Chromatin fiber polymorphism triggered by variations of DNA linker lengths

Rosana Collepardo-Guevara[a] and Tamar Schlick[b,c,1]

[a]Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom; [b]Department of Chemistry, New York University, New York, NY 10003; and [c]Courant Institute of Mathematical Sciences, New York University, New York, NY 10012

Deciphering the factors that control chromatin fiber structure is key to understanding fundamental chromosomal processes. Although details remain unknown, it is becoming clear that chromatin is polymorphic depending on internal and external factors. In particular, different lengths of the linker DNAs joining successive nucleosomes (measured in nucleosome-repeat lengths or NRLs) that characterize different cell types and cell cycle stages produce different structures. NRL is also nonuniform within single fibers, but how this diversity affects chromatin fiber structure is not clear. Here we perform Monte Carlo simulations of a coarse-grained oligonucleosome model to help interpret fiber structure subject to intrafiber NRL variations, as relevant to proliferating cells of interphase chromatin, fibers subject to remodeling factors, and regulatory DNA sequences. We find that intrafiber NRL variations have a profound impact on chromatin structure, with a wide range of different architectures emerging (highly bent narrow forms, canonical and irregular zigzag fibers, and polymorphic conformations), depending on the NRLs mixed. This stabilization of a wide range of fiber forms might allow NRL variations to regulate both fiber compaction and selective DNA exposure. The polymorphic forms spanning canonical to sharply bent structures, like hairpins and loops, arise from large NRL variations and are surprisingly more compact than uniform NRL structures. They are distinguished by tail-mediated far-nucleosome interactions, in addition to the near-nucleosome interactions of canonical 30-nm fibers. Polymorphism is consistent with chromatin's diverse biological functions and heterogeneous constituents. Intrafiber NRL variations, in particular, may contribute to fiber bending and looping and thus to distant communication in associated regulatory processes.

coarse-grained modeling | chromatin polymorphism | nonuniform NRL | chromatin bending and looping

The DNA inside eukaryotic nuclei is not found free, but tightly packed along with histone and nonhistone proteins in the form of chromatin structures. Chromatin organization and structural transitions directly impact fundamental cellular processes such as DNA transcription, replication, repair, and recombination. However, our understanding of chromatin structure, how it is regulated by internal and external factors, and the relationship between structure and biological functions remain elusive. The challenge in solving these questions arises from the complex cellular milieu, chromatin's diverse and varying composition, and the limited resolution of experimental methods for large systems.

Chromatin consists of a repeating sequence of nucleoprotein blocks (or nucleosomes) joined by DNA linker segments. The nucleosome structure is well understood at atomic resolution (1, 2). Its histone protein octamer (two copies each of H2A, H2B, H3, and H4) has ~147 bp of DNA wrapped around it (1) and 10 highly positively charged and flexible tails (two N-terminal domains from each histone dimer and two C-terminal domains from H2A) that mediate interactions with other nucleosomes and the DNA (2).

At low salt concentrations, due to the electrostatic repulsion among DNA linkers, chromatin forms an extended 10-nm wide structure that resembles beads on a string (3). At physiological salt concentrations (100–150 mM NaCl or ~2 mM Mg^{2+}) and in the presence of linker histone (LH) proteins, the extended

conformation is thought to fold into a compact 30-nm wide fiber (4). However, in over three decades of research more questions than answers have arisen concerning various models for compact chromatin [e.g., zigzag (5–8), interdigitated solenoid (9–11), and heteromorphic (12) fibers], not to speak of the actual existence of such a regular higher-order structure (13–16), for instance within interphase chromatin (17) and mitotic chromosomes (18–20).

The realization that chromatin constituents across and within organisms are highly heterogeneous (e.g., in linker DNA length, LH concentration, histone composition, and histone tail modifications) has led to an irregular chromatin fiber architecture (21). This is especially striking for different ionic conditions, where high monovalent salt and a low concentration of divalent ions, as found in vivo, produce an irregular heteromorphic structure (12) blending features of both zigzag (straight DNA linkers) and solenoid (bent DNA linkers).

This ability of the chromatin fiber to adopt a variety of forms is essential to its diverse biological functions. For example, during gene regulation, enhancer and silencers—DNA regulatory elements that activate or repress transcription of their target genes—function at a distance via formation of special chromatin loops (22, 23). Tail-mediated inter- and intramolecular interactions are well known to be important for distant communication in chromatin.

One of the principal factors known to alter the structure of the chromatin fiber is the nucleosome-repeat length (NRL) (24), defined as the wrapped nucleosomal DNA (147 bp) plus the variable linker DNA length; this is due to the NRL changing the spatial organization of successive nucleosomes and the distance between neighboring cores. The average NRL varies across species, tissues, and cell cycle states, ranging from short-to-medium values of ~154 to 189 bp in transcriptionally active cells, to medium-to-long values of ~190 to 240 bp in mature transcriptionally inactive states. Electron microscopy (EM) measurements revealed that a short NRL (167 bp) leads to narrow fibers (21-nm diameter) with a clear zigzag topology, whereas a medium NRL (197 bp)

Significance

The structure of the chromatin fiber remains one of the most fundamental open biological questions because structure dictates many template-directed processes. We use coarse-grained modeling to investigate systematically how variations in the linker DNA length that arise naturally for chromatin in different tissues, species, and cell cycle stages affect fiber architecture. We unravel a natural source of fiber polymorphism, in which irregular interdigitated 10-nm and compact 30-nm fibers coexist. These results suggest how structural diversity can be advantageous for gene regulation activity. Interesting DNA design applications also arise.

BIOPHYSICS AND COMPUTATIONAL BIOLOGY

forms highly compact 30-nm interdigitated solenoid structures (10). All-atom modeling using steric and energetic considerations suggested a wide range of topologies (e.g., one-, two-, and three-start) as a function of NRL (25). Our coarse-grained modeling (26) showed that short NRLs (<182 bp) produce narrow ladder-like forms, medium NRLs (~191–209 bp) zigzag fibers, and long NRLs (>218 bp) heteromorphic fibers. A lower-resolution coarse-grained approach, using the two-angle model for the DNA geometry and a Gay–Berne potential to account for internucleosome interactions, showed that NRLs ranging from 155 to 211 bp produce one-, two-, and three-start forms (27). Besides fiber structure, in vitro and in silico force extension experiments of single chromatin fibers (28) have shown that the NRL also alters the resistance of fibers to unfold (11, 29) and associated unfolding pathways (29, 30).
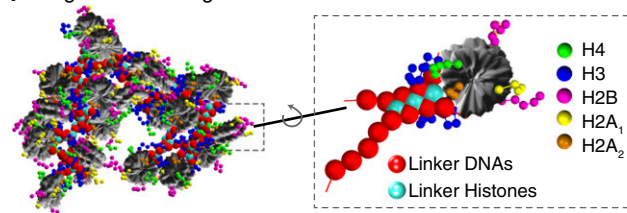
In vivo, the linker DNA length is also nonuniform within single fibers (31, 32). This is especially relevant to proliferating interphase chromatin, to fibers subject to nucleosome wrapping/unwrapping and sliding, and to regulatory DNA sequences. However, how exactly this intrafiber variability impacts chromatin structure remains unclear. Although recent experiments have shown that small NRL deviations (±2 and ±4 bp from the mean repeat) do not change significantly chromatin's folding and compaction (32), it has long been speculated that larger differences (~±10 bp) would lead to polymorphic chromatin organization (8). For example, silencers that flank a *HMR* loci in yeast are connected by a 12-nucleosome nonuniform NRL array (with 5- and ~20- to 30-bp DNA linkers in regular alternation) and are thought to interact through the formation of a chromatin loop (33). Here we explore the effect of a wide range of intrafiber NRL variations in chromatin fiber structure through Monte Carlo (MC) simulations of our mesoscale chromatin model (26, 29, 30, 34–36). Our results identify a remarkable effect of nonuniform NRLs in the organization and compaction of chromatin that introduces polymorphism and more compact, rather than more open, overall fibers. These conformations include distant, tail-mediated chromatin loops and highly bent arrangements essential for distant communication between regulatory elements. Our results support the idea that a polymorphic chromatin fiber is more compatible with the heterogeneous conditions found in vivo. Interesting DNA design implications also arise.
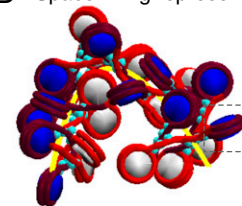
Results

To determine how linker DNA intrafiber variations affect chromatin fiber behavior, we sample by MC 24-core nonuniform NRL oligonucleosomes with LHs (one LH permanently attached to each core) at physiological conditions (0.15 M monovalent salt and room temperature) using our coarse-grained chromatin model (Fig. 1 and Fig. S1), extensively validated against experiments and refined over the past decade (26, 29, 30, 34–36) (*SI Text*). Each nonuniform NRL oligonucleosome contains a combination of two NRLs in regular alternation (Fig. 1*C*). We divide the NRLs we study into short (173 bp), medium (182–209 bp), and long NRLs (218 and 227 bp) (26) and classify the size of the NRL variation from the mean repeat length as moderate (±4.5–9 bp) and large ($\geq\pm9$ bp). Our model can describe NRLs between 173 and 227 bp, as well as explore a wide range of fibers selected to span different combinations of short, medium, and long NRLs and intrafiber NRL variations between ±4.5 and ±27 bp. The full set of fibers studied is listed in Table 1.

Below we describe the major effects of intrafiber NRL variations in fiber structure and compaction. Different NRL combinations give rise to significantly different fiber architectures, that we term bent ladder, canonical, and polymorphic fibers (Table 1). Fig. 2 shows the representative forms according to the NRL variation, and emphasizes the emergence of fiber polymorphism. Fig. 3 quantifies the internal fiber organization through the frequency of near and far-neighbor contacts (see *SI Text*). Finally, Fig. 4 quantifies changes in fiber compaction by measuring overall packing ratios versus the average NRL. We show, surprisingly, how
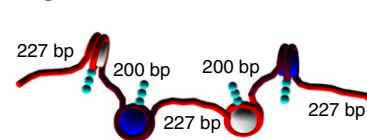

Fig. 1. Representation of the integrated coarse-grained oligonucleosome model with nonuniform NRLs ($NRL_1 = 200$ bp and $NRL_2 = 227$ bp) in regular alternation. (*A*) Nucleosome (with DNA wrapped around) with its irregular surface in gray. (*A, Inset*) The histone tail beads are in green (H4), blue (H3), magenta (H2B), yellow (H2A$_1$, N-termini), and orange (H2A$_2$, C-termini); the LH beads are in turquoise; and the linker DNA beads are in red. (*B*) Space-filling view without tails with alternating DNA shown in red and dark red, successive nucleosomes in blue and white, LHs in turquoise, and fiber axis in yellow. (*C*) Extended conformation illustrating regular alternation of NRLs (NRL_1–NRL_2–...).

larger NRL variations enhance compaction; reference values for uniform NRL fibers are also shown (26).

Bent Ladders for Nonuniform NRL Fibers with One Short Linker DNA. Nonuniform NRL chromatin fibers with one short linker DNA (Table 1; bent ladders) are confined to adopt a ladder-like organization as shown in Fig. 2*A*. Their internucleosome interaction patterns (Fig. 3*A*) reveal the dominant $i \pm 2$ and $i \pm 1$ contacts characteristic of ladder-like structures (e.g., 173-bp uniform fiber); note that the 173- to 227-bp structure is defined instead by peaks at $i \pm 1$ and $i \pm 3$ because of the nonaligned ladder organization imposed by one short and one long NRL.

We term these structures "bent ladders" because they exhibit a remarkably large fiber axis bending; that is, up to 76% higher than uniform NRL structures (Fig. S2*A*). This fiber axis bending yields diverse conformations that resemble side-by-side 10-nm structures (spanning full loops, open/twisted circles, and hairpin forms, among others) and exhibit a large occurrence (more than 40%) of long-range far-neighbor interactions (Fig. 3*B*). Fiber axis bending is favored by the lack of rigid DNA stems. DNA stems are formed when the two LH molecules bound to successive nucleosomes establish contacts with their entering and exiting DNA linkers (37). DNA stems reorganize chromatin because they straighten the linker DNAs and reduce the separation angle between entering and exiting DNA. DNA stems form only when both the entering and exiting linker DNAs are long enough to screen the two LHs in the stem (40 bp approximately); bent ladders cannot form stems because the 26-bp linker DNA is shorter than this length.

The compaction of these bent ladders (Fig. 4 and Fig. S3) is comparable to that of narrow and loose uniform NRL counterparts with short DNA linkers as observed by experiments (10) and simulations (26). As the linker variation between the short and the other linker increases, the packing ratio and sedimentation coefficients significantly decrease. Limited compaction of bent-ladder fibers is due to hampered nucleosome reorganization by short linker DNAs, as shown by EM measurements of 172-bp arrays (32) and the increased stiffness of short versus medium NRL arrays in force extension studies (11, 29). Such loose structures with just 1.2–3.6 nucleosomes per turn have been observed in

Table 1. Characteristics of nonuniform NRL fibers (in base pair units) by fiber type

Fiber	NRL$_1$*	NRL$_2$*	Average NRL*	±$\Delta N^{\dagger}$
	Bent ladders (highly bent 10-nm forms)			
1	Short: 173	Med: 182	Short: 177.5	Mod: 4.5
2	Short: 173	Med: 209	Med: 191	Large: 18
3	Short: 173	Long: 227	Med: 200	Large: 27
	Canonical (irregular zigzag/heteromorphic fibers)			
4	Med: 182	Med: 191	Med: 186.5	Mod: 4.5
5	Med: 182	Med: 200	Med: 191	MtoL: 9
6	Med: 191	Med: 200	Med: 195.5	Mod: 4.5
7	Med: 191	Med: 209	Med: 200	MtoL: 9
8	Med: 200	Med: 209	Med: 204.5	Mod: 4.5
12	Med: 209	Long: 218	Med: 213.5	Mod: 4.5
14	Long: 218	Long: 227	Long: 222	Mod: 4.5
	Polymorphic (from canonical to compact interdigitated fibers)			
9	Med: 191	Long: 218	Med: 204.5	Large: 13.5
10	Med: 200	Long: 218	Med: 209	MtoL: 9
11	Med: 191	Long: 227	Med: 209	Large: 17.5
13	Med: 200	Long: 227	Med: 213.5	Large: 13.5

*NRLs are classified as short, medium (Med), and long.
†Intra-fiber NRL variations from the average repeat length, $\Delta N = (NRL_2 - NRL_1)/2$, are classified as moderate (Mod), intermediate value between moderate and large (MtoL), and large.

transcriptionally active yeast chromatin by analysis of in vivo spatial distances and chromosome conformation capture (3C) (38).

Canonical Fibers for Medium-to-Long Linker DNAs and Moderate NRL Variations. For nonuniform NRL chromatin combining medium and long DNA linkers, interesting structural variability emerges depending on the size of the NRL variation. Among these, fibers characterized by moderate NRL variations (Table 1; canonical) behave very similar to uniform NRL fibers. We name these "canonical" forms since they adopt irregular 30-nm morphologies with packing ratios only slightly (up to 14%) higher than uniform NRL counterparts (Fig. 4) and interaction patterns that overlap with those of uniform NRL zigzag fibers (Fig. 3A). These include dominant $i \pm 2$ zigzag interactions, and moderate peaks at $i \pm 3$ and $i \pm 5$, due to their five-nucleosome-per-turn zigzag organization. A zigzag architecture is favored because symmetric DNA stems form due to the moderate NRL variation.

As observed in uniform NRL fibers (26), the canonical systems with medium average NRLs, do not engage in far-neighbor interactions (Fig. 3B), and have straight fiber axes (Fig. S2A). Also in agreement with uniform systems, the tendency of the linker DNAs to bend (Fig. S2B) and promote fiber axis bending (Fig. S2A) and far-neighbor contacts (Fig. 3B) grows as the average NRL increases, becoming significant for the 218- to 227-bp fiber. In this fiber, the linker DNAs become much longer than the length of two LHs, which allows them to simultaneously form rigid stems and bend in the middle region that is not screened by LH. This feature stabilizes a heteromorphic architecture that combines both straight (zigzag-like) and bent (solenoid-like) linker DNAs, observed before for chromatin with magnesium ions or with long linker DNAs (12, 26). The similarity with uniform NRL fibers (26) suggests that chromatin fiber structure is robust to moderate NRL variations.

Polymorphism in Fibers with Medium-to-Long Linker DNAs and Large NRL Variations. The behavior of fibers combing medium and long DNA linkers changes significantly when the size of the variation increases (Table 1; polymorphic). Large NRL variations promote polymorphic chromatin fibers that are significantly more compact than the corresponding uniform NRL fibers (Fig. 4 and Fig. S3).

In these polymorphic fibers, all nucleosomes are bound to one long linker DNA (Table 1), which imposes fewer constraints for nucleosome reorganization (26) in favor of higher compaction. Indeed, the packing-ratio peak occurs for the 191- to 227-bp fiber (data point 11), which has the highest NRL variation (17.5 bp) of the set and an average NRL of 209 bp; this packing ratio of ~6.3 nucleosomes per 11 nm is 30% larger than the value for the corresponding uniform NRL fiber and very close to 6.5 nucleosomes per 11 nm determined for chicken erythrocyte chromatin (NRL 206–210 bp) (39, 40).

Within the polymorphic architecture, the range of possible forms includes canonical zigzag fibers and also densely packed conformations with significant fiber axis bending (Fig. S2A), such as sharply bent fibers, hairpin-like structures, and even compact loops (selected representative snapshots are shown in Fig. 2C). These structures are stabilized by multiple types of near-neighbor internucleosome contacts (see Fig. 3A and the diversity of internucleosome interaction patterns for selected snapshots in Fig. S5). Large error bars in the interaction patterns highlight the diversity of near-neighbor contacts. Multiple interactions are likely facilitated by the heterogeneity of the linker DNA lengths and by an increased content of bent DNA linkers (Figs. S2B and S4). DNA bending is produced by the formation of strongly imbalanced DNA stems, in which one of the linker DNAs contains a long flexible region in the middle not reached by the rigidifying LHs.

An additional structural feature of polymorphic fibers is their strikingly high occurrence of interfiber interactions (Fig. 3B); these long-range interactions increase with the NRL variation (from 24% for a variation of 9 bp to 88% for a variation of 17.5 bp). In comparison, uniform NRL fibers with similar average NRLs
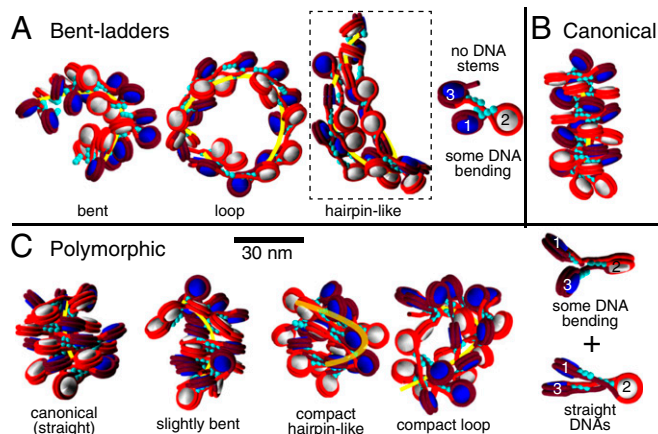
Fig. 2. Representative equilibrium snapshots for nonuniform NRL chromatin fibers. The snapshots are space-filling models (for color code, see Fig. 1). (A) Three bent-ladder conformations observed in the equilibrium ensembles of fibers with one short NRL (fibers 1–3 in Table 1). These snapshots are termed "bent," "loop," and "hairpin-like" simply to aid visualization and exemplify the narrow and highly bent conformations adopted by nonuniform NRL chromatin with one short NRL. The snapshot inside the dashed box is used to illustrate far-neighbor nucleosome contacts in Fig. 3. A nucleosome triplet (consecutive nucleosomes are numbered) is also shown to illustrate the lack of DNA stem formation due to the short linker DNA involved, and the occurrence of some DNA bending. (B) Representative canonical zigzag configuration for fibers combining medium-to-long NRLs with a moderate NRL variation. Only one snapshot is shown because canonical fibers are homogeneous. They exhibit full DNA stems with straight DNA linkers. Fibers combining long NRLs form additional stems with some DNA bending (see triplets in C). (C) Four observed polymorphic equilibrium snapshots for fibers combining two medium-to-long NRLs with a large NRL variation. The structures represent the compact and diverse forms adopted, including a canonical (irregular/heteromorphic zigzag) fiber, a slightly bent form, a sharply bent (hairpin-like) structure, and a compact chromatin loop. Polymorphic fibers form full DNA stems, and exhibit both straight and bent DNA linkers, as illustrated in the two nucleosome triples on the right.
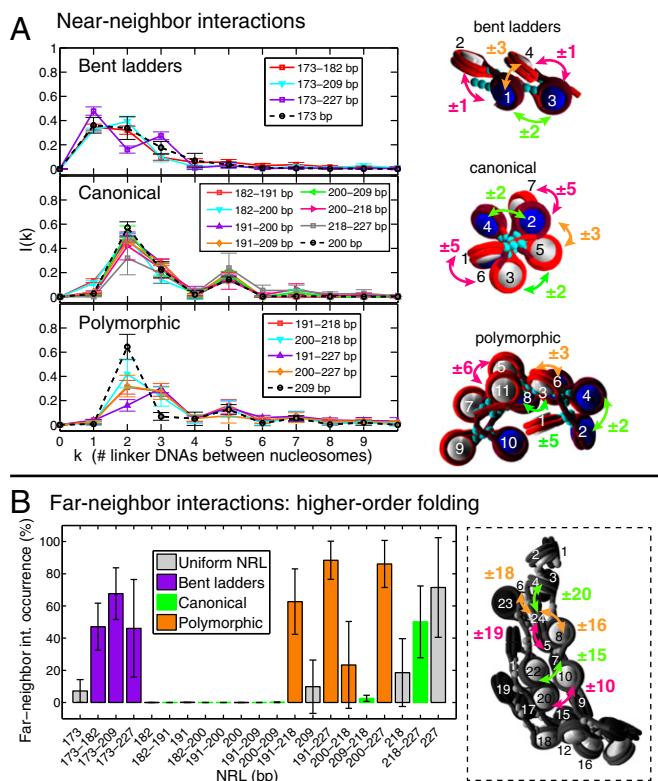
Fig. 3. Chromatin fiber structure as characterized through the frequency of near- and far-neighbor interactions. (*A*) Internucleosome interaction patterns (equilibrium ensemble average and SD) categorized by fiber types (bent ladder, canonical, and polymorphic). For each fiber type, a section of the fiber is shown to exemplify common interactions (selected nucleosomes are numbered) among neighbors separated by k DNA linker segments. Bent ladders have strong $i \pm 1$, $i \pm 2$, and $i \pm 3$ contacts, typical of ladder-like structures. Canonical fibers are characterized by $i \pm 2$, $i \pm 3$, and $i \pm 5$ zigzag contacts. Contrary to the other fiber types, polymorphic structures have internucleosome interaction patterns that do not overlap with that of the uniform NRL fiber; larger error bars and a wide range of strong contacts present highlight the structural diversity. (*B*) Ensemble average and SD of the percentage of conformations with far-neighbor interactions (i.e., conformations with at least one contact between nucleosomes separated by more than nine linker DNA segments) for the three fiber types identified, and the reference uniform fibers (26). Polymorphic fibers, followed by the bent ladders, have the highest occurrence of far-neighbor interactions. One snapshot illustrating the nature of far-neighbor interactions (selected nucleosomes are numbered) is shown (for the color version, see Fig. 2).

and a regular zigzag organization have less than 19% of far-neighbor contacts (Fig. 3*B*, gray bars for NRLs ≤ 218 bp). Long-range interfiber contacts as found in interdigitated fibers are thus favored for large variations of the NRL; such interdigitation corresponds to a higher intensity of histone tail interactions with nonparental DNAs with respect to uniform NRL fibers (see Fig. S6 and *SI Text* for a discussion of the role of histone tails therein). Indeed, long-range internucleosome interactions through histone tails are known to be indispensible to looping and regulatory processes (35, 41). Therefore, by stabilizing interfiber contacts, large NRL variations compromise the formation of regular fibers and instead favor a polymorphic organization.

That both canonical 30-nm structures and a diverse configurational ensemble of interdigitated side-by-side 10-nm aggregates form agrees with propositions based on cryo-EM images of mitotic chromosomes that an irregular organization is the predominant state of compact chromatin, and that 30-nm fibers can form transiently in vivo, especially in crowded environments (18, 42, 43). In such organization, nucleosomes may fold irregularly

and self-associate or interdigitate with one another, leading to lateral interactions between 10-nm fibers and to contacts between distant segments of the same fiber (13, 44). The importance of chromatin loops for metaphase chromatin condensation has emerged recently in 3C experiments combined with polymer simulations (45).

The ability of chromatin to adopt such a high diversity of forms could also be exploited in DNA design applications. Chromatin forms could be controlled by carefully selecting the distribution of nonuniform NRLs across the fiber. For instance, we have produced bent fibers in which nonuniform NRLs, resulting from the removal of selected nucleosomes from uniform NRL fibers, adopt different levels of bending depending on their starting NRL and the number of nucleosomes removed (Fig. S7). Such ideas have potential applications in DNA nanotechnology via introduction of nucleosomes.

Physical Origin of Chromatin Fiber Polymorphism. Through further simulations with altered potential energies (see Fig. S8), we demonstrate that chromatin polymorphism is driven by electrostatics, and is not a result of intrinsic torsion nor topological connectivity. Chromatin polymorphism emerges from a balance between the electrostatic energy and the ability of nucleosomes to reorganize irregularly (and tune this electrostatic internucleosome energy). Such an irregular nucleosome organization, and hence chromatin polymorphism, may occur not only through nonuniform linker DNAs but also through changes in the DNA persistence length, LH removal, or histone epigenetic modifications.

Discussion

Our work highlights a key internal mechanism that controls chromatin fiber structure and compaction: the intrafiber variation of the NRL. Modulation of the intrafiber NRL variation induces a wide range of different fiber forms, including narrow bent ladders, canonical fibers, and polymorphic structures.

When one of the linkers is short, nonuniform-NRL fibers form heterogeneous bent ladders, rather than compact 30-nm regular forms. Such structural fluidity has been proposed for mitotic chromosomes and active interphase nuclei, where nucleosomes
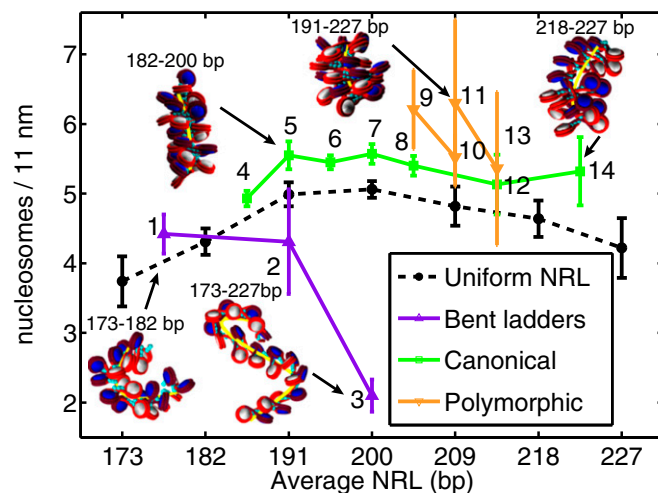
Fig. 4. Fiber compaction as a function of the average NRL per fiber, from equilibrium ensemble averages and SD. Compaction is assessed through the nucleosome linear packing ratio measured as the number of nucleosomes per 11 nm of fiber axis length. Selected simulation snapshots (space-filling models) illustrate differences in compaction (for a description of space-filling model, see the color code in Fig. 1). Data for nonuniform NRL fibers are classified into three types (Table 1): bent ladder (purple), canonical (green), and polymorphic (orange) fibers. The black dashed lines represent the uniform NRL fibers (26). Fiber numbers correspond to those in Table 1. Polymorphic fibers are the most compact.

may be rearranging (46). Chromatin structures with low packing ratios have been observed in vivo not only in transcriptionally active chromatin in yeast (38), but also in transcriptionally inactive regions in the human genome (47). Our results also show that the level of compaction of a short-NRL fiber can be reduced by introducing NRL variations. Thus, simple organisms with higher reproduction rates, such as yeast (average NRL = 168 bp) (48), might exploit NRL variations at selective locations to control fiber opening.

Structures with highly curved fiber axes, such as the bent ladders, have been observed for minichromosomes. Minichromosomes are found in simple organisms and consist of a circular DNA/nucleosome chain. The genome of the simian virus 40, for instance, forms a 5.2-kb minichromosome consisting of looped (regular circles and circles twisted around themselves) chromatin structures (49) with ~20 nucleosomes connected by irregular NRLs (49, 50); these characteristics are close to our 24-unit 173- to 182-, 173- to 209-, and 173 to 227-bp arrays which contain 4.2, 4.5, and 4.8 kb of DNA, respectively. Extrachromosomal yeast chromatin also forms 1.4-kb circles with about nine nucleosomes joined by nonuniform linker DNAs of ~160 and ~180 bp in length (51). In addition, formation of chromatin loops in the context of distant communication has been suggested for a silent yeast locus (23), where a short NRL of 152-bp alternates with ~167- to 177-bp NRL (33).

Chromatin fibers with medium-to-long NRLs are robust enough to accommodate moderate NRL variations, but are highly sensitive to large NRL variations. That is, (i) if the NRL variation is moderate (up to ±9 bp), the fiber retains a canonical zigzag/heteromorphic architecture—analogous to those observed in the well-studied uniform NRL fibers; and (ii) if the NRL variation is larger, a multitude of highly compact forms emerge. This trend is consistent with the speculation made in the seminal tetranucleosome crystal paper (8) that in vivo NRL variations of up to ±5 bp could be absorbed locally, whereas larger ~±10-bp differences would lead to polymorphic fibers.

Because NRL variations of up to ~±4 bp are found extensively in native chromatin (32), the robust stability of chromatin against moderate NRL variations is reasonable. Recently, EM and determination of sedimentation coefficients have also shown that NRL deviations of ±2 and ±4 bp from a medium average NRL do not change chromatin compaction compared with uniform NRL arrays (32). Absorption of moderate NRL variations also implies that chromatin structure is stable against the spontaneous wrapping/unwrapping of a few base pairs from the nucleosome observed through FRET (52, 53).

Large NRL variations might be maintained in vivo by relatively strong positioning sequences. In addition, during the cell cycle, the spacing between nucleosomes can change due to nucleosome sliding and repositioning events, as well as effects of the transcription machinery. High-resolution mapping reveals that most nucleosomes adopt multiple positions (54), and the nucleosome crystal structure suggests that the last 10–20 bp of nucleosomal DNA might not be always wrapped and could have variable conformations in chromatin higher structures (55). It is thus plausible that, at different times in the cell cycle, the combined effects of transcription and remodeling would produce transiently large NRL variations.

The compact polymorphic structures that emerge for large NRL variations include 30-nm irregular fibers but also interdigitated and highly bent conformations that are similar to the irregular and compact aggregations of 10-nm arrays observed in cryo-EM images of mitotic chromosomes (18, 42). Highly bent chromatin is consistent with 3C studies demonstrating that metaphase chromosomes are formed by chromatin loops (45). The highly bent compact conformations we observe are stabilized by both strong internucleosome interactions within fibers (intrafiber) and between fibers (interfiber). However, the intensity of the intrafiber interactions is reduced and the occurrence of the interfiber contacts enhanced, with respect to uniform NRL arrays. Whereas short-range intrafiber contacts leading to regular 30-nm fibers are favored in dilute chromatin conditions, long-range interfiber interactions become increasingly dominant in the crowded environments of mitotic chromosomes (18). Transient NRL variations may thus help initiate folding of higher-level chromatin structures in more crowded environments via long-range internucleosome and tail-mediated communication.

Our results also suggest that electrostatic forces play a key role behind this polymorphism. They drive nonuniform linkers to organize nucleosomes irregularly and promote a variety of long-range nucleosome contacts.

The large intrafiber NRL variations that promote loops and other highly bent structures might be useful for the establishment of tail-mediated interactions between enhancers/silencers and distant promoters. Communication between enhancers and promoters over large genomic distances involves the bending of the chromatin fiber axis (41), and such conformations are thought to form within a compact but dynamic chromatin fiber structure (22). The loops and bent conformations we observe present long-range internucleosome contacts over genomic distances between 1.6 and 5 kb (9–24 nucleosomes); thus intrafiber NRL variations can serve as one factor to enhance communication between elements separated by several kilobases. Consistently, mesoscale modeling combined with experimental rates of communication in chromatin constructs have shown that enhancer–promoter communication is efficient at distances between 0.7 and up to 4.5 kb (41) and involves transient histone tail-mediated internucleosome interactions.

A polymorphic structure has other important biological implications. First, it allows chromatin to pack its gene material more tightly (30% higher packing ratios than regular zigzag arrays). Second, local structural variability simultaneously implies a higher and more homogeneous degree of DNA accessibility: A diverse set of fiber forms exposes selectively different regions of the DNA material. In fact, Nishino et al. (20) suggest that irregular folding has several advantages for template-directed processes, as target sequences are more often exposed than in regular 30-nm fibers. Third, the emergence of chromatin fiber polymorphism from the diversity of DNA-linker lengths supports the idea that the heterogeneous in vivo conditions trigger different topologies, enabling chromatin to achieve various biological roles as required according to the transcriptional state, to the cell cycle stage, or in response to environmental signals and damage (15).

Taken together, our results suggest that NRL variations have a profound impact on the structure of the chromatin fiber. These variations might help regulate both fiber compaction and selective DNA exposure through the stabilization of a wide-range of fiber forms. The polymorphic chromatin structure that emerges here supports the idea that chromatin with heterogeneous components, as found in vivo, can adopt dynamic and interdigitated structures as well as canonical forms. The highly bent and diverse forms that originate from large NRL variations also suggest interesting ways to design and control curved oligonucleosome shapes.

Materials and Methods

Mesoscale Chromatin Model. Our mesoscale oligonucleosome model integrates different coarse-grained descriptions for the nucleosome, histone tails, linker DNA, LHs, and the physiological environment (Fig. S1): The nucleosome, minus histone tails, is represented as an electrostatic object with Debye–Hückel charges (34); DNA linkers as chains of charged beads by a combined worm-like chain model (56); histone tails as flexible chains of charged beads with parameters that mimic their atomistic behavior (36); LHs as rigid electrostatic objects; and solvent implicitly with monovalent ions (Debye–Hückel potential). See *SI Text* and refs. 26, 29, and 36 for further details and a description of the sampling approach, validation of our model, and its limitations. Values of all parameters can be found in refs. 26 and 36.

Simulation Details. MC simulations are performed for 24-core oligonucleosomes with LH at 293 K and 0.15 M NaCl. For each system, we run 12 trajectories of 70 million steps (i.e., four random seeds and three DNA twist deviations around the mean to mimic natural variations) starting from idealized zigzag conformations, as detailed in ref. 26. Convergence is reached well before 60 million steps (Fig. S9). For statistical analyses, we use frames separated by 100,000 steps taken over the last 10 million steps.

1. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* 389(6648):251–260.
2. Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. *J Mol Biol* 319(5):1097–1113.
3. Olins AL, Olins DE (1974) Spheroid chromatin units (v bodies). *Science* 183(4122): 330–332.
4. Robinson PJJ, Rhodes D (2006) Structure of the '30 nm' chromatin fibre: A key role for the linker histone. *Curr Opin Struct Biol* 16(3):336–343.
5. Woodcock CL, Frado LL, Rattner JB (1984) The higher-order structure of chromatin: Evidence for a helical ribbon arrangement. *J Cell Biol* 99(1 Pt 1):42–52.
6. Cui Y, Bustamante C (2000) Pulling a single chromatin fiber reveals the forces that maintain its higher-order structure. *Proc Natl Acad Sci USA* 97(1):127–132.
7. Dorigo B, et al. (2004) Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* 306(5701):1571–1573.
8. Schalch T, Duda S, Sargent DF, Richmond TJ (2005) X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature* 436(7047):138–141.
9. Daban JR, Bermúdez A (1998) Interdigitated solenoid model for compact chromatin fibers. *Biochemistry* 37(13):4299–4304.
10. Robinson PJJ, Fairall L, Huynh VAT, Rhodes D (2006) EM measurements define the dimensions of the "30-nm" chromatin fiber: Evidence for a compact, interdigitated structure. *Proc Natl Acad Sci USA* 103(17):6506–6511.
11. Kruithof M, et al. (2009) Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30-nm chromatin fiber. *Nat Struct Mol Biol* 16(5):534–540.
12. Grigoryev SA, Arya G, Correll S, Woodcock CL, Schlick T (2009) Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. *Proc Natl Acad Sci USA* 106(32):13317–13322.
13. Tremethick DJ (2007) Higher-order structures of chromatin: The elusive 30 nm fiber. *Cell* 128(4):651–654.
14. Fussner E, Ching RW, Bazett-Jones DP (2011) Living without 30nm chromatin fibers. *Trends Biochem Sci* 36(1):1–6.
15. Luger K, Dechassa ML, Tremethick DJ (2012) New insights into nucleosome and chromatin structure: An ordered state or a disordered affair? *Nat Rev Mol Cell Biol* 13(7):436–447.
16. Hübner MR, Eckersley-Maslin MA, Spector DL (2013) Chromatin organization and transcriptional regulation. *Curr Opin Genet Dev* 23(2):89–95.
17. Bouchet-Marquis C, Dubochet J, Fakan S (2006) Cryoelectron microscopy of vitrified sections: A new challenge for the analysis of functional nuclear architecture. *Histochem Cell Biol* 125(1-2):43–51.
18. Eltsov M, Maclellan KM, Maeshima K, Frangakis AS, Dubochet J (2008) Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ. *Proc Natl Acad Sci USA* 105(50): 19732–19737.
19. Maeshima K, Eltsov M (2008) Packaging the genome: The structure of mitotic chromosomes. *J Biochem* 143(2):145–153.
20. Nishino Y, et al. (2012) Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. *EMBO J* 31(7):1644–1653.
21. Schlick T, Hayes JJ, Grigoryev S (2012) Toward convergence of experimental studies and theoretical modeling of the chromatin fiber. *J Biol Chem* 287(8):5183–5191.
22. Kulaeva OI, Nizovtseva EV, Polikanov YS, Ulianov SV, Studitsky VM (2012) Distant activation of transcription: Mechanisms of enhancer action. *Mol Cell Biol* 32(24): 4892–4897.
23. Valenzuela L, Dhillon N, Dubey RN, Gartenberg MR, Kamakaka RT (2008) Long-range communication between the silencers of HMR. *Mol Cell Biol* 28(6):1924–1935.
24. Szerlong HJ, Hansen JC (2011) Nucleosome distribution and linker DNA: Connecting nuclear function to dynamic chromatin structure. *Biochem Cell Biol* 89(1):24–34.
25. Wong H, Victor JM, Mozziconacci J (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(9):e877.
26. Perišić O, Collepardo-Guevara R, Schlick T (2010) Modeling studies of chromatin fiber structure as a function of DNA linker length. *J Mol Biol* 403(5):777–802.
27. Aumann F, Sühnel J, Langowski J, Diekmann S (2010) Rigid assembly and Monte Carlo models of stable and unstable chromatin structures: The effect of nucleosomal spacing. *Theor Chem Acc* 125:217–231.
28. Collepardo-Guevara R, Schlick T (2013) Insights into chromatin fibre structure by in vitro and in silico single-molecule stretching experiments. *Biochem Soc Trans* 41(2):494–500.
29. Collepardo-Guevara R, Schlick T (2011) The effect of linker histone's nucleosome binding affinity on chromatin unfolding mechanisms. *Biophys J* 101(7):1670–1680.
30. Collepardo-Guevara R, Schlick T (2012) Crucial role of dynamic linker histone binding and divalent ions for DNA accessibility and gene regulation revealed by mesoscale modeling of oligonucleosomes. *Nucleic Acids Res* 40(18):8803–8817.
31. Prunell A, Kornberg RD (1982) Variable center to center distance of nucleosomes in chromatin. *J Mol Biol* 154(3):515–523.
32. Correll SJ, Schubert MH, Grigoryev SA (2012) Short nucleosome repeats impose rotational modulations on chromatin fibre folding. *EMBO J* 31(10):2416–2426.
33. Ravindra A, Weiss K, Simpson RT (1999) High-resolution structural analysis of chromatin at specific loci: Saccharomyces cerevisiae silent mating-type locus HMRa. *Mol Cell Biol* 19(12):7944–7950.
34. Beard DA, Schlick T (2001) Modeling salt-mediated electrostatics of macromolecules: The discrete surface charge optimization algorithm and its application to the nucleosome. *Biopolymers* 58(1):106–115.
35. Arya G, Schlick T (2006) Role of histone tails in chromatin folding revealed by a mesoscopic oligonucleosome model. *Proc Natl Acad Sci USA* 103(44):16236–16241.
36. Arya G, Schlick T (2009) A tale of tails: How histone tails mediate chromatin compaction in different salt and linker histone environments. *J Phys Chem A* 113(16): 4045–4059.
37. Bednar J, et al. (1998) Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proc Natl Acad Sci USA* 95(24):14173–14178.
38. Dekker J (2008) Mapping in vivo chromatin interactions in yeast suggests an extended chromatin fiber with regional variation in compaction. *J Biol Chem* 283(50): 34532–34540.
39. Williams SP, et al. (1986) Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length. *Biophys J* 49(1): 233–248.
40. Scheffer MP, Eltsov M, Frangakis AS (2011) Evidence for short-range helical order in the 30-nm chromatin fibers of erythrocyte nuclei. *Proc Natl Acad Sci USA* 108(41): 16992–16997.
41. Kulaeva OI, et al. (2012) Internucleosomal interactions mediated by histone tails allow distant communication in chromatin. *J Biol Chem* 287(24):20248–20257.
42. McDowall AW, Smith JM, Dubochet J (1986) Cryo-electron microscopy of vitrified chromosomes in situ. *EMBO J* 5(6):1395–1402.
43. Maeshima K, Hihara S, Eltsov M (2010) Chromatin structure: Does the 30-nm fibre exist in vivo? *Curr Opin Cell Biol* 22(3):291–297.
44. Grigoryev SA (2004) Keeping fingers crossed: Heterochromatin spreading through interdigitation of nucleosome arrays. *FEBS Lett* 564(1-2):4–8.
45. Naumova N, et al. (2013) Organization of the mitotic chromosome. *Science* 342(6161): 948–953.
46. Takata H, Maeshima K (2011) Irregular folding of nucleosomes in the cell: Comment on "Cracking the chromatin code: Precise rule of nucleosome positioning" by Edward N. Trifonov. *Phys Life Rev* 8(1):51–52, discussion 69–72.
47. Gilbert N, et al. (2004) Chromatin architecture of the human genome: Gene-rich domains are enriched in open chromatin fibers. *Cell* 118(5):555–566.
48. Downs JA, Kosmidou E, Morgan A, Jackson SP (2003) Suppression of homologous recombination by the Saccharomyces cerevisiae linker histone. *Mol Cell* 11(6):1685–1692.
49. Griffith JD (1975) Chromatin structure: Deduced from a minichromosome. *Science* 187(4182):1202–1203.
50. Ambrose C, Lowman H, Rajadhyaksha A, Blasquez V, Bina M (1990) Location of nucleosomes in simian virus 40 chromatin. *J Mol Biol* 214(4):875–884.
51. Thoma F, Bergman LW, Simpson RT (1984) Nuclease digestion of circular TRP1ARS1 chromatin reveals positioned nucleosomes separated by nuclease-sensitive regions. *J Mol Biol* 177(4):715–733.
52. Li G, Widom J (2004) Nucleosomes facilitate their own invasion. *Nat Struct Mol Biol* 11(8):763–769.
53. Poirier MG, Bussiek M, Langowski J, Widom J (2008) Spontaneous access to DNA target sites in folded chromatin fibers. *J Mol Biol* 379(4):772–786.
54. Clark DJ (2010) Nucleosome positioning, nucleosome spacing and the nucleosome code. *J Biomol Struct Dyn* 27(6):781–793.
55. Luger K (2006) Dynamic nucleosomes. *Chromosome Res* 14(1):5–16.
56. Jian H, Schlick T, Vologodskii A (1998) Internal motion of supercoiled DNA: Brownian dynamics simulations of site juxtaposition. *J Mol Biol* 284(2):287–296.