Article

Topological Polymorphism of the Two-Start Chromatin Fiber

Davood Norouzi[1] and Victor B. Zhurkin[1,*]
[1]Laboratory of Cell Biology, National Cancer Institute NIH, Bethesda, Maryland

ABSTRACT Specific details concerning the spatial organization of nucleosomes in 30 nm fibers remain unknown. To investigate this, we analyzed all stereochemically possible configurations of two-start nucleosome fibers with short DNA linkers L = 13–37 bp (nucleosome repeat length (NRL) = 160–184 bp). Four superhelical parameters—inclination of nucleosomes, twist, rise, and diameter—uniquely describe a regular symmetric fiber. The energy of a fiber is defined as the sum of four terms: elastic energy of the linker DNA, steric repulsion, electrostatics, and a phenomenological (H4 tail–acidic patch) interaction between two stacked nucleosomes. By optimizing the fiber energy with respect to the superhelical parameters, we found two types of topological transition in fibers (associated with the change in inclination angle): one caused by an abrupt 360° change in the linker DNA twisting (change in the DNA linking number, $\Delta Lk = 1$), and another caused by overcrossing of the linkers ($\Delta Lk = 2$). To the best of our knowledge, this topological polymorphism of the two-start fibers was not reported in the computations published earlier. Importantly, the optimal configurations of the fibers with linkers L = 10*n* and 10*n* + 5 bp are characterized by different values of the DNA linking number—that is, they are topologically different. Our results are consistent with experimental observations, such as the inclination 60° to 70° (the angle between the nucleosomal disks and the fiber axis), helical rise, diameter, and left-handedness of the fibers. In addition, we make several testable predictions, among them different degrees of DNA supercoiling in fibers with L = 10*n* and 10*n* + 5 bp, different flexibility of the two types of fibers, and a correlation between the local NRL and the level of transcription in different parts of the yeast genome.

INTRODUCTION

The structure of nucleosome, the basic unit of DNA organization in eukaryotic chromatin, was determined with atomic resolution more than a decade ago (1). However, stereochemical details of the next level in the hierarchical organization of DNA in chromatin are still the subject of intense debate (reviewed in (2–5)). Among different models proposed to describe the spatial arrangement of nucleosomes in the so-called 30 nm fiber, the solenoid (6–8) and the two-start (9–13) conformations are probably the best known. The x-ray structure of a tetranucleosome solved by Schalch et al. (10) and Cryo-EM images obtained by Song et al. (11) support the two-start organization. On the other hand, the Cryo-EM data presented by Robinson et al. (7) support the solenoid (interdigitated) model. Based on single-molecule pulling experiments, Kruithof et al. (8) suggested that both spatial nucleosome arrangements occur, and that the linker length, L, determines which arrangement is more favorable: the two-start fiber is more stable for short linkers (e.g., L = 20 bp), whereas the solenoid is formed when linkers are 50 bp or longer.

All the results mentioned above were obtained using arrays of strongly positioned 601 nucleosomes (14), with the nucleosome repeat length (NRL) varying from 167 to 237 bp in increments of 10 bp. Assuming that the nucleosome core is 147 bp, this means that the linker L belongs to the {10*n*} series. Only recently, Correll et al. (12) analyzed nucleosome arrays with L = 25 bp (NRL = 172 bp) and demonstrated that they have a stronger preference for the unfolded state compared with the arrays with L = 20 or 30 bp. The increased fiber plasticity observed for L = 25 bp may be functionally significant because this linker length belongs to the same {10*n* + 5} series as the L = 15 bp, known to be predominant in yeast chromatin (15–17). According to the earlier biochemical data (18), linker sizes in the higher eukaryotes are also close to {10*n* + 5}.

The distinct charge-dependent folding pathways observed for fibers with L = 10*n* and L = 10*n* + 5 (12) suggest different morphologies, and in particular, different spatial arrangements of DNA linkers in these fibers. Various mechanical models of the fiber have been proposed earlier (13,19,20), with linking number ΔLk varying from 0 to –2 per nucleosome, depending on the linker DNA configuration. In principle, computational analysis could shed light on the molecular mechanisms responsible for this topological polymorphism.

However, most of the computational studies of chromatin fibers did not focus on their topological properties. The early studies were based on the two-angle model introduced by Woodcock et al. (21), where the linker DNA was considered to be a straight rod (22,23). Later, elastic energy terms were added to the model, to account for the linker bending and twisting (24–26). Eventually, the number of independent parameters in the regular fiber model was increased up to six in the study by Koslover et al. (27). The regular fibers were also

Submitted January 26, 2015, and accepted for publication April 15, 2015.
*Correspondence: zhurkin@nih.gov
Editor: Jason Kahn.

0006-3495/15/05/2591/10 $2.00
http://dx.doi.org/10.1016/j.bpj.2015.04.015

considered by Scipioni et al. (28) who optimized torsional energy of the linker DNA both for the two-start helices and for the solenoid (interdigitated) structures. In addition, several groups used Monte Carlo simulations to study the influence of charges on structural variability of chromatin fibers (29,30) including formation of loops (31). Recently, Nam and Arya (32) analyzed topological variants of the chromatin fibers under external torsional forces, applying Brownian dynamics to an array of nucleosomes with the fixed linker size L ≈ 60 bp. Still, it should be noted that the topological polymorphism inherent in the chromatin fibers (and its relation to the linker length) was not addressed explicitly in the computer modeling studies mentioned above.

In this study, we make a systematic analysis of all stereochemically possible two-start left-handed configurations of the chromatin fiber, paying special attention to topological polymorphism. We focus on the two-start fibers, as it allows us to compare our results with the high-resolution models based on x-ray (10) and Cryo-EM data (11). We limit our analysis to the relatively short linkers, L = 13 to 37 bp, to exclude the possible formation of solenoid-type structures observed for the long linkers (7,8). We are mostly interested in the left-handed configurations, because the small-angle x-ray scattering analysis performed by Williams et al. (13) suggests formation of left-handed nucleosome fibers in solution. In addition to DNA elastic energy and steric restraints (27,28), our model includes electrostatic and H4 tail–acidic patch interactions. We demonstrate that for each linker L, in addition to the conformers described earlier (10,11,27–31) there is a novel family of topoisomers with different ΔLk that may be biologically relevant. We also show that the right-handed fibers are energetically less favorable than the left-handed ones. Finally, we propose specific ways to check the validity of our theoretical findings.

MATERIALS AND METHODS

Arrangement of nucleosomes in chromatin fiber

We consider symmetric regular fibers—that is, the fibers that retain the same shape if they are turned upside down, and whose linkers have the same conformation. In this case, the fiber geometry is described by four parameters, three of which specify positions of nucleosome centers in a cylindrical coordinate frame (Fig. 1 *A*): radius *r*, rise *h*, and polar angle φ or the twist angle between consecutive nucleosomes (27). This angle determines the number of stacks, N, in the fiber: $N \approx 360^\circ/\varphi$. For instance, $\varphi = 120^\circ$ corresponds to the three-stack fiber (N = 3).

The superhelical fiber twist is defined as $\omega = 2\varphi - 360^\circ$. Since we are mostly interested in the two-stack left-handed fibers, we set the value of φ to vary between 150° and 180°. (The values of $\varphi > 180^\circ$ will give rise to right-handed structures). The radius and rise values were not restrained.

The fourth parameter, ρ, defines inclination of nucleosomes relative to the fiber axis (Fig. 1 *B*). In general, orientation of a nucleosome would be defined by three rotations around the *X*, *Y*, and *Z* axes (Fig. 1 *A*). If, however, symmetric fibers are considered, then only one rotation around the dyad axis *X*, ρ, is permitted (Fig. 1 *B*); rotations around the *Y* and *Z* axes would break the symmetry.

If the nucleosomes were disconnected, variation of the inclination angle ρ from 0° to 180° would cover all possible structures of a chromatin fiber. However, because the DNA linkers connect the exit and the entry points in consecutive nucleosomes in a fixed order (see the *red* and *yellow balls* in Fig. 1 *A*), variation of the angle ρ by 180° is not sufficient. A comprehensive analysis of all possible fiber structures requires variation of this angle by 360°. This consideration distinguishes our approach from the studies published earlier (27,28).

To calculate the optimal geometry of the internucleosomal linkers, we used numeric minimization (see Supporting Material).

Energy terms

Four energy terms are considered in our calculations: elastic, electrostatic, and histone H4 tail–acidic patch interactions, as well as the steric hindrance term.

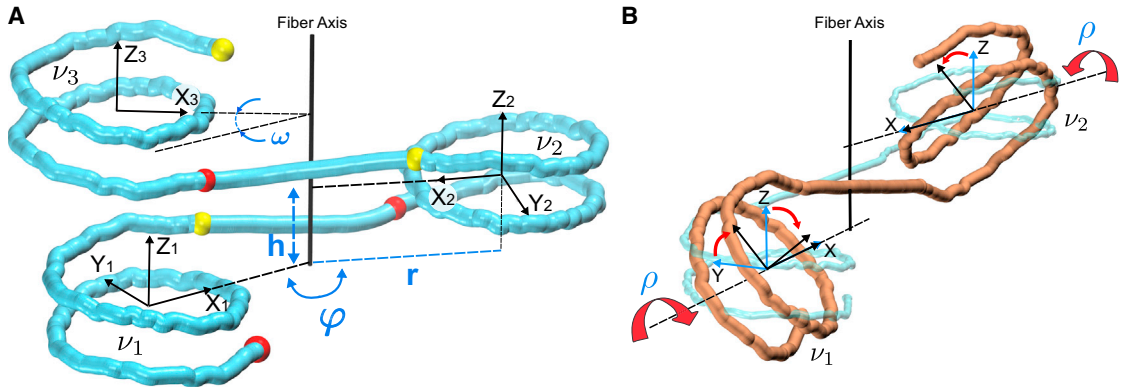

FIGURE 1 Definition of the chromatin fiber parameters. (*A*) Cylindrical parameters *h* (rise), *r* (radius), and polar angle φ. This angle determines the superhelical fiber twist: $\omega = 2\varphi - 360^\circ$ (the polar angle between the centers of nucleosomes ν1 and ν3). In this particular case, $\varphi = 165^\circ$ and $\omega = -30^\circ$. The entry points of nucleosomes are shown as red balls, and the exit points as yellow balls. Each nucleosome is associated with a right-handed coordinate frame in which axis *Z* represents the superhelical axis of the nucleosome (calculated as described in (44)), axis *X* points toward the nucleosome dyad, and axis *Y* is perpendicular to *X* and *Z*. (*B*) Inclination angle ρ (defining rotation around the dyad axis *X*) is related to the angle γ measured by electric dichroism (39–41), which is the angle between the *Y* axis and the fiber axis. These two angles are related by equation $\gamma = |\, 90^\circ - |\, \rho\, |\,|$. Increasing the inclination angle by $\Delta\rho$ leads to decreasing the total twisting of the linker between nucleosomes ν1 and ν2 by ~$2\Delta\rho$. The images were prepared with the VMD software package (45). To see this figure in color, go online.

1) The elastic energy of the linker DNA deformations is calculated using the knowledge-based potential functions introduced by Olson et al. (33) (see Supporting Material).
2) The electrostatic energy is calculated using the Coulomb potential with 30 Å cutoff and the water dielectric constant $\varepsilon = 78$. The DNA charges are partially neutralized because of the salt screening effect (see Supporting Material). In addition, we made computations with the Debye-Huckel potential (see Supporting Material).
3) The H4 tail–acidic patch interactions are modeled phenomenologically, with the minimum of the internucleosome interaction –8 *kT* (see Supporting Material).
4) Steric clashes are modeled by a van der Waals-like repulsion potential (see Supporting Material). Note that the nucleosome core particles (DNA and histones) are considered to be rigid in our computations.

Computation of DNA writhing and linking number

To evaluate the topological changes occurring in DNA on formation of a nucleosome fiber, we use three topological parameters: ΔTw (the change in DNA twisting); DNA writhing, *Wr*; and the change in the linking number, ΔLk (compared with the relaxed state of DNA) (34–36). To calculate the DNA writhing, *Wr*, we use the quadrangle method described by Levitt (37) and Klenin and Langowski (38). Topological parameters *Wr* and ΔLk are only meaningful for closed or restrained-end DNA. Therefore, we added extra base pairs to close the DNA chain. There is no imposed superhelical density in our model. We find the minimum energy conformation for the free DNA and then close the ends and calculate the topological parameters. (For details, see Supporting Material.)

RESULTS AND DISCUSSION

The twisting- and writhing-induced transitions in nucleosome fiber

The numerical results presented in this section were obtained mostly for the fibers with linker L = 20 bp (sometimes, for L = 25 bp); however, the observed trends are valid for the other L values as well. First, consider variability of the two-start nucleosome fiber (Fig. 2) when the angle ρ defining inclination of the nucleosome disk relative to the fiber axis (Fig. 1) changes from –180° to 180°. The fibers with $\rho < -90°$ are energetically unfavorable (Fig. 3 *A*). Nevertheless, they are shown to illustrate the whole spectrum of the fiber conformations. Note that

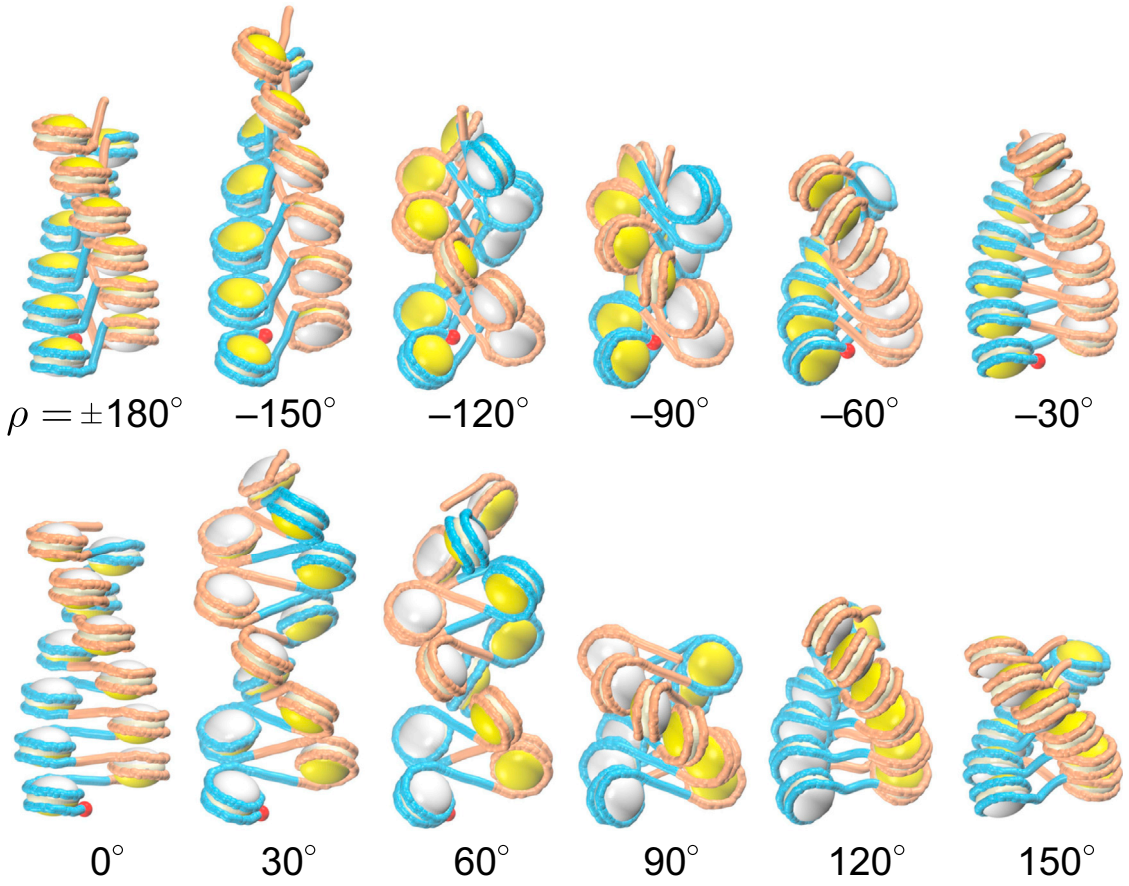

FIGURE 2 Two-start chromatin fibers with the linker L = 20 bp, with the inclination angle ρ changing from ±180° to 150°. All conformations correspond to the energetically optimal points (Fig. 3 *A*). Note that all these structures are different—in particular, the fiber with $\rho = -90°$ differs from the fiber with $\rho = 90°$ (see positions of the DNA entry points shown as *red balls*). The DNA is presented in alternating blue and orange colors, to emphasize the two stacks of nucleosomes and to clarify the passage of linkers in the middle of the fiber. The histone cores are also shown in two colors: the entry sides of nucleosomes (interacting with the DNA entry points shown as *red balls*) are colored in yellow, and the exit sides are in white. In this way, it is easier to distinguish the fiber configurations. For example, for $\rho = 0$ the white sides of the cores face upward, whereas for $\rho = 180°$ they face downward. To see this figure in color, go online.

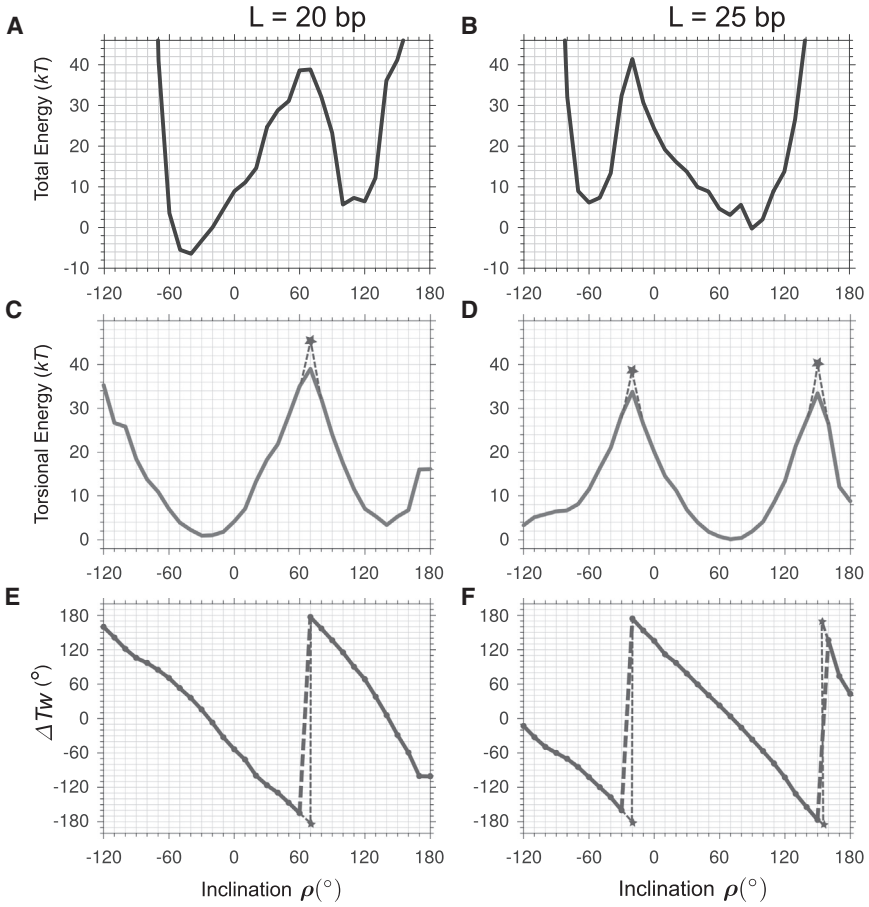

FIGURE 3 Energy profiles and DNA twisting for the chromatin fibers with the linkers L = 20 (*A*, *C*, and *E*) and 25 bp (*B*, *D*, and *F*). Each point corresponds to a minimum of the conformation energy in the space of three fiber parameters—radius, rise, and polar angle φ (Fig. 1 *A*). Energy is very high for $\rho < -120°$ so this region is omitted for clarity. (*A* and *B*) Total energy. (*C* and *D*) Torsional energy of the linker DNA. (*E* and *F*) Changes in the excess twisting, ΔTw, of the linker DNA as a function of the inclination angle ρ. The points shown by asterisks in (*C*)–(*F*) correspond to the linker DNA undertwisted or overtwisted by more than 180°. (See the main text for details.)

all these fibers have different geometry—see, for example, positions of the DNA entry points shown as red balls. This explains why one should consider the 360° variation of the inclination angle. (Also, see Movie S1 and Supporting Material.)

The changes in the fiber shape and dimension presented in Fig. 2 seem to be gradual, if we consider mainly the orientation of nucleosomes and linkers. If, however, we focus on the details of the DNA conformation and its energy, it becomes clear that the linker DNA undergoes two sharp transitions, one close to $\rho = 60°$ and another around $\rho = -120°$. We will consider these transitions in more detail.

Type I transition

The increase in the angle ρ from 60° to 70° is accompanied by an abrupt ~360° change in the DNA twisting (Fig. 3 *E*). This effect can be explained as follows. Increasing the inclination angle ρ by some amount $\Delta\rho$ leads to decreasing the total twisting of the linker by $\sim 2\Delta\rho$ (Fig. 1 *B*). The optimal torsional energy of DNA corresponds to $\rho \approx -30°$ (Fig. 3 *C*). An increase in ρ by 90°, from –30° to 60°, brings about a significant ~180° untwisting of the linker DNA (Fig. 3 *E*). As a consequence, the torsional energy of the DNA linker increases by ~40 *kT* (Fig. 3 *C*).

A further increase in ρ, up to 70°, would be even more costly energetically if the changes in DNA conformation followed the same trend—that is, if the linker DNA untwisted by 183°, the torsional energy would exceed 43 *kT* (see the points shown by *asterisks* in Fig. 3, *C* and *E*). If, however, the linker DNA abruptly changes its conformation and becomes overtwisted at $\rho = 70°$, the DNA torsional energy will be less than that described above because overtwisting of the linker by 177° is more favorable than its undertwisting by 183° (Fig. 3, *C* and *E*). This conformational transition is illustrated in Fig. 4. An increase in the inclination angle ρ beyond the transition point brings additional relief in the DNA twisting tension (Fig. 3 *C*). As a result, the second minimum of the DNA torsional energy is achieved for the inclination angle $\rho = 140°$, whereas the local minimum of the total energy is observed for $\rho = 100°$ to 120° (Fig. 3 *A*).

Note that for closed DNA (or for DNA with the fixed ends) the twist-induced transition described above is practically impossible without type I topoisomerase because in such a case, each nucleosome has to rotate by ~360° relative to its neighbor to achieve this topological change. Moreover, this twisting will accumulate, that is, the third nucleosome will rotate relative to the first one by ~720°, and the fourth nucleosome will rotate by ~1080°, and so on.

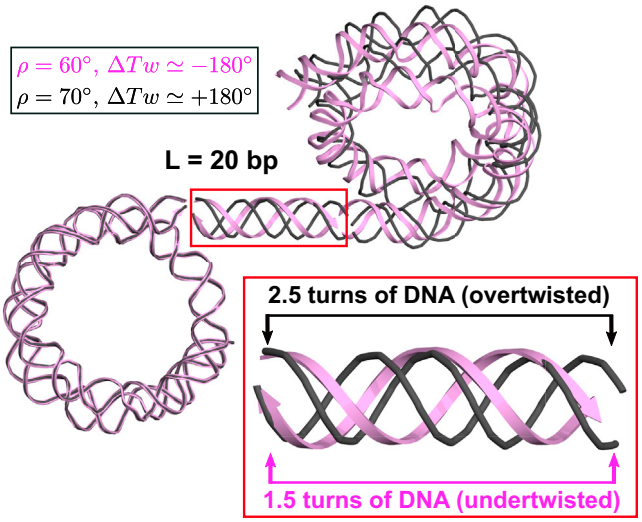

FIGURE 4 Abrupt change in the linker DNA excess twisting, ΔTw, caused by increase in the inclination angle ρ. For $\rho = 60°$ (*pink ribbon*) the linker DNA is undertwisted by ~180°. For $\rho = 70°$ (*black wire*) the linker DNA is overtwisted by ~180°. The nucleosomes on the left are superimposed. The images were prepared with the Chimera software package (46). Inset: the DNA twisting differs by one helical turn (~360°) in the two linkers. Compare the black wires representing ~2.5 turns of DNA in the case of $\rho = 70°$ with the pink ribbons corresponding to ~1.5 helical turns ($\rho = 60°$). To see this figure in color, go online.

Type II transition

At approximately $\rho = -120°$ we encounter another type of sharp transition caused by the change in DNA writhing, *Wr*. If we compare the two fiber structures with $\rho = -130°$ and $\rho = -110°$ (Fig. 5, *A* and *C*), in the first case the two bottom linkers shown in blue are positioned above the linker shown in orange, whereas in the second case, the blue linkers are positioned below the orange one. This change in the spatial arrangement of linkers is accompanied by DNA passing through itself (Fig. 5 *B*) and leads to a sharp increase in DNA writhing by ± 2. This is a typical example of the topological transition in DNA induced by type II topoisomerase (Topo II).

Thus, we see that the conformational space of nucleosome fibers with linker L = 20 bp is characterized by two local energy minima surrounded by topological barriers. In this regard, the fibers with L = 25 bp are qualitatively similar to the fibers with L = 20 bp, but the torsional energy minima and maxima are positioned differently (Fig. 3 *D*). The exact values of DNA writhing and twisting, as well as their dependence on the linker length and inclination angle ρ are discussed below.

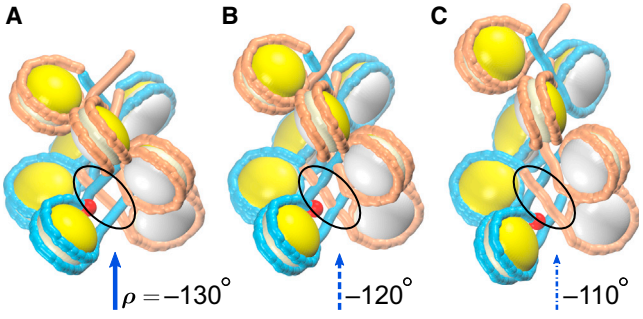

FIGURE 5 The writhing-induced transition occurring in the chromatin fiber with the linker L = 20 bp. Note the different mutual positioning of linkers in the fibers with $\rho = -130°$, $-120°$, and $-110°$ (*A*, *B*, and *C*, respectively). (*A*) The two bottom linkers shown in blue are positioned above the linker shown in orange (*solid blue arrow*). (*B*) Overcrossing (*intersection*) of the blue and orange DNA linkers (*dotted-line blue arrow*). (*C*) The two bottom blue linkers are positioned below the orange linker (*dash-and-dot blue arrow*). To see this figure in color, go online.

Energetically optimal conformations

As shown above, the torsional energy profiles for fibers with linkers L = 20 and 25 bp have two minima spaced ~180° apart, with the energy values close to zero for both minima (Fig. 3, *C* and *D*). Consideration of other energy terms changes the balance between the two minima (Figs. 3, *A* and *B*, and S5 in the Supporting Material), while preserving the bimodality of the energy profiles. For other linker sizes, the same bimodal effect is visible in the energy landscape in the two-dimensional (2D) space (ρ, L) shown in Fig. 6. In terms of the inclination angle ρ, the dark blue regions with the lowest energy span either the $-80° < \rho < -30°$, or $80° < \rho < 120°$ interval. In terms of the linker length (along the vertical axis), this 2D landscape has a clear 10 to 11 bp periodicity, which reflects the helical period of B-DNA. In particular, the energy profiles for L = 20 and 30 bp are similar to each other, with the global energy minima appearing on the left side of the plot (negative ρ values), whereas for L = 15, 25, and 35 bp, the minima are on the right side (positive ρ values).

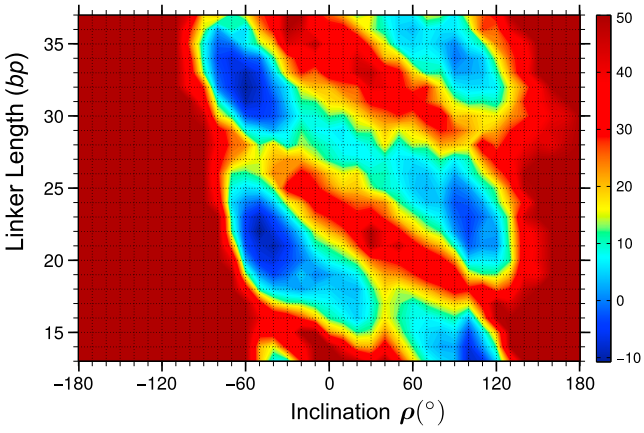

FIGURE 6 Total energy of the two-start fibers as a function of the inclination angle ρ and the linker length L. Optimization is made in the space of three remaining fiber parameters, i.e., radius, rise, and polar angle φ. Dark blue regions (*negative energy*) represent stable structures, whereas dark brown areas are those with energies higher than 50 *kT*. Cross-sections of this 2D map for linkers L = 20 and 25 bp are shown in Fig. 3, *A* and *B*. To see this figure in color, go online.

The optimal conformations for L = 20 and 30 bp have the inclination angle $\rho \approx -45^\circ$ and -60°, respectively, which is in accord with models based on x-ray (10) and Cryo-EM data (11). Note that the blue and red elongated regions (with low and high energy, respectively) are tilted from top-left to bottom-right. This skewed energy landscape pattern is entirely consistent with the tilt of the contour lines for excess twisting, ΔTw, in Fig. 7 *A*.

The brown areas to the left and right of the blue islands of stability have extremely high energy. The $\rho < -90^\circ$ region is unfavorable because of linker-linker and linker DNA-nucleosome clashes (see Fig. S5). For $|\rho| > 130^\circ$, the DNA elastic energy is prohibitively high because of strong bending of the linkers (see the kinky DNA conformations in Fig. 2 and Movie S1). For details of the contributions of each energy term, see Fig. S5.

As to the central area between the low-energy blue regions, the fiber conformations with $\rho = -20^\circ$ to 60° are unfavorable for a number of reasons. The region around $\rho = 0$ is characterized by a partial loss of internucleosome stacking and high elastic energy of DNA (compare the structures with $\rho = -30^\circ$ to 30° in Fig. 2 to see the loss of stacking, and see Fig. S5). For the ρ angles between 20° and 60°, the adjacent nucleosomal DNA in the same stack (e.g., the two orange nucleosomes in Fig. 2, $\rho = 30^\circ, 60^\circ$) are colliding, thereby inducing electrostatic repulsion. In general, our results are consistent with the electric dichroism (ED) data suggesting strong inclination of nucleosomes in the fiber (39–41).

In addition, we analyzed the right-handed two-start fibers ($180^\circ < \varphi < 210^\circ$) and found that the linker-nucleosome clashes prevent close stacking of nucleosomes in this case; as a consequence, the minimum energy is higher than in the left-handed fibers (see Supporting Material). Apparently, the chirality of nucleosomes dictates the left-handedness of the chromatin fiber.

Topological changes in chromatin fibers with various linker lengths

To evaluate topological changes occurring in DNA on formation of a nucleosome fiber, we used three topological parameters: the change in DNA twisting, ΔTw; DNA writhing,

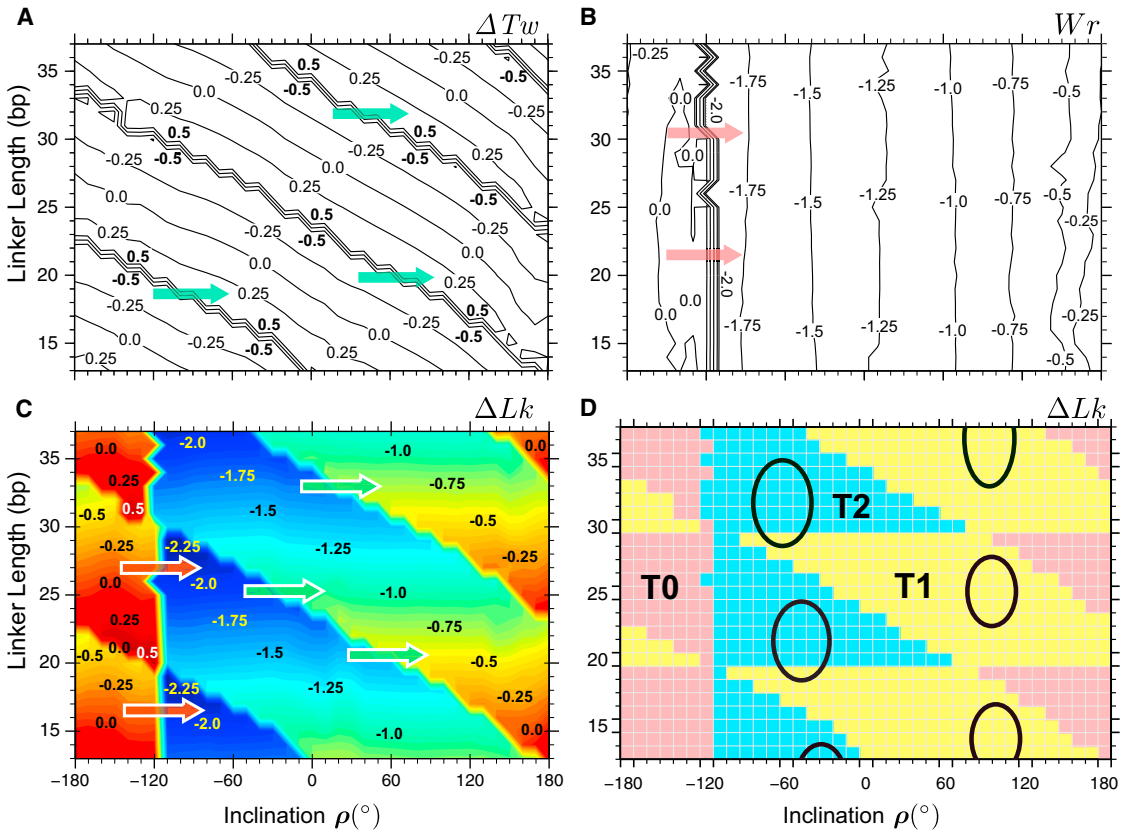

FIGURE 7 Variation of the three topological parameters as a function of the inclination angle ρ and the linker length L. The contour plots are consistent with the conformation energy map presented in Fig. 6. (*A*) The excess twisting of the linker DNA, ΔTw (compare with Fig. 3, *E* and *F*). The ΔTw values are expressed in the number of DNA helical turns (or degrees/360). (*B*) The writhing number per nucleosome, *Wr*. (*C*) Change in the linking number per nucleosome, ΔLk (relative to the relaxed state of DNA). See Fig. S12 for additional details. (*D*) Classification of the fiber conformations: three families of topoisomers. The cyan region represents family T2 with the lowest ΔLk values (approximately from –2.2 to –1.5). The yellow region corresponds to the family T1 with intermediate ΔLk (from –1.2 to –0.5). The highest ΔLk values define family T0 (approximately from 0 to 0.5). The ovals denote energetically optimal regions shown in Fig. 6. To see this figure in color, go online.

Wr; and the change in the linking number, *ΔLk*—see Figs. S10 and S12 and Table S2.

Variation in linker DNA twisting as a function of the inclination angle ρ and the linker length L is presented in Fig. 7 *A*. This plot is a generalization of Fig. 3, *E* and *F*, where variation of *ΔTw* is shown for L = 20 and 25 bp. The difference is that in Fig. 3, *E* and *F* the excess twist values are given in degrees, whereas in Fig. 7 *A* they are given in DNA helical turns (i.e., degrees/360). The triple tilted contour lines denote the abrupt change in *ΔTw* from –0.5 to 0.5 (corresponding to the sharp increase in excess twisting from –180° to 180° in Fig. 3, *E* and *F*, see *green arrows*). The tilted contour lines also imply that *ΔTw* depends both on the angle ρ and the linker L.

By contrast, the contour lines in Fig. 7 *B* are nearly vertical. This means that the DNA writhing varies substantially with changes in the inclination, ρ, but practically does not depend on the linker length L. In particular, the abrupt changes in *Wr* from 0 to –2 occur at $\rho = -120° (\pm 10°)$ for all values of L (see *red arrows*). This sharp transition is accompanied by the over-crossing of DNA linkers as was described above for L = 20 bp (Fig. 5).

Adding together the excess twisting of the linker, *ΔTw*, and the writhing, *Wr*, we find the change in the DNA linking number, $\Delta Lk = \Delta Tw + Wr$ (Fig. 7 *C*). The resulting contour plot represents a rather complicated pattern containing vertical, horizontal, and tilted lines. The tilted contour lines in Fig. 7 *C* originate from the skewed triple lines in Fig. 7 *A*, which highlight the abrupt changes in *ΔTw* from 0.5 to –0.5. Accordingly, *ΔLk* also changes by ~1 (see *green arrows* intersecting *tilted contour lines* in Fig. 7, *A* and *C*). The multiple vertical lines at $\rho \approx -120°$ demarcate a stepwise transition in *ΔLk,* which originates from a similarly sharp transition in *Wr* visible in Fig. 7 *B*; both *ΔLk* and *Wr* values are changed by ~2 (see *red arrows*).

The horizontal contour lines in Fig. 7 *C* are the result of summation of the *ΔTw* and *Wr* values. Despite the fact that both *ΔTw* and *Wr* strongly correlate with the inclination angle ρ (Fig. 7, *A* and *B*), the sum of these parameters looks like a piece-wise constant function of the angle ρ (Fig. S12).

All these findings taken together allow us to divide the two-start fiber topoisomers into three families, T2, T1, and T0 (Fig. 7 *D*). The structures with the lowest *ΔLk* values (from approximately –2.2 to –1.5) belong to the family T2. The fibers with intermediate *ΔLk* (from –1.2 to –0.5) comprise family T1. The highest *ΔLk* values define family T0 (from ~0 to 0.5). The transition between families T2 and T1 is related to abrupt changes in DNA twisting (shown by *green arrows* in Fig. 7, *A* and *C*). The sharp transition between the families T0 and T2 is related to the sudden increase in DNA writhing (*red arrows* in Fig. 7, *B* and *C*). The ovals correspond to energetically optimal regions (Fig. 6). Importantly, the optimal topoisomers for L = 20 and 30 bp belong to the family T2, whereas for L = 15, 25, and 35 bp they belong to the family T1.

Optimal conformations for L = 20 and 25 bp are presented in Fig. 8. In terms of topology, the difference between the two families of structures is more obvious in the stretched forms. In particular, it is clear that in the case of T2 (Fig. 8 *C*) the nucleosomal DNA makes almost two complete turns—apparently more than in T1 (Fig. 8 *F*). Hence, the DNA writhing and linking number are larger in absolute value in the T2 structures (compare with Fig. 7).

In terms of stability, the two types of structures are also different. Direct visual comparison shows that the internucleosome stacking is stronger for L = 20 bp: the contact area is larger (Fig. 8 *A*) and the H4 tail–acidic patch bridge length (Fig. 8 *B*) is closer to the optimal value in this case. Accordingly, the numerical results indicate that the stacking energy is lower for L = 20 bp (Fig. S5) and the total energy

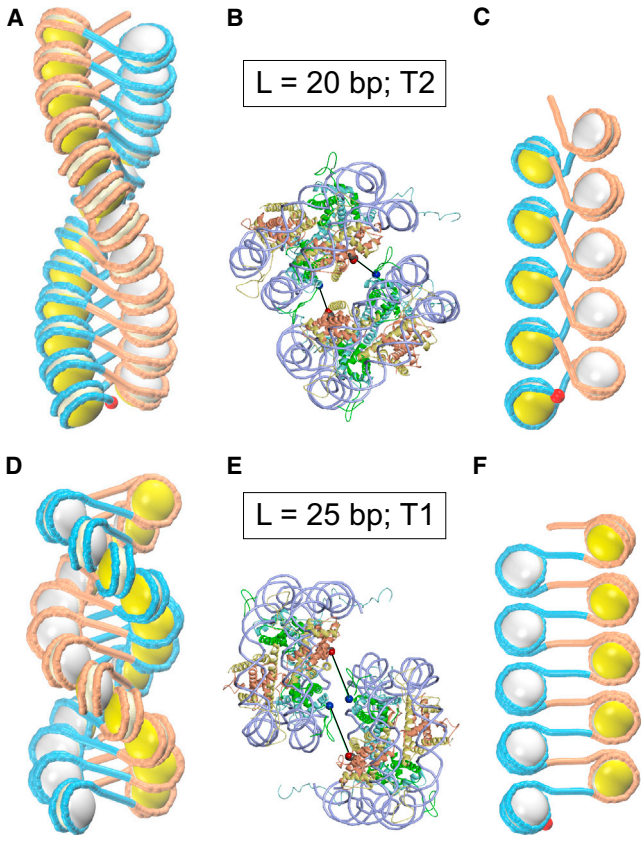

FIGURE 8 Topoisomers T2 and T1 for linkers L = 20 bp (*A–C*) and L = 25 bp (*D–F*), respectively. The optimal structure for L = 20 bp has inclination angle $\rho = -45°$ and rise $h = 27$ Å (*A* and *B*); the optimal structure for L = 25 bp has $\rho = 90°$ and $h = 23$ Å (*D* and *E*). The difference between the two families is clearly visible in the stretched forms with $h = 55$ Å (*C* and *F*). Note a more extensive folding of DNA in the nucleosomal disks in (*C*). This is consistent with the DNA writhing: Wr = –1.7 in (*C*) and Wr = –0.9 in (*F*). Red balls indicate the DNA entry points. The detailed representations in (*B*) and (*E*) show that there are no steric clashes between the stacked nucleosomes ν1 and ν3 (Fig. 1 *A*). The internucleosome stacking is stabilized by the H4 tail-acidic patch bridges shown by black lines; the distances between Asp24 (H4) and Glu61 (H2A) are 23 Å in (*B*) and 34 Å in (*E*) (see Fig. S3 for details). To see this figure in color, go online.

minimum is deeper for L = 20 bp compared with L = 25 bp (Figs. 3, *A* and *B*, and 6). Overall, our computations strongly suggest that the T2 topoisomer for L = 20 bp is the most favorable energetically, whereas the T1 topoisomer for L = 25 bp is the best one.

The x-ray (10) and Cryo-EM data (11) provide strong support for stabilization of T2 topoisomers for L = 10*n*. In contrast, there is no direct evidence for formation of the T1 topoisomers. We believe this is because in all imaging studies except the one reported by Correll et al. (12), the linkers of type 10*n* + 5 were not analyzed. However, the linking number measurements suggest the presence of the T1 topoisomer, at least in SV40 minichromosomes (42,43).

CONCLUSIONS

We performed a thorough analysis of stereochemically feasible two-start chromatin fibers with internucleosome linkers varying from 13 to 37 bp. For each linker length, we found three families of topoisomers characterized by different linking number values (Fig. 7 D). The families are separated by topological barriers—that is, conformational transition between the structures belonging to different families is only possible in the presence of cutting-closing enzymes (topoisomerases of type I or II) for closed or restrained-end fiber. Depending on the linker length, the most favorable topoisomer belongs either to family T2 ($\Delta Lk \approx -1.5$ to -2.2), or to family T1 ($\Delta Lk \approx -0.5$ to -1.2) (Figs. 6 and 7 *D*). Existence of the two families of structures with different linking numbers is consistent with the earlier observations of Williams et al. (13) who built space-filling models of the crossed-linker fibers with $\Delta Lk = -1$ and -2. The family T0 comprises highly unfavorable conformations, at least concerning the short linkers in this study (Fig. 7 *D*).

In this study we focused on the relationship between the DNA linker length and the fiber topology. We found that although in terms of torsional energy both topoisomers T1 and T2 are equally feasible, adding all energy terms together makes one topoisomer preferable to the other. One of our most important conclusions is that for linkers L = 10*n*, the topoisomer T2 is more favorable, whereas for linkers L = 10*n* + 5 the topoisomer T1 is the optimal conformation. The superhelical density of DNA in the optimal fiber structures ($\sigma = -0.09$ and $\sigma = -0.06$ for L = 20 and 25 bp, respectively) is in the range of values observed experimentally (47) (see Supporting Material). These values give an estimate of the torsional stress experienced by DNA were the histones to be removed.

The T1 topoisomer (with L = 25 bp) demonstrates a higher bending and stretching flexibility compared with the T2 topoisomer (with L = 20 bp). This may be explained by the fact that nucleosomes are strongly inclined in the T1 form (Fig. 8), and the internucleosome stacking interactions are less sensitive to deformations of the fiber (data not shown). Our data complement what was recently found by Clauvelin et al. (48) about the effects of linker length on the interactions and deformability of simulated nucleosome arrays. Note that our results are in accord with the findings of Correll et al. (12) suggesting that in the absence of the linker histone the fibers with L = 10*n* +5 are less stable than the fibers with L = 10*n*. In fact, our study was motivated by these findings (12); before 2012 there were no experimental data indicating that folding of the chromatin fiber periodically depends on the changes in linker DNA length.

To the best of our knowledge, such a detailed computational search of energetically favorable two-start fiber conformations has not been previously carried out. The critical difference between our approach and those used by others (27–31) is that we considered all possible orientations of the nucleosomal disks relative to the fiber axis, resulting in all possible topologies. (That is, we allowed a 360° range of variation for the inclination angle ρ—see Figs. 1 and 2.) By contrast, Koslover et al. (27) varied the nucleosome inclination angle β in the interval only between 0° and 180° (see their Fig. 5). Similarly, Scipioni et al. (28) also did not investigate the entire phase space of the fiber conformations. The two-start fibers analyzed in these studies closely resemble the x-ray-based model by Schalch et al. (10) and belong to the family T2; the T1 conformations were not considered.

According to the ED measurements (39–41), nucleosomes are strongly inclined in fibers. The angle between the nucleosome disks and the fiber axis, $\gamma = |\,90^\circ - |\,\rho\,|\,|$, is estimated to be ~20° to 30° for various chromatin fibers with mixed internucleosome spacing. In terms of the inclination angle ρ used in this study (Fig. 1), the average value of $|\rho|$ is ~60° to 70°. Our computations also suggest that the angle ρ significantly deviates from zero in the optimal conformations from both topological families: $|\rho| = 40^\circ$ to 60° for the T2, and 80° to 90° for the T1 family (Fig. 6), or $|\rho| \approx 65^\circ$ on average. Assuming that the fractions of topoisomers T1 and T2 are comparable in the fibers used for ED measurements, we conclude that our results are generally consistent with experimental data.

It can be argued that there are multiple ways to take into account the electrostatic potential. In this study we chose to employ the Coulomb potential with a 30 Å cutoff and charge reduction because of salt screening. We also used the Debye-Huckel potential; importantly, the overall results did not change significantly (see Supporting Material). The other critical issue is modeling of the internucleosome stacking interaction. Here, it is described with a phenomenological potential (see Supporting Material) whose depth, $-8\ kT$, is taken as the average of the two experimental values obtained by fiber stretching (8,49). It would be interesting to see how using more rigorous physical approximations would change the exact energy landscape for the internucleosome stacking interaction. We found, however, that

varying the potential depth practically does not affect the total energy profiles (see Fig. S4). Entropy has not been considered in our model explicitly. However, at the qualitative level we can estimate the entropy of conformational states comparing the widths of corresponding energy minima. In this way, one can see that the optimal energy minima (Fig. 3, *A* and *B*) are more favorable entropically than the secondary minima (because their energy profiles are wider). In summary, we conclude that the final results, especially the topological classification of the stereochemically feasible fiber conformations, are essentially independent of the numerical details of the potentials selected to model the histone-DNA interactions.

Potentially, any kind of chromatin polymorphism (structural and/or topological) is interesting from the point of view of gene regulation. The expected general implications of this study is that it may reveal new mechanisms for encoding structural information in a nucleosome array in the form of alternative topological states (i.e., topological switch between the T1 and T2 topoisomers).

The topological and structural aspects of our findings can be tested both in vitro and in vivo. First, employing the approach used by Simpson et al. (50) and Norton et al. (51), one can measure the number of superhelical turns in covalently closed DNA induced by formation of precisely positioned 601 nucleosomes (14) separated by linkers L = 20 or 25 bp. In this way, it is possible to detect the difference in linking number between the topoisomers T1 and T2 predicted by our model (Figs. 7 *D* and 8).

Second, the increased flexibility of the T1 topoisomer (with $L = 10n + 5$) compared with the T2 topoisomer (with $L = 10n$) may be utilized by cell in differential folding of the functionally distinct parts of the genome. In particular, we hypothesize that the highly and lowly expressed genes may have different average NRL values. To test this hypothesis, we analyzed high-resolution nucleosome positioning data in yeast (17,52) and found that there is a strong correlation between the level of expression and local nucleosome spacing: the average NRL = 161 to 162 bp for highly active genes (i.e., linker L = 14 to 15 bp), whereas NRL = 167 to 168 bp (i.e., linker L = 20 to 21 bp) for less frequently transcribed genes (A. Katebi, D.N., F. Cui, and V.B.Z., unpublished observation). A possible explanation of this correlation is that the greater plasticity of the fibers with $L \approx 10n + 5$ may facilitate formation of gene loops (53,54), thereby inducing transcription of the corresponding genes.

The observed correlation may reflect a more general tendency of chromosomal domains containing active or repressed genes to retain topologically distinct higher-order structures. It remains to be seen whether this topological mechanism of transcription regulation is applicable to higher eukaryotes, but the available information (55) suggests that the changes in NRL on cell differentiation in mice follow the same trend as described above for yeast.

SUPPORTING MATERIAL

Supporting Materials and Methods, 12 figures, two tables, and two movies are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(15)00394-X.

AUTHOR CONTRIBUTIONS

V.B.Z. designed research; D.N. performed research; and D.N. and V.B.Z. analyzed data and wrote the article.

ACKNOWLEDGMENTS

We are grateful to Wilma Olson and Sergei Grigoryev for sharing their unpublished observations and valuable discussions and to George Leiman for text editing. We appreciate constructive critical remarks by anonymous reviewers. The research was supported by the NIH Intramural Research Program.

REFERENCES

1. Davey, C. A., D. F. Sargent, …, T. J. Richmond. 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. *J. Mol. Biol.* 319:1097–1113.
2. Tremethick, D. J. 2007. Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell.* 128:651–654.
3. van Holde, K., and J. Zlatanova. 2007. Chromatin fiber structure: where is the problem now? *Semin. Cell Dev. Biol.* 18:651–658.
4. Schlick, T., J. Hayes, and S. Grigoryev. 2012. Toward convergence of experimental studies and theoretical modeling of the chromatin fiber. *J. Biol. Chem.* 287:5183–5191.
5. Luger, K., M. L. Dechassa, and D. J. Tremethick. 2012. New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat. Rev. Mol. Cell Biol.* 13:436–447.
6. Routh, A., S. Sandin, and D. Rhodes. 2008. Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc. Natl. Acad. Sci. USA.* 105:8872–8877.
7. Robinson, P. J. J., L. Fairall, …, D. Rhodes. 2006. EM measurements define the dimensions of the '30-nm' chromatin fiber: evidence for a compact, interdigitated structure. *Proc. Natl. Acad. Sci. USA.* 103:6506–6511.
8. Kruithof, M., F. T. Chien, …, J. van Noort. 2009. Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30-nm chromatin fiber. *Nat. Struct. Mol. Biol.* 16:534–540.
9. Dorigo, B., T. Schalch, …, T. J. Richmond. 2004. Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science.* 306:1571–1573.
10. Schalch, T., S. Duda, …, T. J. Richmond. 2005. X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature.* 436:138–141.
11. Song, F., P. Chen, …, G. Li. 2014. Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science.* 344:376–380.
12. Correll, S. J., M. H. Schubert, and S. A. Grigoryev. 2012. Short nucleosome repeats impose rotational modulations on chromatin fibre folding. *EMBO J.* 31:2416–2426.
13. Williams, S. P., B. D. Athey, …, J. P. Langmore. 1986. Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length. *Biophys. J.* 49:233–248.
14. Lowary, P. T., and J. Widom. 1998. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* 276:19–42.

15. Lohr, D. E. 1981. Detailed analysis of the nucleosomal organization of transcribed DNA in yeast chromatin. *Biochemistry.* 20:5966–5972.
16. Wang, J. P., Y. Fondufe-Mittendorf, …, J. Widom. 2008. Preferentially quantized linker DNA lengths in Saccharomyces cerevisiae. *PLOS Comput. Biol.* 4:e1000175.
17. Cui, F., H. A. Cole, …, V. B. Zhurkin. 2012. Transcriptional activation of yeast genes disrupts intragenic nucleosome phasing. *Nucleic Acids Res.* 40:10753–10764.
18. Strauss, F., and A. Prunell. 1983. Organization of internucleosomal DNA in rat liver chromatin. *EMBO J.* 2:51–56.
19. Grigoryev, S. A., and L. B. Ioffe. 1981. The dependence of the linking number of a circular minichromosome upon the shape and the orientation of its nucleosomes. *FEBS Lett.* 130:43–46.
20. Worcel, A., S. Strogatz, and D. Riley. 1981. Structure of chromatin and the linking number of DNA. *Proc. Natl. Acad. Sci. USA.* 78:1461–1465.
21. Woodcock, C. L., S. A. Grigoryev, …, N. Whitaker. 1993. A chromatin folding model that incorporates linker variability generates fibers resembling the native structures. *Proc. Natl. Acad. Sci. USA.* 90:9021–9025.
22. Schiessel, H., J. Widom, …, W. M. Gelbart. 2001. Polymer reptation and nucleosome repositioning. *Phys. Rev. Lett.* 86:4414–4417.
23. Barbi, M., J. Mozziconacci, and J. M. Victor. 2005. How the chromatin fiber deals with topological constraints. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 71:031910.
24. Wedemann, G., and J. Langowski. 2002. Computer simulation of the 30-nanometer chromatin fiber. *Biophys. J.* 82:2847–2859.
25. Wong, H., J. M. Victor, and J. Mozziconacci. 2007. An all-atom model of the chromatin fiber containing linker histones reveals a versatile structure tuned by the nucleosomal repeat length. *PLoS ONE.* 2:e877.
26. Lanzani, G., and H. Schiessel. 2012. Out of register: how DNA determines the chromatin fiber geometry. *EPL.* 97:38002.
27. Koslover, E. F., C. J. Fuller, …, A. J. Spakowitz. 2010. Local geometry and elasticity in compact chromatin structure. *Biophys. J.* 99:3941–3950.
28. Scipioni, A., G. Turchetti, …, P. De Santis. 2010. Geometrical, conformational and topological restraints in regular nucleosome compaction in chromatin. *Biophys. Chem.* 148:56–67.
29. Kepper, N., D. Foethke, …, K. Rippe. 2008. Nucleosome geometry and internucleosomal interactions control the chromatin fiber conformation. *Biophys. J.* 95:3692–3705.
30. Perišić, O., R. Collepardo-Guevara, and T. Schlick. 2010. Modeling studies of chromatin fiber structure as a function of DNA linker length. *J. Mol. Biol.* 403:777–802.
31. Kulaeva, O. I., G. Zheng, …, W. K. Olson. 2012. Internucleosomal interactions mediated by histone tails allow distant communication in chromatin. *J. Biol. Chem.* 287:20248–20257.
32. Nam, G. M., and G. Arya. 2014. Torsional behavior of chromatin is modulated by rotational phasing of nucleosomes. *Nucleic Acids Res.* 42:9691–9699.
33. Olson, W. K., A. A. Gorin, …, V. B. Zhurkin. 1998. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. USA.* 95:11163–11168.
34. Fuller, F. B. 1978. Decomposition of the linking number of a closed ribbon: a problem from molecular biology. *Proc. Natl. Acad. Sci. USA.* 75:3557–3561.
35. Crick, F. H. 1976. Linking numbers and nucleosomes. *Proc. Natl. Acad. Sci. USA.* 73:2639–2643.
36. Prunell, A. 1998. A topological approach to nucleosome structure and dynamics: the linking number paradox and other issues. *Biophys. J.* 74:2531–2544.
37. Levitt, M. 1983. Protein folding by restrained energy minimization and molecular dynamics. *J. Mol. Biol.* 170:723–764.
38. Klenin, K., and J. Langowski. 2000. Computation of writhe in modeling of supercoiled DNA. *Biopolymers.* 54:307–317.
39. McGhee, J. D., D. C. Rau, …, G. Felsenfeld. 1980. Orientation of the nucleosome within the higher order structure of chromatin. *Cell.* 22:87–96.
40. Yabuki, H., N. Dattagupta, and D. M. Crothers. 1982. Orientation of nucleosomes in the thirty-nanometer chromatin fiber. *Biochemistry.* 21:5015–5020.
41. McGhee, J. D., J. M. Nickol, …, D. C. Rau. 1983. Higher order structure of chromatin: orientation of nucleosomes within the 30 nm chromatin solenoid is independent of species and spacer length. *Cell.* 33:831–841.
42. Keller, W. 1975. Determination of the number of superhelical turns in simian virus 40 DNA by gel electrophoresis. *Proc. Natl. Acad. Sci. USA.* 72:4876–4880.
43. Germond, J. E., B. Hirt, …, P. Chambon. 1975. Folding of the DNA double helix in chromatin-like structures from simian virus 40. *Proc. Natl. Acad. Sci. USA.* 72:1843–1847.
44. Tolstorukov, M. Y., A. V. Colasanti, …, V. B. Zhurkin. 2007. A novel roll-and-slide mechanism of DNA folding in chromatin: implications for nucleosome positioning. *J. Mol. Biol.* 371:725–738.
45. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38.
46. Pettersen, E. F., T. D. Goddard, …, T. E. Ferrin. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25:1605–1612.
47. Bauer, W. R. 1978. Structure and reactions of closed duplex DNA. *Annu. Rev. Biophys. Bioeng.* 7:287–313.
48. Clauvelin, N., P. Lo, …, W. K. Olson. 2015. Nucleosome positioning and composition modulate in silico chromatin flexibility. *J. Phys. Condens. Matter.* 27:064112.
49. Cui, Y., and C. Bustamante. 2000. Pulling a single chromatin fiber reveals the forces that maintain its higher-order structure. *Proc. Natl. Acad. Sci. USA.* 97:127–132.
50. Simpson, R. T., F. Thoma, and J. M. Brubaker. 1985. Chromatin reconstituted from tandemly repeated cloned DNA fragments and core histones: a model system for study of higher order structure. *Cell.* 42:799–808.
51. Norton, V. G., B. S. Imai, …, E. M. Bradbury. 1989. Histone acetylation reduces nucleosome core particle linking number change. *Cell.* 57:449–457.
52. Cole, H. A., B. H. Howard, and D. J. Clark. 2011. The centromeric nucleosome of budding yeast is perfectly positioned and covers the entire centromere. *Proc. Natl. Acad. Sci. USA.* 108:12687–12692.
53. O'Sullivan, J. M., S. M. Tan-Wong, …, N. J. Proudfoot. 2004. Gene loops juxtapose promoters and terminators in yeast. *Nat. Genet.* 36:1014–1018.
54. Ansari, A., and M. Hampsey. 2005. A role for the CPF 3′-end processing machinery in RNAP II-dependent gene looping. *Genes Dev.* 19:2969–2978.
55. Teif, V. B., Y. Vainshtein, …, K. Rippe. 2012. Genome-wide nucleosome positioning during embryonic stem cell development. *Nat. Struct. Mol. Biol.* 19:1185–1192.

Supplementary Information

Topological Polymorphism of the Two-start Chromatin Fiber

Davood Norouzi and Victor B. Zhurkin

A: Calculating optimal geometry of the inter-nucleosomal linkers

Finding conformation of the DNA linker connecting two nucleosomes is the special case of a general polymer chain closure problem. This problem has been solved for the all-atom models of the polypeptide [S1] and nucleic acid [S2] chains previously. Here, because of an extremely large size of the system, we use a 'mesoscopic' approach [S3] where DNA is modeled at the level of dimeric steps, and its trajectory is described by the six base-pair step parameters Twist, Roll, Slide, etc. [S4]. The main idea remains the same, however – one has to join the two chain ends. To build the linker connecting nucleosomes ν1 and ν2, we start at the exit point of ν1 (Fig. 1A) and, for any given set of the DNA parameters, generate positions of the L base pairs of the linker plus one 'virtual' base pair #(L+1). Our goal is to find the values of DNA parameters that bring the base pair #(L+1) at the end of the linker in the same position and direction as the base pair #1 at the entry point of ν2 (Fig. 1A).

To this aim, the penalty function (evaluating the difference between the positions of the two base pairs) is calculated as the sum of squares of the distances between the ends of the reference vectors X, Y and Z for the two base pairs. This penalty is added to the elastic energy of the linker DNA [33] and the net function is minimized using a standard method of numeric minimization (described in detail earlier [S5]). Typically, in the end of minimization the distances between the ends of the two sets of the vectors X, Y and Z are less than 0.05 Å. The DNA linker minimization is nested in an outer cycle in which the total energy of the nucleosome fiber is minimized as a function of the four parameters defining the fiber configuration.

B: Energy terms

(I) DNA Elastic energy. The elastic energy of the linker DNA deformation is calculated using the knowledge-based potential functions introduced by Olson et al. [33]. The stiffness constants, including the cross correlations (such as Twist-Roll) are taken as averages for all 16 dinucleotides. As the rest-state values we use the average helical parameters of B-DNA: Twist = 34.5° and Rise = 3.35 Å (the other rest-state values, such as Slide, are taken to be zero).

(II) Electrostatic energy. The electrostatic energy is calculated using the Coulomb potential with 30 Å distance cutoff and the water dielectric constant ε = 78. (In addition, we made computations with the Debye-Huckel potential; see section D) We assume that it is the DNA surface which is the subject to salt screening, whereas most of the histone charges are buried inside nucleosomes and don't get screened. We chose partial charges in such a way that the nucleosomes remain 'slightly' negatively charged, which is consistent with electrophoresis experiments [S6]. The centers of charges considered in our calculations are: Cz, Nz in Arginine and Lysine with corresponding partial charge +1; Cd, Cg in Glutamate and Aspartate with partial charge –1, and the P atoms in nucleosomal DNA with partial charge –0.3. This level of neutralization is predicted in numerical computations [S7].

The long and flexible tails of H3 histones are cut away in our model, but their effect has been taken into account implicitly. According to the "coarse grained" MC simulations [30] the positively charged H3 tails are likely to align along the linkers, resulting in a significant neutralization of linker DNA. Therefore, the linker DNA was modeled with the partial charges –0.25 per nucleotide. (In addition, we varied these charges from –0.05 to –0.5; see section D.)

(III) Steric clashes. Steric clashes are modeled by a van der Waals-like repulsion potential. All the centers of charges considered above are included here, as well as the centers of the DNA base pairs. This is necessary because the P-P distances are relatively large when measured across the major groove, and it might happen that the nucleosomal DNA and the linkers penetrate if the base pairs are not considered. The van der Waals radii are assumed to be 3.0 Å for the centers of charges and 8.0 Å for the DNA base pair centers. The repulsion potential is calculated as

$$E_{vdW} = 10.0 \times \sum_{ij} \left(\frac{\sigma_i + \sigma_j}{r_{ij}}\right)^{12}$$

where σ_i , σ_j are the van der Waals radii, and r_{ij} is the distance between the corresponding pseudo-atoms.

(IV) H4 tail – acidic patch interactions.

Our model of the H4 tail – acidic patch interactions is based on three assumptions:

(i) To stabilize the inter-nucleosome stacking, the positively charged Lys16 (H4) of one nucleosome has to be in the immediate vicinity of the 'acidic patch center' of adjacent nucleosome [S8].

(ii) The acidic patch center coincides with Cd atom of Glu61 (H2A).

(iii) The flexible H4 tail can rotate freely around the hinge located at Asp24 (H4). This is based on comparison of H4 conformations in five crystal structures (Table S1 and Figure S1). In particular, the

distance between Asp24 (the presumed location of the hinge on H4 tail) and Lys16 (H4) varies from ~10 to ~35 Å. (In these calculations, the residues Asp24 and Lys16 were represented by their Ca and Nz atoms, respectively.)

Table S1. List of PDB structures used to calculate the Asp24 (H4) – Lys16 (H4) distances

PDB ID	Minimal distance (Å)	Maximal distance (Å)
T1- 1KX5	11	32
T2- 1AOI	19	24
T3- 1ZBB	20	29
T4- 3UT9	24	29
T5- 4KUD	19	27

References:

[T1] Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., Richmond, T.J. (2002) *J. Mol. Biol.* **319:** 1097-1113

[T2] Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., Richmond, T.J. (1997) *Nature* **389:** 251-260

[T3] Schalch, T., Duda, S., Sargent, D.F., Richmond, T.J. (2005) *Nature* **436:** 138-141

[T4] Chua, E.Y.D., Vasudevan, D., Davey, G.E., Wu, B., Davey, C.A. (2012) *Nucl. Acids Res*. **40:** 6338-6352

[T5] Yang, D., Fang, Q., Wang, M., Ren, R., Wang, H., He, M., Sun, Y., Yang, N., Xu, R.M. (2013) *Nat. Struct. Mol. Biol.* **20**: 1116-1118

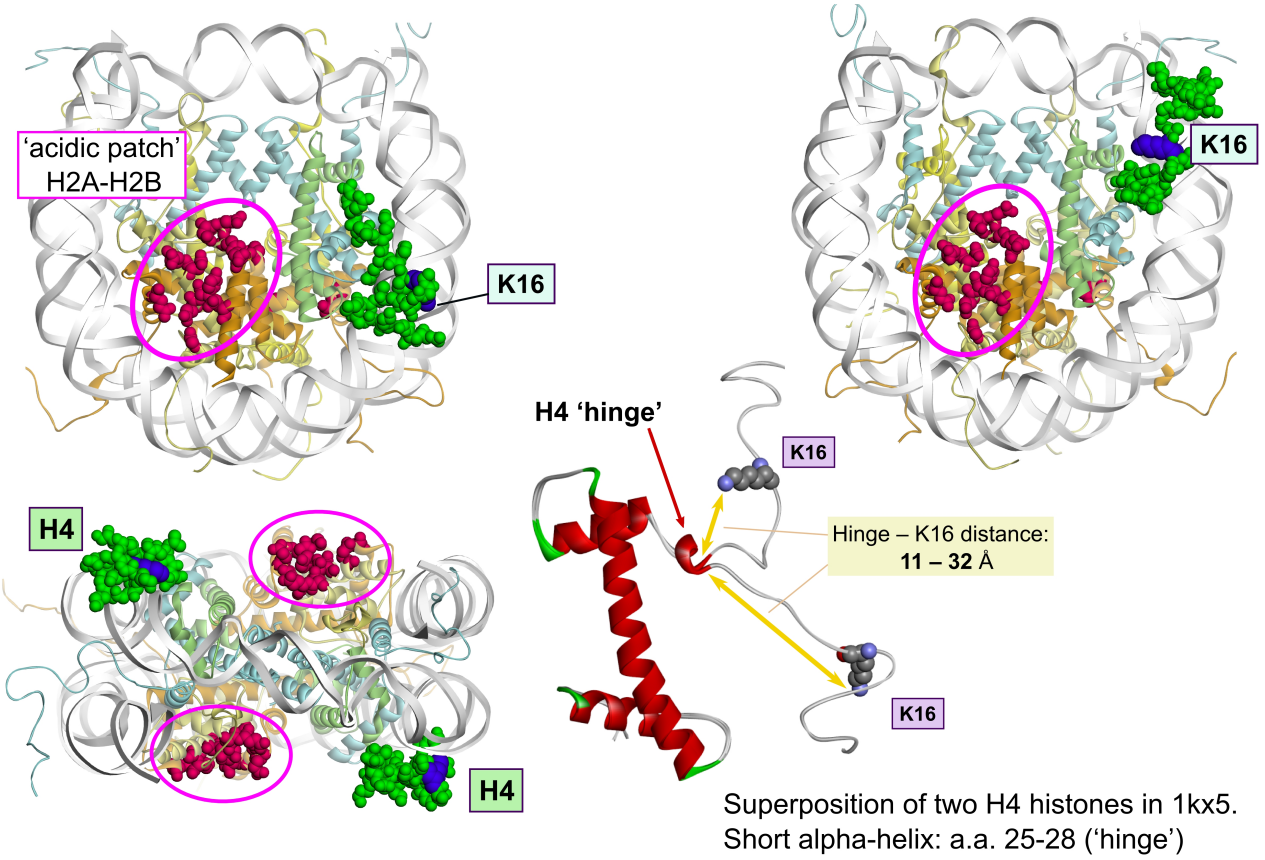

Figure S1. Two conformations of the H4 tail in nucleosome 1kx5.

The top view (left) and the bottom view (right) demonstrate different conformations of the H4 histone N-tail, in green (chains B and F, respectively). The acidic patches (histones H2A-H2B, red CPK atoms) are indicated by pink ellipses.

Center: Superposition of the two H4 chains shows a significant difference between the two N-tail conformations. At the same time, the short α-helix (residues 25-28) retains practically the same conformation in both H4 chains. Therefore, we assume that the N-tail can rotate freely around the 'hinge' located at the N-end of this α-helix, namely, at Asp24 (H4). Note that the distance between the positively charged Lys16 (atom Nz) and the hinge at Asp24 (atom Ca) varies from 11 to 32 Å in the two conformations of histone H4.

Bottom left: The side view of nucleosome 1kx5 [1], with the H4 histone N-tails and acidic patches indicated.

The attractive interactions between the H4 tail and the acidic patch are modeled phenomenologically. We calculate the distance x between the hinge, Asp24 (H4), and the patch center, Glu61 (H2A), located on two adjacent nucleosomes. The energy of the tail-patch interaction as a function of the distance x is approximated by the flat-well potential:

$$E_{tail-patch} = C \times \left(\frac{x-d_2}{1+|x-d_2|} + \frac{d_1-x}{1+|d_1-x|} \right)$$

where C = 4.35 kT defines the depth of the potential and d_1 = 10 Å and d_2 = 35 Å are the positions of two walls of the well (Figure S2). The energy calculated in this way corresponds to formation of two 'bridges' between two stacked nucleosomes (Figure S3). Note that in addition to the electrostatic attraction between the basic H4 tails and the acidic patches on H2A/H2B histones, this potential implicitly includes the van der Waals and hydrophobic interactions stabilizing the inter-nucleosome stacking. The depth of the potential, –8 kT, was selected as an intermediate between two experimentally measured values for the inter-nucleosome stacking, –3.4 kT [49] and –13.8 kT [8]. In addition, we made computations with the stacking energy –5 kT, and found that the main results (such as the positions of the total energy minima shown in Figure 3) remain practically the same. Figure S4 shows how the total energy varies by the changes in the depth of the stacking potential.

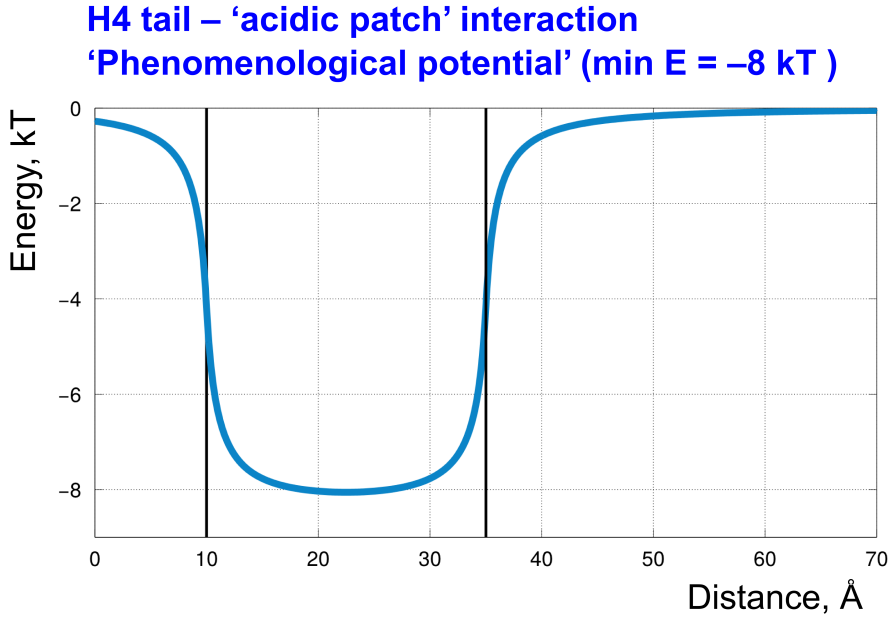

Figure S2. Phenomenological potential describing the H4 tail – acidic patch interactions.
Note that the interval between two walls of this flat-well potential, d_1 = 10 Å and d_2 = 35 Å, covers the minimal and maximal distances Asp24 (H4) – Lys16 (H4) presented in Table S1.

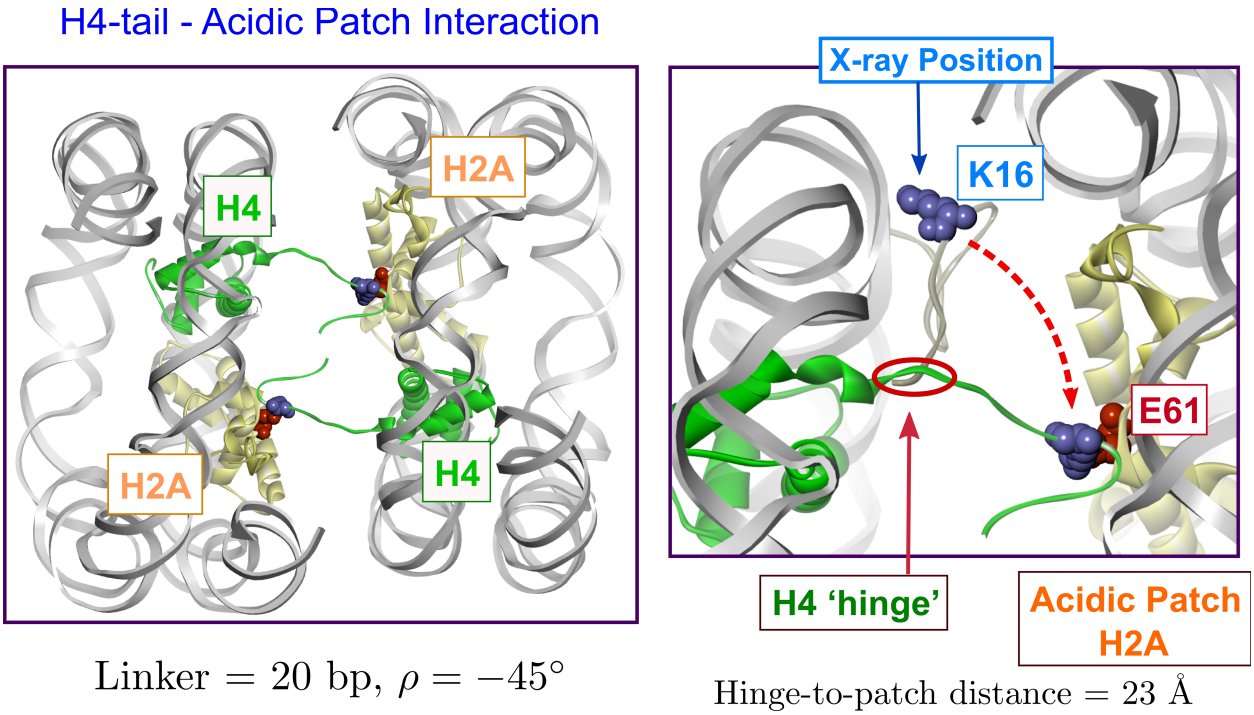

Figure S3. Visualization of the H4 tail – acidic patch 'bridges.'

Left: Two H4 – H2A 'bridges' stabilize the inter-nucleosome stacking. The optimal configuration of the nucleosome fiber with linker L=20 bp is shown, with the inclination angle $\rho = -45°$.

Right: The H4 tail rotates around the 'hinge' to bring the positively charged Lys16 (H4) in the close vicinity of the 'acidic patch' of adjacent nucleosomes. In this particular structure, the distance between the 'hinge' at Asp24 (H4) and the patch center at Glu61 (H2A) is 23 Å. Note that this distance corresponds to the minimum of the potential presented in Figure S2.

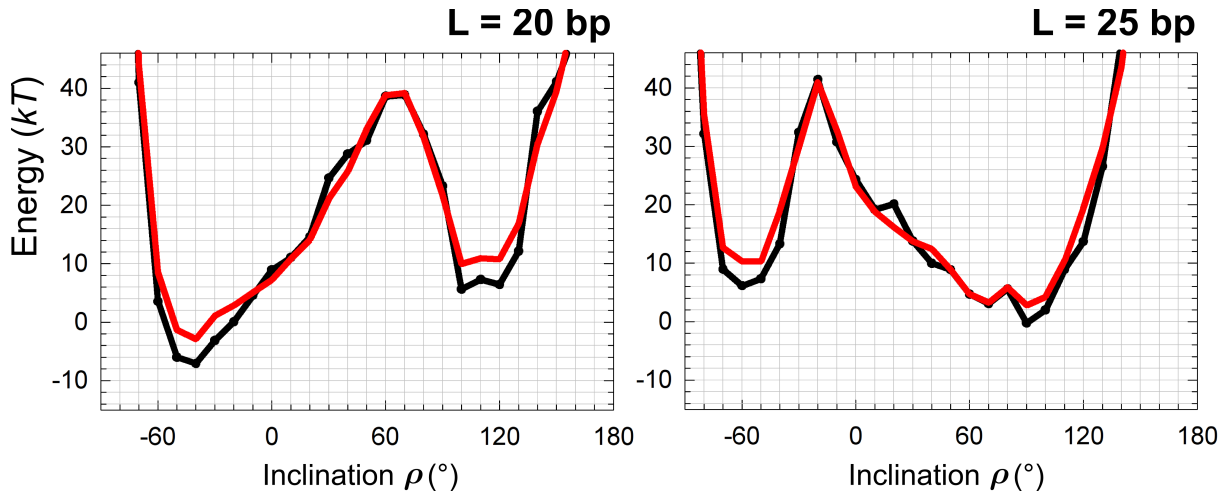

Figure S4. Total energy profiles for two H4 tail – acidic patch interaction potentials.

The black curves correspond to stacking energy –8 *kT* used in the main text (compare with Figure 3). Changing this value to –5 *kT* (red curves) increases the energy values in local minima. Note that positions of the minimal points and general form of energy profiles do not change.

C: Total energy minimization

The energy terms described above are calculated per fiber asymmetric unit, consisting of the nucleosome core and the 'downstream' linker (Fig. 1A). The energy of interaction between units #i and #j is denoted by E(i,j), and the internal energy of unit #i is denoted by E(i). Then, the total energy per unit can be written as E_{total} = E(i) + E(i,i+1) + E(i,i+2). The interactions with units #i+3, #i+4, etc. are ignored due to the distance cutoffs. Interactions with the 'preceding' nucleosomes are not considered due to the symmetry and regularity of the fiber. For example, interaction E(i-1,i) is equivalent to E(i,i+1) and it is assigned to the unit #i-1.

During minimization of the total energy, for each selected set of the four superhelical parameters (Fig. 1), the linker DNA is optimized as described above. The DNA elastic energy is calculated during this cycle of minimization. Another, outer cycle of minimization, is used to optimize the total energy of the fiber. To make sure that the optimized conformations are not just randomly selected local minima, ~300,000 points in the 4D space of fiber parameters were considered as starting points.

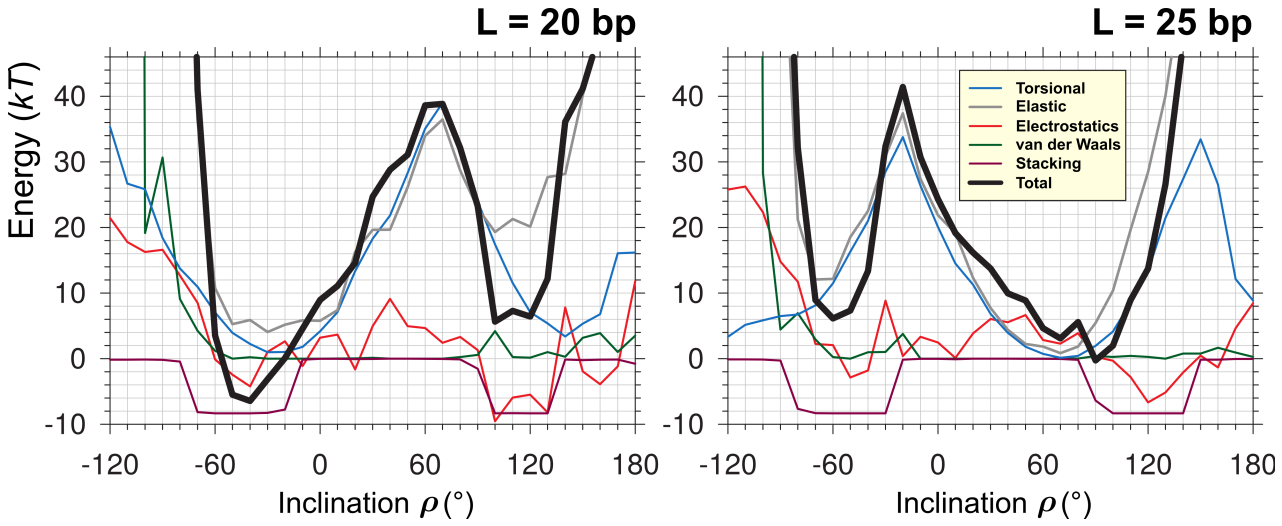

Figure S5. Total fiber energy as a function of inclination angle ρ. Optimization is made in the space of three remaining fiber parameters, i.e. radius, rise, and polar angle φ (Fig. 1). All energy terms are shown separately. The torsional energy profiles (blue curves) have two approximately equal minima separated by topological barriers. Adding other energy terms changes the minima depths and positions, but bimodality of the energy profiles remains.

For example, adding the van der Waals (green) and electrostatic (red) energy terms makes the structures with $\rho < -90°$ and $\rho > 130°$ extremely unfavorable. Furthermore, electrostatic interactions are repulsive for the inclination angle in the interval $30° < \rho < 60°$, therefore, the 'right' minimum for L=25 bp is shifted to $\rho = 90°$ (compared to the torsional energy minimum at $\rho = 70°$). Finally, the regions of optimal stacking interactions (purple curves) are nearly the same for both L=20 and 25 bp, $-70° < \rho < -30°$ and $90° < \rho < 130°$.

D: Electrostatic Interactions

We used two ways to account for the electrostatic interactions in the nucleosome fibers. One approach (presented in the main text) is using Coulomb potential with distance cutoff 30 Å, along with the charge reduction due to the counterion screening. The second approach is using Debye-Huckel potential with the (in vitro) Debye length 8.3 Å [31]. Comparison between energy profiles obtained by the two methods is shown in Figure S6. Importantly, we found that changing the details of electrostatic potential does not change qualitatively the total energy profiles.

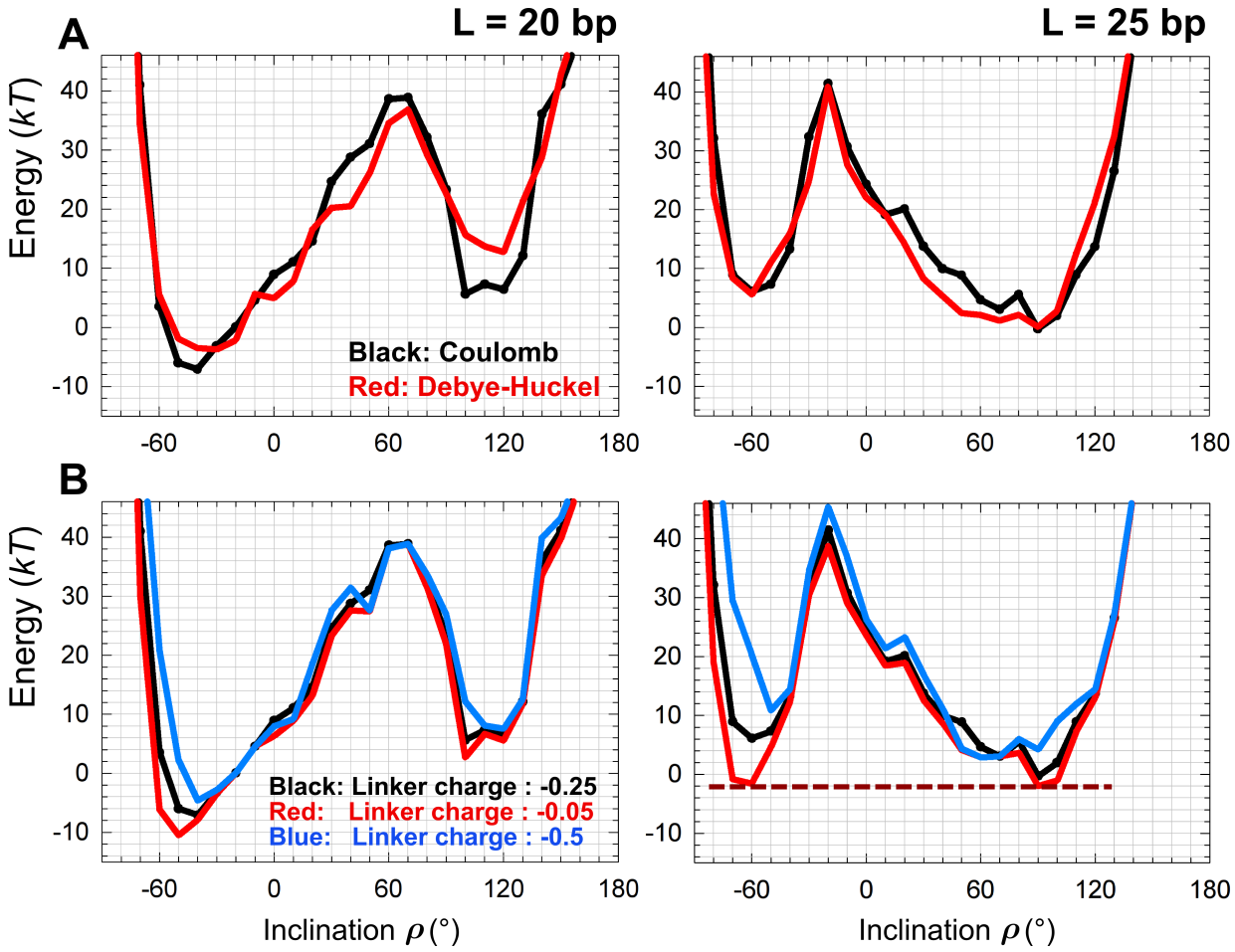

Figure S6. Dependence of the total energy profiles on electrostatic potential.

(A) Black curves correspond to Coulomb potential and red curves correspond to Debye-Huckel potential. Charge on the linker DNA is –0.25 per phosphate (as in the main text).

(B) Effect of the linker DNA charges on the relative stability of the fiber topoisomers (Coulomb potential). Black curves reproduce the main text results, for the linker DNA charge –0.25 per phosphate, blue curves are for the charge –0.5 per phosphate (weak DNA neutralization), and red curves are for the charge –0.05. In the latter case the linker DNA neutralization is the strongest; it may be caused, for example, by linker histones or histone H3 tails (Figure S8). Note that the energy profile is more affected for L=25 than for L=20 bp. In particular, when the charge density is the smallest, –0.05, the left minimum at $\rho = -60°$ becomes as deep at the right one at $\rho = 90°$ (L=25 bp). Figure S7 shows why electrostatic effects are more pronounced in the T2 structures with $\rho < 0$.

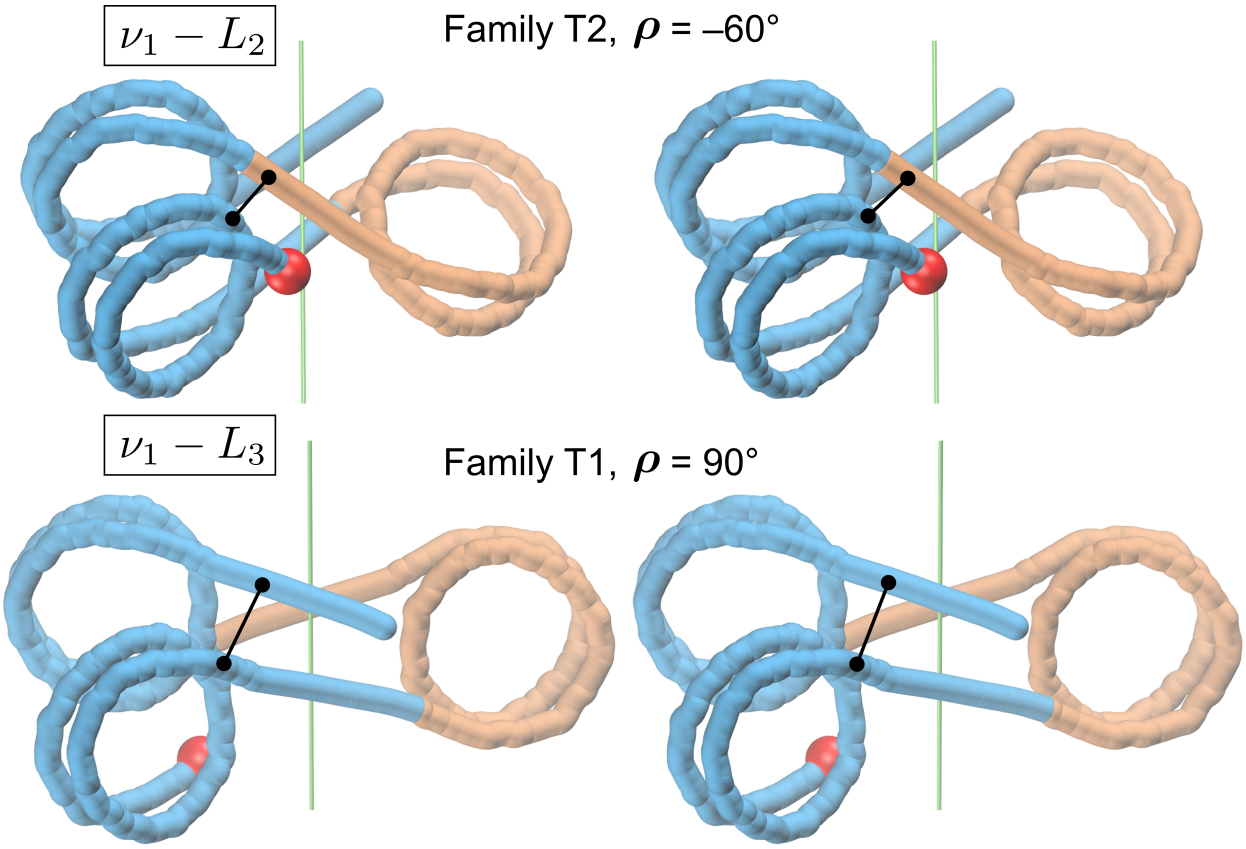

Figure S7. Electrostatic DNA-DNA interactions in the optimal fiber structures with L=25 bp. The T2 (top) and T1 (bottom) conformations are presented in stereo. The closest distances between the core and linker DNA are shown by black segments: in the T2 structure, the distance is between the ν1 core DNA and the linker L2, while in the T1, it is between ν1 and L3. The distances for the T2 structure are shorter than for the T1 (~30 and ~40 Å, respectively); therefore electrostatic repulsion is stronger in T2. This difference explains the result presented in Figure S6B – the decrease in the DNA charge density deepens the left energy minimum (family T2, $\rho < 0$) more than the right minimum (family T1, $\rho > 0$; L=25 bp). The larger DNA-DNA distances in the T1 conformations would make the linker DNA more accessible for various transcription factors. The red balls denote entry points of ν1, see Figure 1. The green vertical lines indicate the fiber axes.

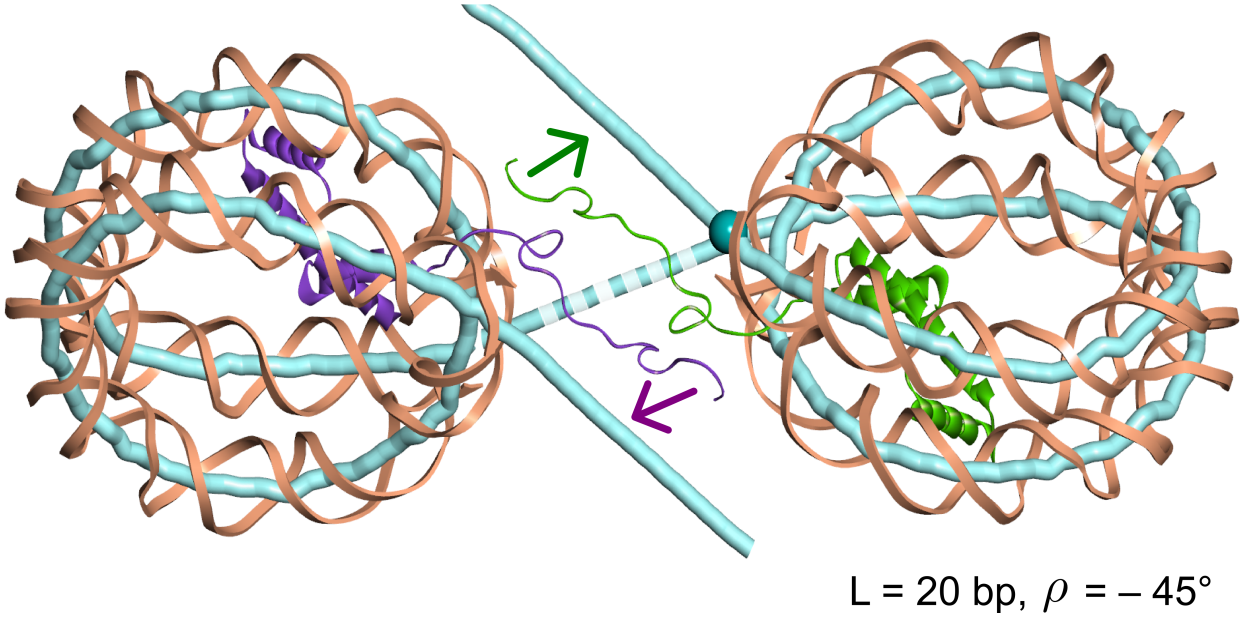

Figure S8. H3 tail – linker DNA interaction.

Mutual orientation of the shown nucleosomes corresponds to the optimal fiber conformation for L=20 bp (compare with Figure S9). In both nucleosomes, the histone H3 is represented by the chain A in the 1kx5 structure [1]. In this conformation, the H3 tail 'covers' 10-15 bp fragment of linker DNA. Our model is consistent with that proposed by Perisic et al. [30] based on Monte Carlo simulations. In yeast, where the linker histone is strongly underrepresented compared to the core histones [S9,S10], the shown scheme is likely to be one of the main mechanisms of the linker DNA neutralization.

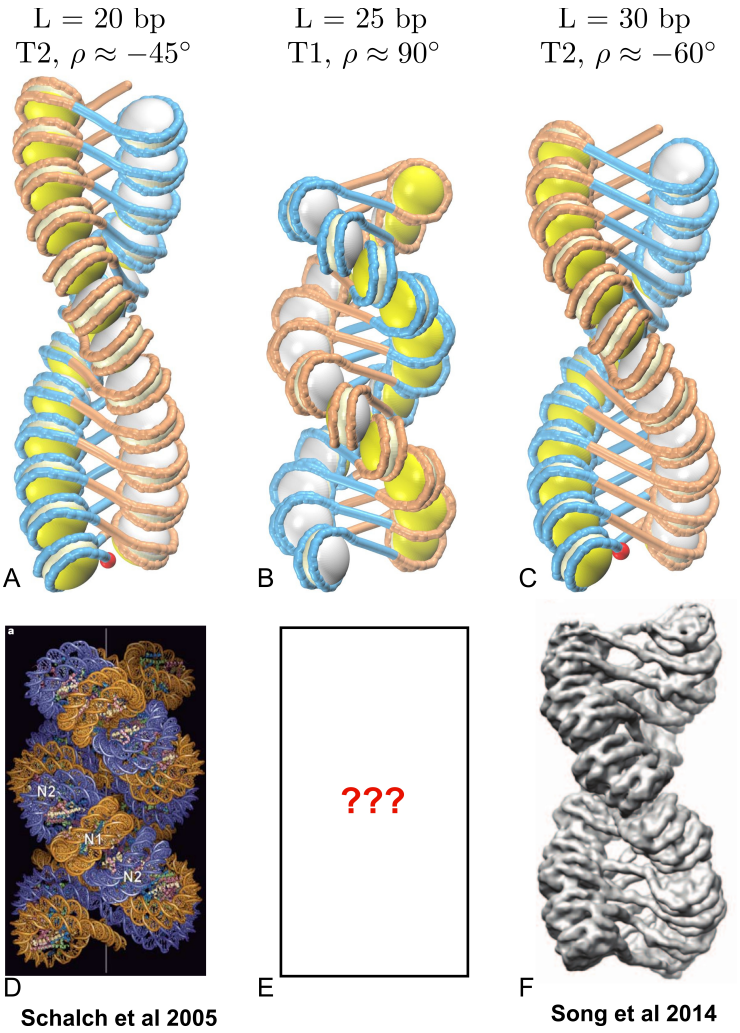

Figure S9. Comparison with the experiment-based models.

Our optimal structures for L = 20, 25, and 30 bp are shown in the first row. In the second row the published experiment-based models are given for comparison. For L=20 bp, our optimal structure has inclination angle –45° and rise = 27 Å (Figure 8). The 'direct' model by Schalch *et al.* [10] is based on the X-ray tetranucleosome structure; it has an inclination angle about –50°/–40° and rather low rise = 17 Å. Note, however, that this model has "steric overlaps" [10]; the authors suggest that these overlaps can be "relieved by increasing the separation of "tetranucleosomes." On the other hand, using the EM images for the '601' nucleosome arrays with NRL=167 [7,12], one can estimate the rise value as 25 Å, which is close to our evaluation of rise = 27 Å. For L=30 bp, Song et al. [11] built their model based on the Cryo-EM data, with the inclination angle –60° and rise = 24 Å. Our optimal structure is very close to this model: $\rho \approx -60°$ and rise = 25 Å. For L=25 bp, our optimal structure has $\rho = 90°$ and rise = 23 Å. Note that there are no experiment-based models for L=25 bp.

E: Symmetrization of the nucleosome structure

To make sure that our results are consistent with the requirement of symmetry of a nucleosome fiber, we used the symmetrized version of the 1kx5 structure [1]. To this aim, the chains E, F, G and H were replaced by the superimposed copies of the chains A, B, C and D (histones H3, H4, H2A and H2B respectively). The symmetrization procedure was made using MatchMaker tool of Chimera [S11].

F: Computation of DNA writhing and linking number

To evaluate the topological changes occurring in DNA upon formation of a nucleosome fiber, we are using the three topological parameters: ΔTw (the change in DNA twisting), DNA writhing, Wr, and the change in the linking number, ΔLk (compared to the relaxed state of DNA), which are related by well-known equation: $\Delta Lk = \Delta Tw + Wr$ [34-38]. This equation is valid for the closed circular DNA, therefore, we need to find an effective way to calculate Wr of DNA packaged in a nucleosome fiber, so that the result does not depend on selection of the closed DNA trajectory.

To calculate the DNA Writhing we use the quadrangle approach described by Levitt [37] and Klenin and Langowski [38]. The DNA trajectory is described as a polygonal chain with the vertex points at the centers of base pairs. The DNA twisting is determined using the Euler angle formalism [S12] implemented in CompDNA [S13] and 3DNA [S14] software. (*– see below.) After generating the fiber as described in the main text we added 25 or 200 extra points connecting the ends of DNA in a way that the closing chain does not pass through nucleosomes (Figure S10). We found that the open 10-nucleosome trajectories and the closed ones (with 25 extra points) produce different average Wr values, varying by as much as 0.23 per nucleosome (Table S2). On the other hand, the open and the closed 100-nucleosome trajectories produce very close Wr values, with the difference not exceeding 0.05. In the case of 200 extra points (connecting the ends of DNA) this difference does not exceed 0.03.

Therefore, we conclude that calculating the DNA writhing for the open 100-nucleosome fibers is an adequate way to estimate the topological changes in DNA upon formation of a nucleosome fiber.

(*) The DNA twisting is computed only for the linkers. Twisting of the nucleosomal DNA may be somewhat different from that of free DNA [S15], but this difference is not essential for our purpose, because it remains the same for different fiber structures, as long as the nucleosome cores retain the same conformation.

Table S2

DNA Writhing per nucleosome

Linker (bp)	Angle ρ	10 nsm open	10 nsm closed *	$\delta(Wr)$ 10 nms	100 nsm open	100 nsm closed *	$\delta(Wr)$ 100 nsm	100 nsm closed, 200 points added
20	-70	-1.57	-1.69	0.12	-1.67	-1.71	0.03	-1.69
20	-40	-1.35	-1.58	**0.23**	-1.50	-1.55	**0.05**	-1.53
20	40	-1.09	-1.10	0.01	-1.13	-1.09	0.04	-1.10
20	70	-0.99	-0.92	0.07	-1.00	-0.98	0.02	-1.00
25	-70	-1.57	-1.69	0.12	-1.67	-1.70	0.03	-1.69
25	-40	-1.39	-1.57	0.17	-1.52	-1.55	0.03	-1.52
25	40	-1.09	-1.11	0.02	-1.12	-1.12	0.01	-1.12
25	70	-0.96	-0.92	0.05	-1.00	-0.98	0.02	-0.99

* 25 points added, connecting the ends of DNA in nucleosome fiber (Figure S10).
$\delta(Wr)$ is the difference between the two *Wr* values obtained for the open and closed nucleosome fiber.
The largest values are in boldface underlined.

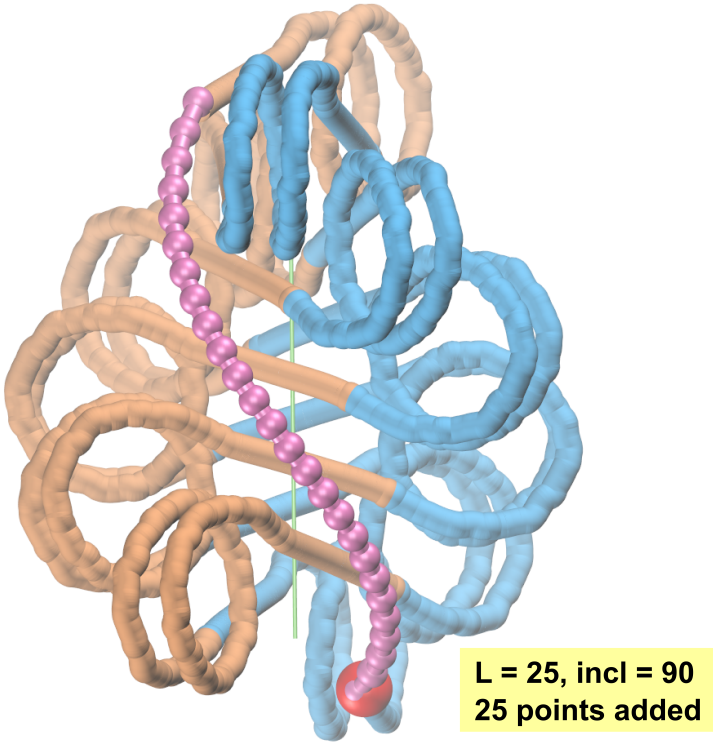

Figure S10. Illustration of the DNA chain closure, used to calculate DNA writhing in the nucleosome fiber. 25 extra points are added (shown in pink) to make the closed DNA trajectory.

G: Movie. Fiber conformation as a function of the inclination angle

The movie "30-nm-Rho-Variation" visualizes changes in the two-start fiber conformation caused by a gradual increase in the inclination angle ρ from –180° to 180° (linker L=20 bp). The increase in inclination angle is accompanied by the clock-wise rotation of the red ball at the bottom, denoting the DNA 'entry point.' Notice over-crossing of the linkers at $\rho \approx -120°$ (shown by ellipses). All the shown structures are optimal for the given inclination (Figure 3A). There are two versions of the movie -- one is for PC users (with extension .avi) and the other is for MAC users (with extension .mov).

H: Right-handed two-start fibers

Small angle X-ray scattering analysis of the nucleosome fibers performed by Williams et al. [13] suggests formation of the left-handed two-start superhelices in solution. The X-ray crystallography and Cryo-EM imaging also support the left-handed organization of two-start fibers [10-11]. To check if our model is consistent with these results, we analyzed the right-handed fibers using the same minimization procedure (described in the main text). To this aim, we assumed that the polar angle φ varies between 180° and 210° (Figure 1). (Note that for the left-handed two-start fibers the angle φ varies between 150° and 180°.)

We found that the optimal energy for the right-handed helices is significantly higher than for the left-handed ones (Figure S11). This energy difference is more pronounced for L=25 bp (~12 kT) than for L=20 bp (~6 kT). As follows from the comparison of the two types of fibers, the linker-nucleosome clashes prevent formation of a strong stacking between nucleosomes in the right-handed superhelices. In summary, we conclude that the chirality of nucleosomes dictates the left-handedness of the chromatin fiber.

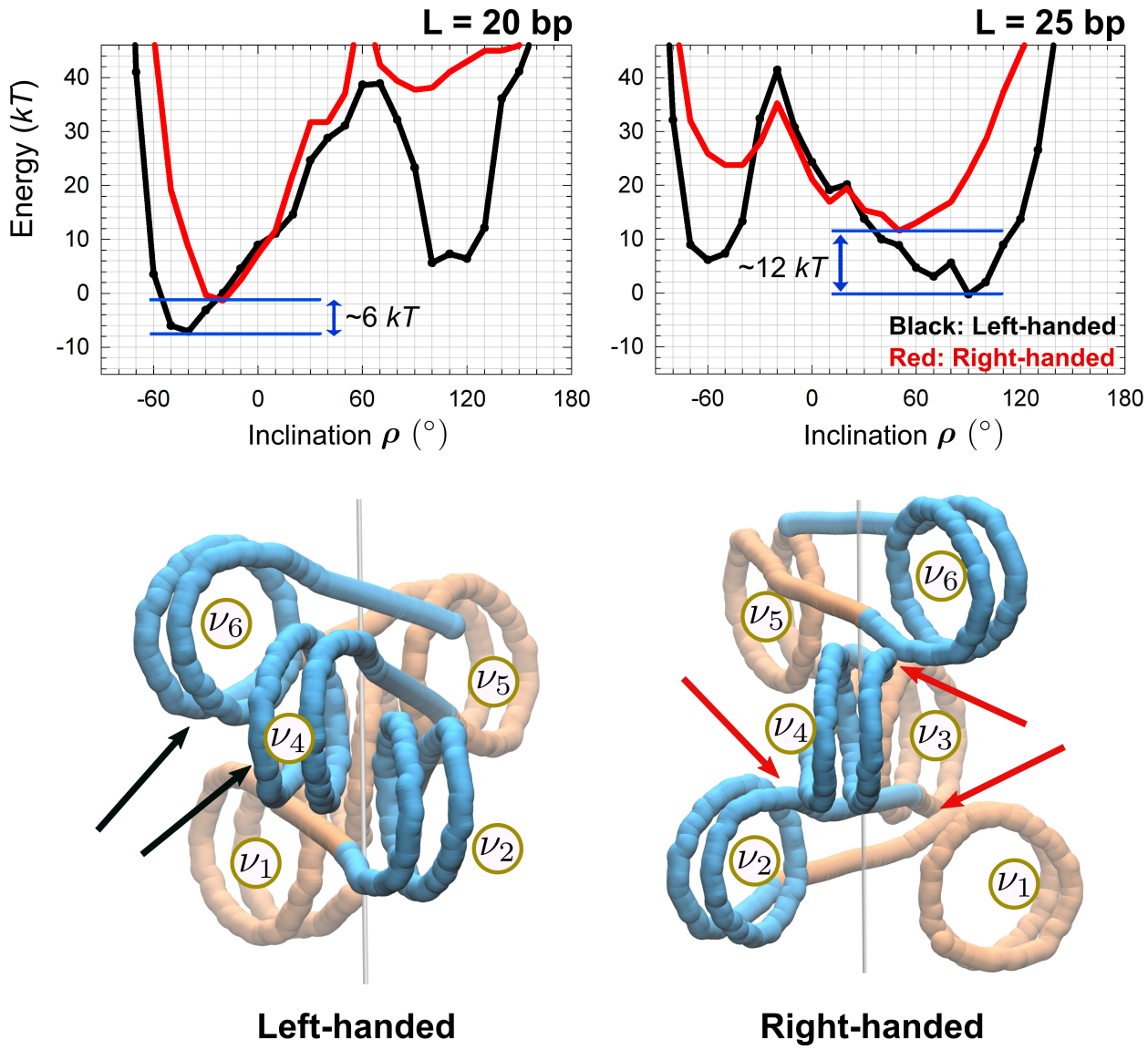

Figure S11. Comparison between the left-handed and right-handed nucleosome fibers.

Top. Total energy profiles for the left-handed (black curves) and right-handed (red curves) fibers. The minimal energy is higher in the right-handed structures, mostly due to electrostatic repulsion between the core DNA and the linkers (shown in the bottom). Note that the right minimum (at $\rho \approx 90°$) is influenced more by change in the fiber handedness.

Bottom. Optimal structures with $\rho = 90°$, L = 25 bp are shown. The left-handed fiber is tightly packed with a strong inter-nucleosome stacking (black arrows) while the right-handed fiber needs a large rise to avoid clashes (close contacts shown by red arrows); as a result, the stacking is disrupted. The nucleosomes are numbered to clarify their connectivity in fibers.

I: Linking number in chromatin fibers with various linker lengths

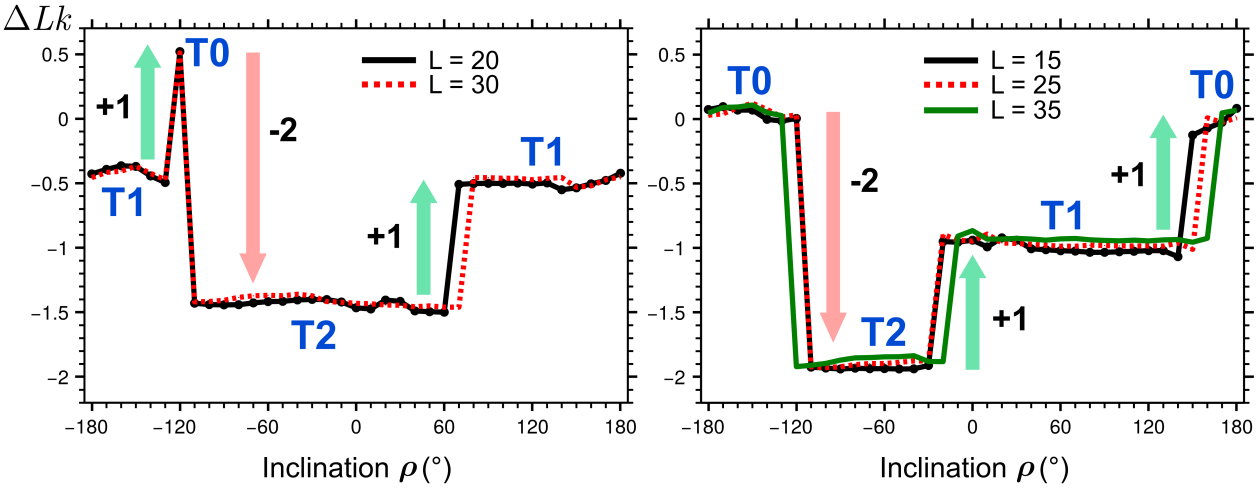

Figure S12: Linking number per nucleosome as a function of inclination angle ρ for selected linker lengths, L. *Left:* the data for L = 10n (20 and 30 bp); *right:* the data for L = 10n+5 (15, 25, and 35 bp). The linking number is (almost) piecewise constant in the three regions corresponding to the three topological families, T2, T1 and T0 (see Figure 7D). The T2 family consists of the fiber conformations with the lowest ΔLk values (from approximately −2.2 to −1.5). The fibers with intermediate ΔLk (from −1.2 to −0.5) comprise family T1. The family T0 corresponds to the highest ΔLk values (from approximately 0 to 0.5). Note that for L = 20 and 30 bp, the T0 family is represented by a single point $\rho = -120°$ (in this particular presentation with the ρ increment of 10°).

The inter-family transitions are accompanied by abrupt changes in ΔLk values shown by red arrows (decrease in ΔLk) and green arrows (increase in ΔLk). The sharp transition between the families T0 and T2 (red arrows) is accompanied by $\Delta(\Delta Lk) \approx -2$ and occurs around $\rho = -120°$ for all linkers. By contrast, the T2-T1 and T1-T0 transitions occur at different ρ values depending on the linker length, L; these transitions are accompanied by $\Delta(\Delta Lk) \approx +1$. The structural nature of these transitions is discussed in the Results section, see Figure 7.

Using the ΔLk values, we can calculate the superhelical density of DNA, σ, defined as (number of superhelical turns) divided by (number of turns of DNA in relaxed state, N_r). Normalizing both values per one nucleosome, we have $\sigma = \Delta Lk / N_r = \Delta Lk /$ (NRL/10.45), where NRL is nucleosome repeat length, and 10.45 is the average number of DNA base pairs per turn.

The energetically optimal fibers are characterized by the values:

L = 20 bp, NRL =167 bp, $\Delta Lk = -1.5$ (Figures 7C, D), $\sigma = -0.09$
L = 25 bp, NRL =172 bp, $\Delta Lk = -1.0$ (Figures 7C, D), $\sigma = -0.06$

Note that these values of σ are consistent with the experimental measurements for both pro- and eukaryotes [47]; in particular, $\sigma \approx$ -0.06 for *E. coli*. In other words, the torsional stress experienced by DNA if the histones are removed, is comparable with the torsional stress in bacteria.

Supplementary References

S1. Go, N. & Scheraga, H.A. Ring Closure and Local Conformational Deformations of Chain Molecules. *Macromolecules* 1970, 3:178-187.

S2. Zhurkin, V.B., Lysov, Y.P. & Ivanov, V.I. Different families of double-stranded conformations of DNA as revealed by computer calculations. *Biopolymers*. 1978, 17:377-412.

S3. Olson, W. K., Marky, N. L., Jernigan, R. L. & Zhurkin, V. B. Influence of fluctuations on DNA curvature. A comparison of flexible and static wedge models of intrinsically bent DNA. *J. Mol. Biol*. 1993. 232: 530-554.

S4. Dickerson, R. E., Bansal, M., Calladine, C. R., Diekmann, S., Hunter, W. N., Kennard, O. Lavery, R., Nelson, H. C. M., Olson, W. K., Saenger, W., Shakked, Z., Sklenar, H., Soumpasis, D. M., Tung, C -S., Kitzing, E. V., Wang, A. H. -J. & Zhurkin, V. B. Definitions and nomenclature of nucleic acid structure parameters. *EMBO J*. 1989. 8: 1-4.

S5. Kosikov, K.M., Gorin, A.A., Zhurkin, V.B. & Olson, W.K. DNA stretching and compression: large-scale simulations of double helical structures. *J Mol Biol*. 1999, 289:1301- 1326.

S6. Hewish, D.R. & Burgoyne, L.A. Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochem Biophys Res Commun*. 1973, 52:504-510.

S7. Materese, C.K., Savelyev, A. & Papoian, G.A. Counterion atmosphere and hydration patterns near a nucleosome core particle. *J Am Chem Soc*. 2009, 131:15005-15013.

S8. Yang, D. & Arya, G. Structure and binding of the H4 histone tail and the effects of lysine 16 acetylation. *Phys Chem Chem Phys*. 2011, 13:2911-2921.

S9. Freidkin, I. & Katcoff, D.J. Specific distribution of the Saccharomyces cerevisiae linker histone homolog HHO1p in the chromatin. *Nucleic Acids Res*. 2001, 29, 4043–4051.

S10. Downs, J.A., Kosmidou, E., Morgan, A. & Jackson, S.P. Suppression of homologous recombination by the Saccharomyces cerevisiae linker histone. *Mol. Cell* 2003, 11, 1685–1692

S11. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. & Ferrin, T.E. UCSF Chimera -- a visualization system for exploratory research and analysis. *J Comput Chem*. 2004, 25:1605-1612.

S12. Zhurkin, V.B., Lysov, Y.P., & Ivanov, V.I. Anisotropic flexibility of DNA and the nucleosomal structure. *Nucleic Acids Res*. 1979. 6:1081-1096.

S13. Gorin, A. A., Zhurkin, V. B., & Olson, W. K. B-DNA twisting correlates with base pair morphology. *J. Mol. Biol*. 1995, 247: 34-48.

S14. Lu, X. & Olson, W.K. 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat Protoc* 2008, 3:1213-1227.

S15. Richmond, T.J. & C.A. Davey, The structure of DNA in the nucleosome core. *Nature*, 2003. 423(6936): p. 145-50.