their points of attachment to the nucleosome: $\{a_i^{DNA}, \ b_i^{DNA}, \ c_i^{DNA}\}$ represents the direction from the attachment point of the exiting linker DNA to the center of the i+1 DNA bead; $\{a_i^{+}, \ b_i^{+}, \ c_i^{+}\}$ represents the local tangent on the nucleosome core at the point of attachment of the exiting linker DNA; and $\{a_i^{-}, \ b_i^{-}, \ c_i^{-}\}$ represents the tangent corresponding to the entering linker DNA. To transform the coordinate system of one linker DNA to that of the next (or to that of the entering point of attachment to the core) along the oligonucleosome chain (i.e., $\{a_i, b_i, c_i\} \rightarrow \{a_{i+1}, b_{i+1}, c_{i+1}\}$), we define the Euler angles α_i, β_i, and γ_i as follows:

$$\alpha_i = \begin{cases} cos^{-1}\left(\frac{a_i \cdot a_{i+1}}{\sin(\beta_i)}\right) & \text{if } a_{i+1} \cdot c_i > 0 \\ -cos^{-1}\left(\frac{a_i \cdot a_{i+1}}{\sin(\beta_i)}\right) & \text{if } a_{i+1} \cdot c_i < 0 \end{cases} \quad \text{(S1)}$$

$$\beta_i = cos^{-1}(a_i \cdot a_{i+1}) \quad \text{(S2)}$$

$$\gamma_i = \begin{cases} cos^{-1}\left(\frac{b_i \cdot b_{i+1} + c_i \cdot c_{i+1}}{1 + a_i \cdot a_{i+1}}\right) - \alpha_i & \text{if } \frac{b_i \cdot b_{i+1} + c_i \cdot c_{i+1}}{1 + a_i \cdot a_{i+1}} > 0 \\ -cos^{-1}\left(\frac{b_i \cdot b_{i+1} + c_i \cdot c_{i+1}}{1 + a_i \cdot a_{i+1}}\right) - \alpha_i & \text{if } \frac{b_i \cdot b_{i+1} + c_i \cdot c_{i+1}}{1 + a_i \cdot a_{i+1}} < 0 \end{cases} \quad \text{(S3)}$$

To implement the correct non-integral twist for each DNA segment, we first estimate the actual number of turns, τ_{ns}, that each DNA linker should make according to its length by dividing the linker length over the number of base pairs per turn for DNA in chromatin (l_r); that is, $\tau_{ns} = l_{ns}^{DNA}/l_r$. Here, we use l_r = 10.3 bp/turn for DNA in chromatin, based on experimental observations (5, 6). Note that a range of 10.2–10.5 bp/turn has been reported for DNA of chromatin, which is different from the twist for nucleosome-free DNA. The resulting τ_{ns} values are non-integral for all the NRL studied, except for NRL=209 bp, where the linker length corresponds to six full helical turns. When the length of the linker DNA corresponds to an integral number of turns, the average mean twist of that DNA section is exactly zero. However, a nonintegral number of turns shifts the average twist of the DNA linker involved. Thus, to model the different DNA linker lengths, we incorporate the appropriate equilibrium twist per DNA linker segment to accommodate nonintegral numbers of DNA turns. In practice, we accomplish this by including a penalty term in the total torsional energy of the bead segments. This torsional energy is summed over the DNA beads

$$E_t = \frac{s}{2l_o}\sum_{i=1}^{N-1}(\alpha_i + \gamma_i - \varphi_{ns})^2, \quad \text{(S4)}$$

where *s* is the torsional rigidity of DNA, *N* is the number of beads in the oligonucleosome chain, φ_{ns} is the twist deviation penalty term per segment, and α_i and γ_i are two of the Euler angles defined above. The sum $\alpha_i + \gamma_i \in [-\pi, \pi]$ gives the linker DNA twist at each bead location. Thus, subtracting φ_{ns} from this sum of angles shifts the average linker DNA twist per segment from zero to the required value. The detailed description of the twist penalty term per segment calculation is given in (1).

Nucleosome: The nucleosome particle with wrapped DNA, but without flexible histone tails is represented by 300 rigidly positioned charged beads that robustly reproduce the electrostatic field of the nucleosome at physiological monovalent salt concentrations (7–9). The irregular shape for the nucleosome and the 300 surface charges are derived by our discrete surface charge optimization (DiSCO) algorithm (8), which approximates the electric field of the atomistic nucleosome (PDB 1KX5) by placing pseudocharges along the surface of the complex as a function of monovalent salt (10, 11). A full discussion is given in ref (1).

Histone tails: There are 10 histone tails per nucleosome: tails belonging to N-termini of H2A (denoted H2A1), H2B, H3, and H4 histones, plus C-termini tails of H2A histones (denoted as

H2A2). Each tail is modeled as a flexible chain of charged spherical beads, with each bead representing five adjacent amino acids. Each nucleosome has two of H2A1, H2A2, H2B, H3, and H4 histone tails, represented using 4, 3, 5, 8, and 5 beads, respectively, for a total of 50 tail beads per nucleosome (equivalent to the approximately 250 histone tail residues per each nucleosome). The center of each macro bead is placed at the C_β atom of the third amino acid. Each bead is assigned a charge equal to the sum of the charges on the five amino acids it represents, and scaled to adjust charges to the monovalent salt environment (for salt concentrations of 0.01, 0.15, and 0.2 M, the scaling factors for the bead charges are 0.75, 1.12, and 1.2, respectively). The excluded volume of each tail bead is modeled through a Lennard–Jones potential with fixed parameters.

The histone tails are rigidly fixed to their idealized position on the nucleosome surface by stiff springs between the core and the first tail bead. For tail beads not attached to the core, the stretching and bending harmonic potentials between beads and bond angles between three consecutive beads are tuned to reproduce configurational properties of the full-atom histone tail models used in Brownian dynamics simulations (11–13).

Linker histones H1C and H1E: There are two linker histone types represented in our model, H1E (based on rat H1.4) and H1C (based on mouse H1.2). H1E is the first flexible LH model implemented with our system (14). The rigid GH of H1E was modeled on the basis of the rat H1.4 via 6 beads (14–16). The flexible H1E CTD tail was modeled following the same approach as histone tails. The 111 amino acids of H1.4 CTD are modeled via 22 beads, with one bead per 5-amino acids. The fully extended CTD chain is 25nm long, but the starting configuration is compressed to 10 nm (14). Each CTD bead is connected to adjacent sequential neighbors by identical harmonic and bending potentials, see Fig. 1.

Energy terms: Energy terms include bend, stretch, and twist terms for linker DNA and histone tail beads, a Debye−Hückel electrostatic interaction term for all charged segments, and excluded volume terms in the form of a modified Lennard-Jones potential for all beads (1). Coordinates are computed and propagated via a local coordinate frame: Euler vectors are used to track the pitch, roll, and twist of each nucleosome, and then to calculate the corresponding linker DNA and tail coordinates, as detailed previously (1).

Chromatin electrostatics: Physiological salt conditions with monovalent and divalent cations are indispensable for compacting chromatin by screening the highly charged chromatin components

(e.g., nucleosomal and linker DNA). We treat the counterions implicitly using mean field theories. Specifically, our DiSCO algorithm (8, 9) parameterizes the screening potential from the Poisson–Boltzmann equation (PBE) using a Debye–Hückel approximation with salt-dependent effective charges, obtained by minimizing the difference between the electric fields from PBE and the (linear) Debye–Hückel approximation using our efficient TNPACK (truncated Newton) optimization package (17). Thus, DiSCO is used to evaluate the effective charges on the nucleosome core, LHs, and histone tails; the effective charges for DNA beads are obtained using an analytical method by Stigter (3). For the nucleosome core, we use 300 effective charges uniformly distributed across the nucleosome surface; this produces a robust approximation, with < 10% error in the DH approximation over a large range of monovalent salt concentrations (8, 9). The DiSCO approach has been implemented for monovalent ions and assumes that the screening potential is independent of chromatin conformation.

Chromatin energy function: The total potential energy is expressed as the sum of stretching, bending, and torsional components of linker DNA, stretching of histone tails, intramolecular bending of the histone tails, total electrostatic energy, and excluded volume terms (11):

$$E = E_S + E_B + E_T + E_{tS} + E_{tB} + E_C + E_V. \quad \text{(S5)}$$

The first three terms denote stretching,

$$E_s = \frac{h}{2}\sum_{i=1}^{N-1}(l_i - l_o)^2 \ . \quad \text{(S6)}$$

bending

$$E_B = \frac{g}{2}\left[\sum_{i=1}^{N}(\beta_i)^2 + \sum_{i=i\in I_c}^{N}(\beta_i^+)^2\right]. \quad \text{(S7)}$$

and torsional energy of linker DNA (equation 1). Here, h and g denote the stretching and bending rigidities of DNA, l_i denotes the separation between the DNA beads, and I_C denotes a nucleosome particle within the oligonucleosome chain. As mentioned above, N is the total number of beads in the chromatin chain, β_i and β_{i+1} are bending angles, and l_0 is the equilibrium separation distance between beads of relaxed DNA (3 nm).

The fourth term, E_{tS}, represents the total stretching energy of the histone tails, composed of two terms: stretching of tail beads and stretching of the histone tail bead from its assigned attachment site, as given by:

$$E_{tS} = \sum_{i\in I_c}^{N} \sum_{j=1}^{N_T} \sum_{k=1}^{N_{bj}-1} \frac{k_{b_{jk}}}{2} \left(l_{ijk} - l_{jk0}\right)^2 + \frac{h_{tc}}{2} \sum_{i\in I_c}^{N} \sum_{j=1}^{N_T} \left|t_{ij} - t_{ij0}\right|^2. \quad \text{(S8)}$$

Here, $N_T = 10N_C$ is the total number of histone tails, N_{bj} is the number of beads in the *j-th* tail, k_{bjk} is the stretching constant of the bond between the *k*th and (*k+1*)th beads of the *j*th histone tail, and 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, t_{ij} is the position vector of the first tail bead in the coordinate system of its parent nucleosome, and t_{ij0} is the ideal position vector in the crystal configuration.

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

$$E_{tB} = \sum_{i\in I_c}^{N} \sum_{j=1}^{N_T} \sum_{k=1}^{N_{bj}-2} \frac{k_{\theta_{jk}}}{2} \left(\theta_{ijk} - \theta_{jk0}\right)^2, \quad \text{(S9)}$$

where θ_{ijk} and θ_{ij0} 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 sixth term, E_C, represents the total electrostatic interaction energy of the oligonucleosome. All these interactions are modeled using the Debye–Hückel potential that accounts for salt screening:

$$E_c = \sum_i \sum_{i\neq j} \frac{q_i q_j}{4\pi\varepsilon\varepsilon_o r_{ij}} exp\left(-kr_{ij}\right), \quad \text{(S10)}$$

where q_i and q_j are the 'effective' charges separated by a distance r_{ij} in a medium with a dielectric constant of κ and an inverse Debye length of 1/κ, ε_0 is the electric permittivity of vacuum, and ε is the dielectric constant (set to 80). As described above, the salt-dependent effective charges are calculated using DiSCO (8, 9) by matching the electric field from the PBE (solved using the DelPhi software) to the field parameterized using the Debye–Hückel form.

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:

$$E_V = \sum_i \sum_{i\neq j} k_{ij} \left[\left(\frac{\sigma^{ij}}{r^{ij}}\right)^{12} - \left(\frac{\sigma^{ij}}{r^{ij}}\right)^{6}\right], \quad \text{(S11)}$$

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 were all taken from relevant models of the components as described fully in Ref. (4).

Monte Carlo algorithm: Equilibrium chromatin array configurations were sampled using efficient Monte Carlo local (translation and rotation) and global moves with our coarse grained chromatin model. The local translation and rotation moves choose a randomly oriented axis passing through a randomly picked linker bead or nucleosome core. The translation move shifts chosen component along the above chosen random axis by a distance sampled from a uniform distribution. With the rotation move, the chosen polymer segment (DNA bead or nucleosome) is rotated about the random axis by a uniformly sampled angle. The global move randomly choses one of the polymer links (DNA bead or nucleosome) and randomly chooses an axis through that component and rotates a shorter segment of the polymer fiber around that axis. The global and local moves are accepted using the standard Metropolis energy criterion (see details in (1)). The total enthalpy of the chromatin fiber is composed of bonded (stretching, bending and twisting energy) and nonbonded (electrostatic and van der Waals energy) terms.

The tails are regrown using the configurational bias MC method (18, 19). In this approach we randomly select a tail to be regrown and then apply the Rosenbluth-Rosenbluth (20) scheme to regrow it. The tail configuration with the lowest energy is accepted. The volume enclosed within the nucleosomal surface is discretized to prevent histone tail beads from penetrating the nucleosome core during tail regrowth, and any insertion attempts that place the tail beads within this volume are rejected automatically.

Both linker histones are folded using a Monte Carlo scheme in which a randomly chosen bead is stochastically moved along the three spatial dimensions (x, y and z) using uniform distributions (14). In nucleosomes with two LH, both LHs are regrown independently of each other using the same Monte Carlo scheme applied to regrow single LH per nucleosome.

Initial configurations are ideal zigzag configurations constructed by rotating nucleosomes around the fiber axis and moving them downward. Each starting configuration contains coordinates and orientation matrices for each DNA bead and each nucleosome. Initial configurations for tails and LHs are loaded from the corresponding configurational files. Each nucleosome has the same configurations for tails and LH.

The fiber systems were sampled with 35 to 40 million MC steps. The analysis was performed on the last 5 to 10 million steps. Configurations were sampled every 50,000 MC steps.

Calculation of fiber packing ratio: The packing ratio is calculated from the length of fiber axis which is expressed as a three-dimensional parametric curve

$$\mathbf{r}^{ax}(i) = (r_1^{ax}(i), r_2^{ax}(i), r_3^{ax}(i)), \tag{S12}$$

where each $r_j^{ax}(i)$ (j = 1, 3 or 3) is a nonlinear function fitted to the positions of nucleosome centers to one spatial dimension (x, y or z). The functions $r_1^{ax}(i)$ are the variable degree (M) polynomials with coefficients estimated using the least squares procedure. The polynomial degrees (M) was chosen to reduce the standard deviation of the fiber width (M = 3 for 24 core arrays and M = 5 for 100 core arrays). The fiber widths are calculated as the twice the average distance between nucleosome centers and fiber axis (see (1) for details).

Persistence length calculation: The persistence length L_p is calculated by fitting an exponential to the angle defined by two unit tangent vectors ($u(s) = \partial \mathbf{r}^{ax}(s)/\delta s$) of the fiber axis parametric curve $\mathbf{r}^{ax}(i)$ (21, 22):

$$\langle u(s) \cdot u(s') \rangle = exp\big(-|s - s'|/L_p\big), \tag{S13}$$

where $u(s)$ and $u(s')$ are a tangent vector at the beginning of the curve and a tangent vector that corresponds to a highest bending (smallest value of the dot product). $|s - s'|$ is the contour length of the whole fiber.

Contact patterns calculation: The local chromatin topology is described by internucleosome patterns (1, 23) which are extracted from the interaction intensity matrices that measure the fraction of MC configurations that each nucleosome interacts with other nucleosomes in the fiber. The nucleosomes are considered to "interact" if their tail beads or core beads come within 1.8 nm of other nucleosome's core beads. The interaction matrices are decomposed into one-dimensional plots that depict the magnitude of i, $i + k$ interactions, as

$$I(k) = \frac{\sum_{i=1}^{N_C} I'(i,i\pm k)}{\sum_{j=1}^{N_C} I(j)}. \tag{S14}$$

TABLES – Supplementary material

2 LH setup		2 H1C			2 H1C			2 H1C			2 H1C			2 H1C			2 H1C		
NRL (bp)	ρ	–20º/on-dyad			–20º/+20º			on-dyad/–20º			on-dyad/+20º			+20º/–20º			+20º/on-dyad		
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
200	1.3	4.7	143.8	175.6	4.2	134.7	63.5	4.1	131.4	146.3	4.4	142.6	91.9	3.7	138.9	79.2	4.4	138.7	180.4
209	1.3	3.9	133.8	59.6	4.5	140.5	60.7	3.8	134.0	88.9	4.4	132.0	142.9	4.4	158.2	47.4	4.4	141.1	52.9
2 LH setup		**H1C/H1E**			**H1C/H1E**			**H1C/H1E**			**H1C/H1E**			**H1C/H1E**			**H1C/H1E**		
NRL (bp)	ρ	–20º/on-dyad			–20º/+20º			on-dyad/–20º			on-dyad/+20º			+20º/–20º			+20º/on-dyad		
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
200	1.3	4.5	136.1	235.2	4.7	141.8	93.0	4.1	137.8	92.2	4.7	146.6	156.9	4.1	133.5	100.1	4.8	147.2	57.2
209	1.3	4.0	130.9	59.9	4.3	136.1	59.4	4.0	129.1	95.9	4.0	140.1	79.0	4.5	158.1	53.9	4.2	142.2	83.9
2 LH setup		**2 H1C**			**2 H1C**			**2 H1C**			**2 H1C**			**2 H1C**			**2 H1C**		
NRL (bp)	ρ	–20º/on-dyad			–20º/+20º			on-dyad/–20º			on-dyad/+20º			+20º/–20º			+20º/on-dyad		
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
200	1.6	4.0	129.6	68.7	4.4	135.4	146.2	3.9	127.5	134.8	5.0	146.0	63.4	3.8	129.2	66.2	4.9	145.1	156.8
209	1.6	4.2	127.3	74.9	4.0	128.6	166.6	3.8	125.6	85.3	4.6	139.4	72.4	4.4	137.8	69.9	4.2	138.3	74.2
2 LH setup		**H1C+H1E**			**H1C+H1E**			**H1C+H1E**			**H1C+H1E**			**H1C+H1E**			**H1C+H1E**		
NRL (bp)	ρ	–20º/on-dyad			–20º/+20º			on-dyad/–20º			on-dyad/+20º			+20º/–20º			+20º/on-dyad		
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
200	1.6	4.1	131.0	229.3	4.0	129.2	156.9	3.8	130.0	79.9	4.5	138.6	102.4	4.0	132.3	90.5	4.9	145.8	179.1
209	1.6	3.7	141.9	64.5	4.4	138.6	54.1	3.8	125.1	92.2	4.3	131.2	109.0	4.4	137.2	59.9	4.7	139.9	108.2

Table S1. Fiber properties (packing densities in nuc/11 nm, sedimentation coefficients in S, persistence lengths in nm) for 100-nucleosome arrays saturated with LH densities 1.3 and 1.6 LH per nucleosome. For each setup we show which and how two LHs are placed in one nucleosome. Nucleosomes with 1 LH are saturated with H1E and follow orientation of the first LH in 2 LH pair. For example, in top left entry with ρ = 1.3, NRL = 200 bp, two LHs are organized as –20º/on-dyad. That means that nucleosome with 2 LHs have 2 H1C bound –20º/on-dyad, and nucleosomes with 1 LH have H1E bound –20º. Columns are as follows: a – packing densities; b – sedimentation coefficients; c – persistence lengths. The values are color coded by nucleosome packing ranges: blue - low packing (< 4.5 nuc/11 nm), yellow - medium (4.5 > nuc/11 nm < 6), and red - high (> 6 nuc/11 nm).

NRL (bp)	H1C –20°			H1C ON-DYAD			H1C +20°		
	Packing (nuc/11 nm)	Sedimentation (S)	Persistence length (nm)	packing (nuc/11 nm)	Sedimentation (S)	Persistence length (nm)	packing (nuc/11 nm)	Sedimentation (S)	Persistence length (nm)
182	4.2	74.8	30.5	5.1	81.8	88.6	4.1	78.6	47.1
191	4.2	73.1	138.4	4.6	76.4	88.7	4.3	72.8	17.4
200	4.2	72.5	29.9	4.5	72.3	41.5	4.7	73.7	81.6
209	4.4	69.9	24.3	4.2	70.9	29.7	4.3	69.1	32.5
218	4.6	70.8	37.1	4.1	67.5	22.2	4.9	71.3	59.3
NRL (bp)	H1E –20°			H1E ON-DYAD			H1E +20°		
	packing (nuc/11 nm)	Sedimentation (S)	Persistence length (nm)	packing (nuc/11 nm)	Sedimentation (S)	Persistence length (nm)	packing (nuc/11 nm)	Sedimentation (S)	Persistence length (nm)
182	4.8	81.5	27.3	3.8	70.8	40.1	3.6	75.1	36.9
191	5.8	83.3	214.4	4.7	75.4	159.0	4.0	74.2	102.5
200	6.5	81.8	73.4	5.2	74.9	213.2	4.4	72.4	97.6
209	5.2	73.3	31.0	4.3	69.5	58.0	4.3	74.8	43.9
218	7.3	76.7	39.1	4.9	71.2	37.8	5.8	72.5	150.5

Table S2a. Fiber properties (packing densities, sedimentation coefficients, persistence lengths) for 24-nucleosome arrays fully saturated with 1 LH (H1C or H1E) per nucleosome. The values are color coded by nucleosome packing ranges: blue - low packing (< 4.5 nuc/11 nm), yellow – medium (4.5 > nuc/11 nm < 6), and red - high (> 6 nuc/11 nm).

NRL (bp)	H1C –20°			H1C on-dyad			H1C +20°		
	Packing (nuc/11 nm)	Sedimentation (S)	Persistence length (nm)	Packing (nuc/11 nm)	Sedimentation (S)	Persistence length (nm)	Packing (nuc/11 nm)	Sedimentation (S)	Persistence length (nm)
182	3.5	69.9	143.6	3.7	71.5	124.4	4.0	73.2	97.3
191	4.0	72.1	83.9	4.0	70.2	71.1	3.6	72.3	143.8
200	4.0	70.4	60.1	3.7	71.0	38.0	4.1	72.0	36.7
209	3.8	67.4	37.5	3.6	67.6	26.3	3.3	73.7	26.5
218	3.6	68.7	26.9	3.6	69.2	26.2	3.9	67.8	41.4
NRL (bp)	H1E –20°			H1E on-dyad			H1E +20°		
	Packing (nuc/11 nm)	Sedimentation (S)	Persistence length (nm)	Packing (nuc/11 nm)	Sedimentation (S)	Persistence length (nm)	Packing (nuc/11 nm)	Sedimentation (S)	Persistence length (nm)
182	4.6	79.5	116.2	3.6	71.5	86.2	3.7	75.1	27.7
191	4.3	73.3	165.3	4.2	72.2	50.2	3.6	69.6	47.2
200	4.9	76.9	55.6	4.4	71.3	25.7	4.2	71.1	47.6
209	4.5	72.2	24.9	3.8	70.0	21.5	4.6	69.8	20.6
218	3.6	68.6	24.4	3.4	66.5	22.1	4.1	66.6	48.6

Table S2b. Fiber properties (packing densities, sedimentation coefficients, persistence lengths) for 24-nucleosome arrays saturated with 0.5 LH (H1C or H1E) per nucleosomeThe values are color coded by nucleosome packing ranges: blue - low packing (< 4.5 nuc/11 nm), yellow – medium (4.5 > nuc/11 nm < 6), and red - high (> 6 nuc/11 nm).

	2 H1C				2 H1C				2 H1C				2 H1C				2 H1C				2 H1C			
NRL (bp)	–20º/on-dyad				–20º/+20º				on-dyad/–20º				on-dyad/+20º				+20º/–20º				+20º/on-dyad			
	a	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**
200	-98.6	-8.6	-5.2	6.7	-93.1	-8.7	-4.4	6.6	-107.9	-4.5	-3.8	6.2	-105.0	-9.0	-4.7	7.3	-77.0	-17.0	-3.6	4.7	-76.1	-25.4	-5.0	6.8
209	-106.0	-6.8	-3.5	7.7	-99.8	-9.1	-3.5	7.5	-113.4	-5.3	-3.0	7.5	-111.8	-10.5	-3.6	8.6	-89.3	-15.7	-3.5	6.2	-94.4	-17.3	-3.8	7.5

	H1C/H1E				H1C/H1E				H1C/H1E				H1C/H1E				H1C/H1E				H1C/H1E			
NRL (bp)	–20º/on-dyad				–20º/+20º				on-dyad/–20º				on-dyad/+20º				+20º/–20º				+20º/on-dyad			
	a	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**
200	-105.6	-8.0	-4.7	7.4	-98.4	-10.0	-5.0	7.7	-113.5	-4.4	-3.9	6.3	-111.7	-10.9	-4.7	7.5	-80.4	-17.0	-4.0	4.9	-74.6	-27.2	-5.4	6.4
209	-113.0	-6.9	-3.5	8.2	-103.4	-10.1	-3.4	8.3	-119.2	-6.1	-3.1	7.8	-121.3	-8.7	-3.6	8.4	-91.8	-16.9	-3.6	6.7	-92.5	-19.4	-3.7	6.8

Table S3a. Electrostatic energies for 100-nucleosome arrays with 1.3 LH per nucleosome. Nucleosomes with 1 LH are saturated with H1E, and nucleosomes with 2 LH are filled with 2 H1C (top) or H1C/H1E (bottom) per nucleosome. Columns are as follows: a) energies between linker histone and parental DNA strands; b) energies between linker histone and nonparental DNA strands; c) energies between tails and nonparental cores; d) energies between DNA linkers. Colored setups correspond to the colored setups in Table S1.

	2 H1C				2 H1C				2 H1C				2 H1C				2 H1C				2 H1C			
NRL (bp)	–20º/on-dyad				–20º/+20º				on-dyad/–20º				on-dyad/+20º				+20º/–20º				+20º/on-dyad			
	a	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**
200	-104.4	-6.3	-3.8	5.2	-97.6	-8.5	-4.5	6.1	-110.2	-4.2	-3.3	5.0	-109.1	-12.8	-5.1	7.7	-91.1	-12.2	-3.5	5.1	-91.6	-22.5	-5.6	7.2
209	-116.1	-5.7	-3.1	6.7	-99.3	-13.5	-3.3	7.2	-117.4	-5.8	-2.7	6.6	-111.1	-14.4	-3.7	8.0	-97.0	-14.4	-3.4	6.3	-103.3	-18.1	-4.0	7.6

	H1C/H1E				H1C/H1E				H1C/H1E				H1C/H1E				H1C/H1E				H1C/H1E			
NRL (bp)	–20º/on-dyad				–20º/+20º				on-dyad/–20º				on-dyad/+20º				+20/–20º				+20º/on-dyad			
	a	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**	**a**	**b**	**c**	**d**
200	-122.2	-6.3	-4.0	7.1	-103.7	-7.6	-3.9	6.7	-116.7	-4.7	-3.3	5.7	-122.2	-13.9	-4.7	7.8	-91.4	-12.2	-3.9	4.6	-91.1	-21.9	-5.2	6.3
209	-128.7	-6.2	-2.9	7.8	-110.2	-13.1	-3.4	8.1	-125.5	-5.7	-2.6	7.1	-133.3	-12.8	-3.7	8.8	-99.7	-15.2	-3.5	5.8	-105.9	-19.4	-4.3	7.4

Table S3b. Electrostatic energies (in kcal/mol) for 100-nucleosome arrays with 1.6 LH per nucleosome. Nucleosomes with 1 LH are saturated with H1E, and nucleosomes with 2 LH are filled with 2 H1C (top) or H1C/H1E (bottom) per nucleosome. Columns are as follows: a - energies between linker histone and parental DNA strands; b - energies between linker histone and nonparental DNA strands; c - energies between tails and nonparental cores; d - energies between DNA linkers. Colored setups correspond to the colored setups in Table S1.

FIGURES – Supplementary material

Figure S1. Sequence alignment of Rat linker histone H1.4 and Mouse linker histone H1.2.

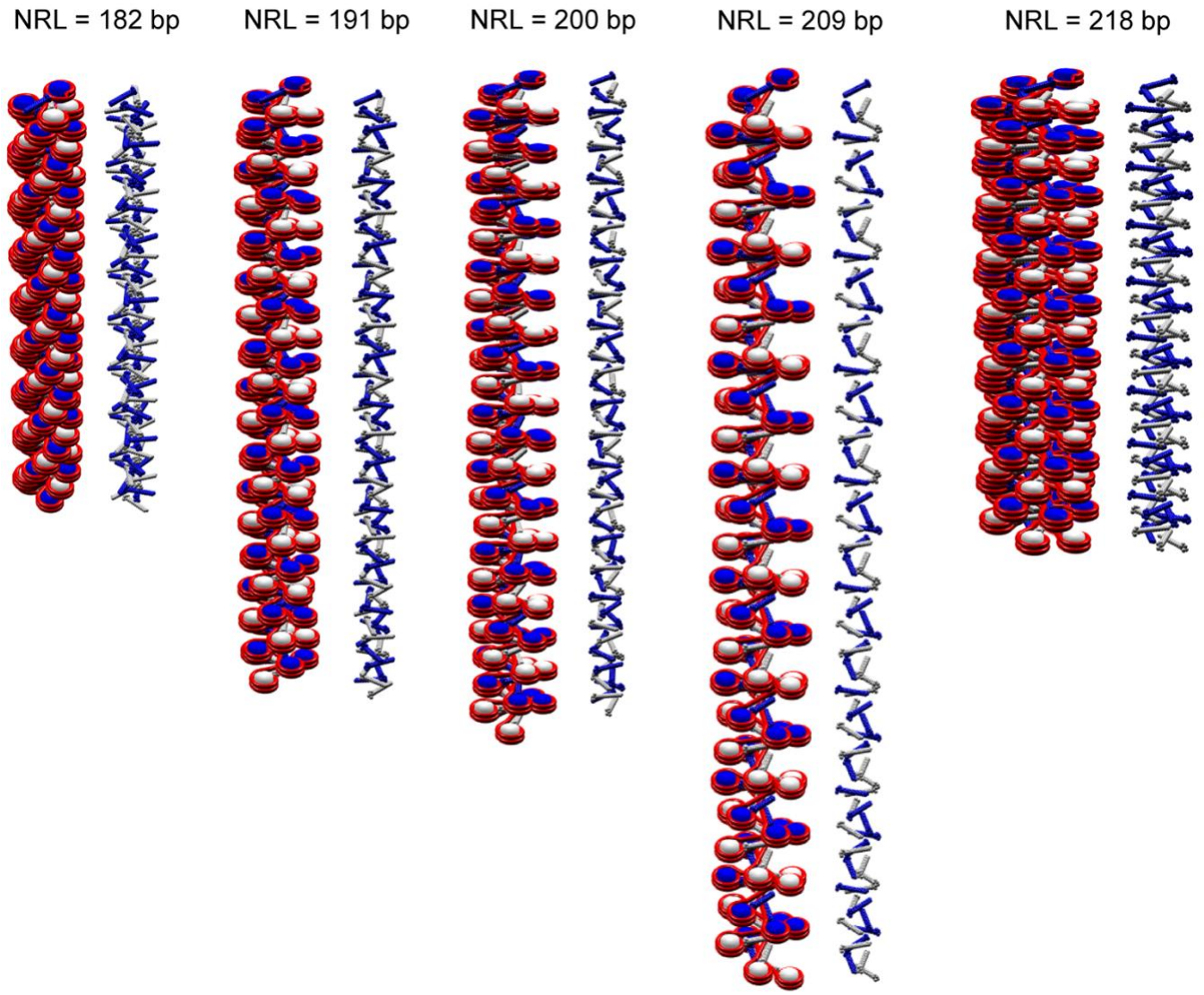

Figure S2. Starting configurations for 100-nucleosome arrays for all covered NRLs (182, 191, 200, 209 and 218 bp). On the left of each pair are full fiber images with LHs, and on the right only the LHs (LH skeletons). Alternative nucleosome are colored white and blue. LHs share the same color as their parental nucleosome core. The depicted fibers have their LHs (H1E) bound –20^{o}. Other fibers share the same nucleosome and DNA starting configurations. The differences are in the orientation and density of LHs.

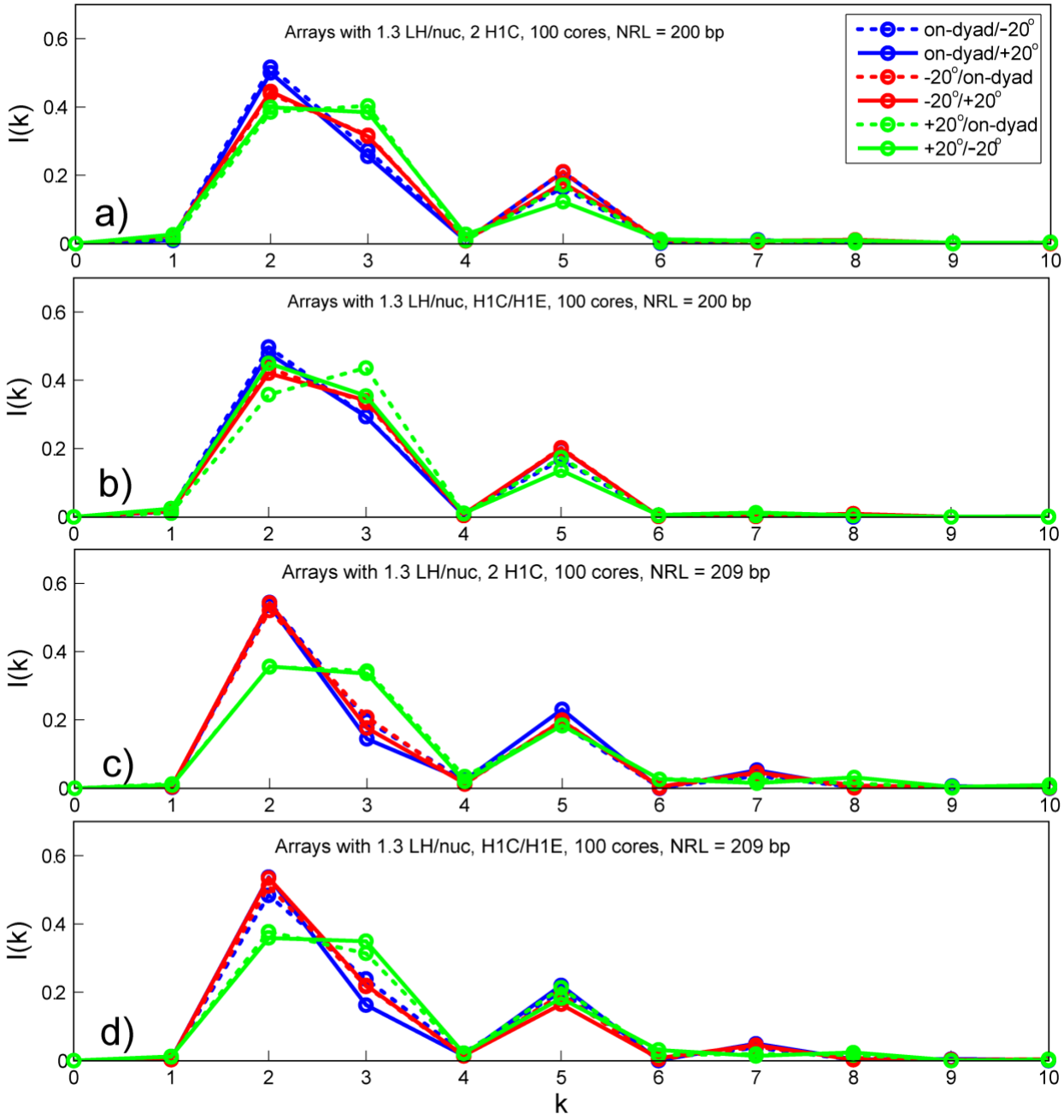

Figure S3. Contact patterns for 100-core arrays with ρ = 1.3. Full and dashed lines correspond to two alternative LH binding modes in respect to the dominant binding mode. For example, in the first subplot the full blue line corresponds to a setup in which the dominant mode is on-dyad, with 100 LHs bound in that mode, and –20° off-dyad as the alternative binding mode (30 LHs bound in that mode). The dashed blue line correspond to the mode with +20° as the alternative binding mode. a) Contact patterns for arrays with NRL = 200 bp, in which nucleosomes with 2 LH have 2 H1C; b) Contact patterns for arrays with NRL = 200 bp, in which nucleosomes with 2 LH have H1C/H1E pair; c) Contact patterns for arrays with NRL = 209 bp, in which nucleosomes with 2 LH have 2 H1C; d) Contact patterns for arrays with NRL = 209 bp, in which nucleosomes with 2 LH have H1C/H1E pair.

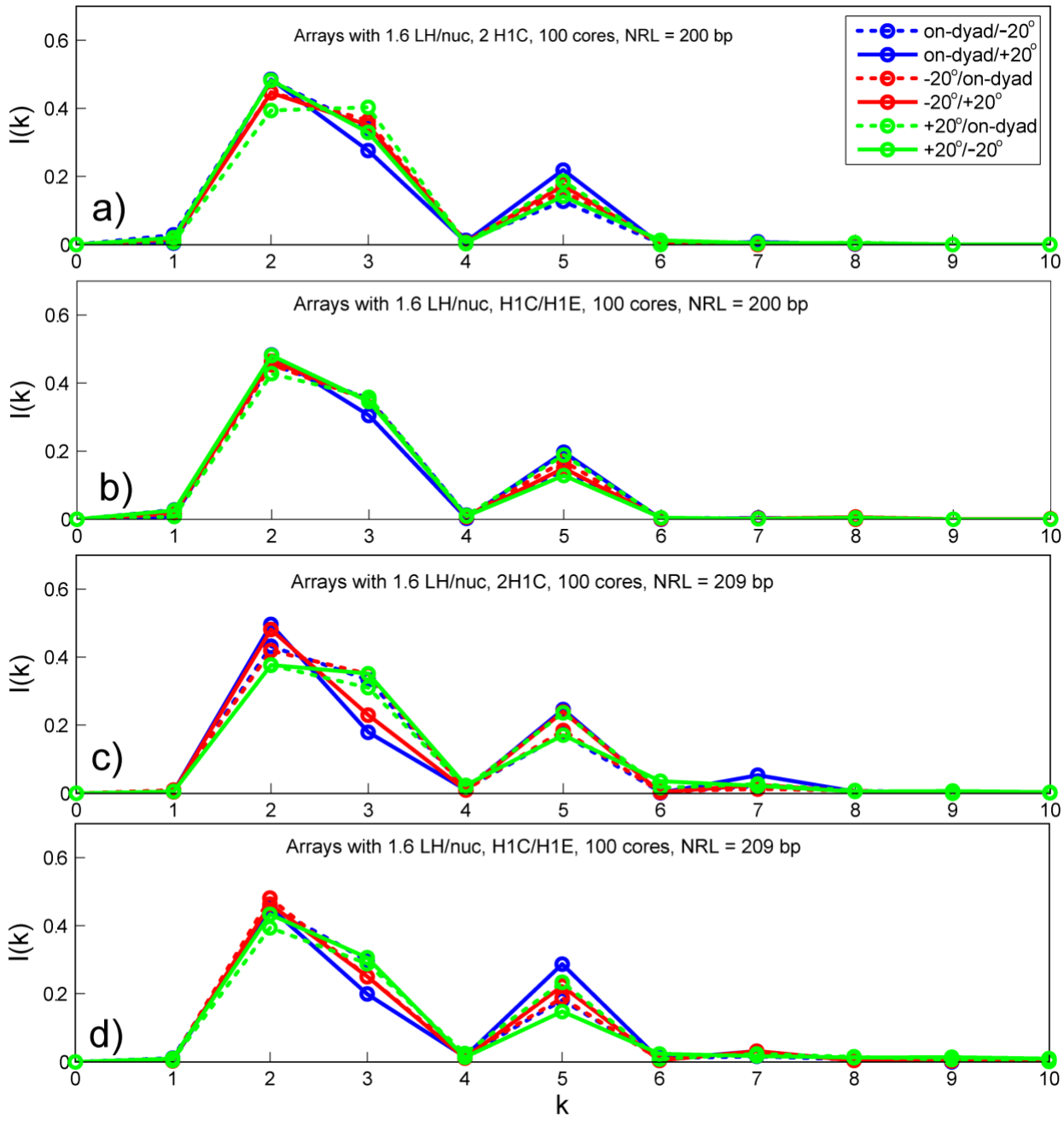

Figure S4. Contact patterns for 100-core arrays with ρ = 1.6. Full and dashed lines correspond to two alternative LH binding modes in respect to the dominant binding mode. For example, in the first subplot the full blue line corresponds to a setup in which the dominant mode is on-dyad, with 100 LHs bound in that mode, and –20° off-dyad as the alternative binding mode (60 LHs bound in that mode). The dashed blue correspond to the setup with +20° as the alternative binding mode. a) Contact patterns for arrays with NRL = 200 bp, in which nucleosomes with 2 LH have 2 H1C; b) Contact patterns for arrays with NRL = 200 bp, in which nucleosomes with 2 LH have H1C/H1E pair; c) Contact patterns for arrays with NRL = 209 bp, in which nucleosomes with 2 LH have 2 H1C; d) Contact patterns for arrays with NRL = 209 bp, in which nucleosomes with 2 LH have H1C/H1E pair.

References

1. Perišić,O., Collepardo-Guevara,R. and Schlick,T. (2010) Modeling Studies of Chromatin Fiber Structure as a Function of DNA Linker Length. *J. Mol. Biol.*, **403**, 777–802.

2. Yamakawa,H. and Yoshizaki,T. (2012) Helical Wormlike Chains in Polymer Solutions. Springer Science & Business Media: New York.

3. Stigter,D. (1977) Interactions of highly charged colloidal cylinders with applications to double-stranded. *Biopolymers*, **16**, 1435–1448.

4. Arya,G. and 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**, 4045–4059.

5. Drew,H.R. and Travers,A.A. (1985) DNA bending and its relation to nucleosome positioning. *J Mol Biol.*, **186**, 773–790.

6. Deng,J., Pan,B. and Sundaralingam,M. (2003) Structure of d(ITITACAC) complexed with distamycin at 1.6 A resolution. *Acta Crystallogr D Biol Crystallogr.*, **59**, 2342–2344.

7. Perišić,O. and Schlick,T. (2016) Computational Strategies to Address Chromatin Structure Problems. *Phys. Biol.*, **13**, 35006.

8. Beard,D.A. and Schlick,T. (2001) Modeling salt-mediated electrostatics of macromolecules: the discrete surface charge optimization algorithm and its application to the nucleosome. *Biopolymers*, **58**, 106–115.

9. Zhang,Q., Beard,D.A., Schlick,T., Q.,Z., D.,B. and T.,S. (2003) Constructing Irregular Surfaces to Enclose Macromolecular Complexes for Mesoscale Modeling Using the Discrete Surface Charge Optimization (DISCO) Algorithm. *J. Comput. Chem.*, **24**, 2063–2074.

10. Davey,C.A., Sargent,D.F., Luger,K., Maeder,A.W. and Richmond,T.J. (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 A resolution. *J. Mol. Biol.*, **319**, 1097–1113.

11. Arya,G., Zhang,Q. and Schlick,T. (2006) Flexible histone tails in a new mesoscopic oligonucleosome model. *Biophys. J.*, **91**, 133–150.

12. Zhang,Q. (2005) Mesoscopic, microscopic, and macroscopic modeling of protein/DNA complexes. *PhD thesis, New York Univ. New York*.

13. Arya,G. and Schlick,T. (2006) Role of histone tails in chromatin folding revealed by a mesoscopic oligonucleosome model. *Proc. Natl. Acad. Sci.*, **103**, 16236–16241.

14. Luque,A., Collepardo-Guevara,R., Grigoryev,S. and Schlick,T. (2014) Dynamic Condensation of Linker Histone C-terminal Domain Regulates Chromatin Structure. *Nucleic Acids Res.*, **42**, 7553–7560.

15. Bharath,M.M., Chandra,N.R. and Rao,M.R. (2002) Prediction of an HMG-box fold in the C-terminal domain of histone H1: insights into its role in DNA condensation. *Proteins*, **49**, 71–81.

16. Bharath,M.M., Chandra,N.R. and Rao,M.R. (2003) Molecular modeling of the chromatosome particle. *Nucleic Acids Res.*, **31**, 4264–4274.

17. Schlick,T. and Fogelson,A. (1992) TNPACK-a truncated Newton minimization package for largescale problems: I. Algorithm and usage. *ACM Trans. Math. Softw*, **18**, 46–70.

18. Frenkel,D., Mooijt,G.C. and Smit,B. (1992) Novel scheme to study structural and thermal properties of continuously deformable molecules. *J. Phys. Condens. Matter*, **4**, 3053–3076.

19. de Pablo J.J., Laso,M. and Suter,U.W. (1992) Simulation of polyethylene above and below the melting point. *J. Chem. Phys.*, **96**, 2395–2403.

20. Rosenbluth,M.N. and Rosenbluth,A.W. (1955) Monte Carlo calculation of the average extension of molecular chains. *J. Chem. Phys.*, **23**, 356–359.

21. Ha,B.-Y.Y. and Thirumalai,D. (1995) Electrostatic Persistence Length of a Polyelectrolyte Chain. *Macromolecules*, **28**, 577–581.

22. Nagashima,H. and Asakura,S. (1980) Dark-field Light Microscopic Study of the Flexibility of F-actin Complexes. *J. Mol. Biol.*, **136**, 169–182.

23. Schlick,T. and Perišić,O. (2009) Mesoscale Simulations of Two Nucleosome-Repeat Length Oligonucleosomes. *Phys. Chem. Chem. Phys.*, **11**, 10729–10737.