bioRxiv preprint doi: https://doi.org/10.1101/453126; this version posted October 26, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Irregular chromatin: packing density, fiber width and occurrence of heterogeneous clusters

[1]Gaurav Bajpai,[1]Ranjith Padinhateeri*

[1]Department of Biosciences and Bioengineering, IIT Bombay, Powai, Mumbai 400076

Abstract

How chromatin is folded in the lengthscale of a gene is an open question. Recent experiments have suggested that, *in vivo*, chromatin is folded in an irregular manner and not as an ordered fiber with a width of 30nm expected from theories of higher order packaging. Using computational methods, we examine how the interplay between DNA-bending non-histone proteins, histone tails, intra-chromatin electrostatic and other interactions decide the nature of packaging of chromatin. We show that while the DNA-bending non-histone proteins make the chromatin irregular, they may not alter the packing density and size of the fiber. We find that the length of the interacting region and intra-chromatin electrostatic interactions influence the packing density, clustering of nucleosomes, and the width of the chromatin fiber. Our results suggest that the actively maintained heterogeneity in the interaction pattern will play an an important role in deciding the nature of packaging of chromatin.

October 2018*ranjithp@iitb.ac.in

1 Introduction

DNA is a very long polymer that contains the genetic code. Inside biological cells, DNA is not in its bare form; rather it is folded, decorated and packaged with the help of myriads of proteins and chemical groups into a functional structure known as chromatin (1). Purpose of folding DNA into chromatin is not just for the packaging and storage of the polymer but also for regulating the accessibility and reading of the genetic code. How is this one-dimensional sequence information is folded and packaged into a 3-dimensional chromatin organization is poorly understood. Research in the last few decades have shown that stretches of DNA are wrapped around a multimer complex of histone proteins leading to formation of an array of nucleosomes (2, 3). However, exactly how this nucleosomal array is further folded and packaged is an open question (4).

In the packaging of chromatin, the nature of interaction and structural details of nucleosomes are thought to play an important role (5, 6). DNA is a negatively charged polymer while histone proteins are predominantly positively charged. In a nucleosome, histone protein subunits (H2A-H2B, H3-H4) have tail-like regions giving raise to a nucleosome complex having multiple tails protruding out of the core region (7). Given that these tails are electrostatically charged and various chemical modifications (e.g. acetylation) can alter the amount of charge, the histone tails and electrostatic interactions are thought to play crucial roles in higher order packaging of chromatin (8–10). Apart from the core histones, the linker histone (H1) is also thought to play an important role in chromatin organization. The H1 protein is known to bind between DNA segments that are entering and exiting a nucleosome, and thought to stabilize the nucleosome structure (11–13).

How a string of nucleosomes, known as 10-nm wide chromatin, gets folded further to form higher order structure is a question that is not settled even after nearly four decades of research (4, 14). In vitro reconstitution of chromatin, based on array of nucleosomes, suggested that the nucleosomes will form a regular zig zag-like structure having a width of 30nm (7, 15, 16). Based on other *in vitro* experiments, there is also an alternate hypothesis that the chromatin should be solenoid with 30nm width (17, 18). Simple polymer physics theories, accounting for certain physical aspects of DNA stiffness and structure of nucleosomes predicted regular 30nm chromatin organization (19–22). The extend of folding of this string of nucleosomes into chromatin is measured by computing the packing density, which is roughly defined as the number of nucleosomes packed in every 11nm effective length of the chromatin fiber. Analysis of *in vitro* experimental data and simulation results showed that the packing density of 30 nm chromatin fiber can vary from 6 nucleosomes/11nm to 12 nucleosomes/11nm depending on different conditions (11, 12, 16, 17, 23–32) .

© 2018 The Authors

bioRxiv preprint doi: https://doi.org/10.1101/453126; this version posted October 26, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

However, most of the recent experiments suggest that, *in vivo*, chromatin does not have any regular structure (33–36). Cryo-electron microscopy (cryo-EM) study on mitotic and interphase cells suggest that there is no 30nm-wide structure *in vivo* (33–35). Recent studies with advanced EM techniques to visualize chromosome in interphase and mitotic cells, suggest that chromatin is folded into an irregular chain having $\approx$ 5-to-24nm wide structures (37). Chromatin conformation capture (Hi-C) experiments suggest that chromatin in interphase is not a homogeneous or a regular structure (38). It is organized into topologically associated domains (TADs) and lamina associated domains (LADs) having open (euchromatin) and compact (heterochromatin) domains distributed across the nucleus and have a fractal nature (38, 39).

There have been many recent theoretical/computational studies trying to understand the irregular nature of chromatin (10, 40–42). It has been hypothesized that molecular crowding in the cell may lead to the formation of irregular chromatin (34). It has been shown that the variability in linker length and other factors give raise a polymorphic structure of chromatin (42, 43). Our own earlier work suggested that DNA-bending non-histone protein can make chromatin irregular (41).

Even though there is a vast literature on chromatin organization, there remains many important questions unanswered. Given that chromatin is irregular in the length-scale of genes, the first question is about its compactness. How compact is the irregular chromatin when compared to the regular structures? How does the irregular nature influence the packing density? Given that recent experiments observed chromatin of width 5nm to 24nm, can we have a theoretical explanation for having this wide variety of widths? Given that chromatin is heterogeneous in terms of interaction potentials (e.g., spatial variation of histone modification), how will this affect chromatin configurations? Do the spatial extend/variation in modification pattern affect the packing density and width?

In this paper, we address these questions by performing coarse-grained molecular simulations and studying chromatin organization in 3D. We simulate a polymer model of chromatin having nucleosomes with explicit histone tails, electrostatic interactions and other intra-chromatin interactions. We also account for the spatial variation of the interaction potentials to mimic interaction heterogeneity due to histone modification patterns. We show how DNA-bending non-histone proteins and heterogeneous interactions among nucleosome particles affect packaging of irregular chromatin. Our work explains why chromatin is less than 30nm wide even when it is in a highly packaged heterochromatin state, as observed in recent experiments.

2 Model and Method

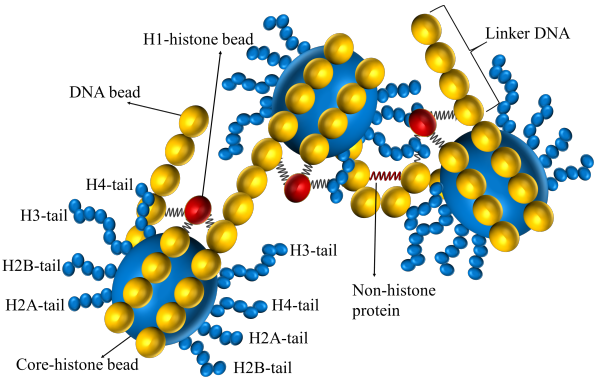

Figure 1. Schematic diagram describing the model: DNA is modeled as a polymer made of type-1 beads(yellow) having diameter of 3.4 nm. 14 DNA-beads wrap around core histone bead (type-2, big blue bead, 5.25 nm in diameter). Histone tails are modeled as flexible bead-spring polymer chains where each histone-tail bead (type-3, small blue bead) is 1.56 nm in diameter and have the following lengths: tail-1 (H2A)= 4 beads, tail-2 (H2B)= 5 beads, tail-3 (H3)= 8 beads, and tail-4 (H4)=5 beads. H1 histone (red bead, type-4) is connected to three DNA beads— the entry bead, exit bead and the central bead of DNA wrapped around the core histone bead.

In this section we describe a coarse-grained model for chromatin, the method of simulating the model, and the parameter values we use. In the coarse-grained model, chromatin is modelled as a set of bead-spring chains having five major parts namely (i) the DNA polymer chain, (ii) the core histone,(iii) explicit histone tails, (iv) the linker histone and (v) the non-histone protein(see Fig. 1).

The DNA polymer is modeled as a bead-spring chain having N_d beads connected by $N_d - 1$ springs (Fig. 1). A nucleosome is modelled as one big bead representing the core histone octamer on which 14 DNA beads are wrapped around in 1.75 turns.

bioRxiv preprint doi: https://doi.org/10.1101/453126; this version posted October 26, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

To explicitly model histone tails, we introduce eight small flexible polymers emanating from the core histone as shown in Fig. 1. The DNA polymer bead has a diameter of 3.4 nm while the octamer core histone bead has a diameter of 5.25 nm. The details of each histone-tail polymer length is as follows: H2A tail= 6.2 nm(4 beads), H2B tail = 7.8 nm(5 beads), H3= 12.6 nm(8 beads) and H4= 7.8 nm(5 beads) (21). We fix the first bead of each histone tail, on the core histone bead such that the first bead and the core histone bead behave like a rigid body. We introduce the linker histone H1 as a separate bead (diameter=2.9nm (44)) which interacts via a harmonic spring with the two DNA beads entering and exiting the nucleosome as well as with the nucleosome. The H1 bead constrains the DNA tangent vectors entering and exiting the nucleosomes ensuring known structural features of nucleosomes.

Now we describe the energetics of the chromatin system that we consider. Throughout this paper, we will use notation with subscript $(\alpha, \beta) \in \{d, h, t, l\}$ where d, h, t, l stand for DNA bead, histone bead, tail bead and linker bead, respectively. For example, diameter of a DNA bead will be represented by σ_d while that of the tail bead will be σ_t, and so on. All beads in the system are connected with their respective neighbours using a harmonic potential given by the general formula as:

$$U^{spring} = \sum_{\alpha,\beta} \sum_{i,j} \frac{k^{spring}_{(\alpha,\beta)}}{2} \left[|\mathbf{r}^{(i)}_{\alpha} - \mathbf{r}^{(j)}_{\beta}| - \left(\frac{\sigma_\alpha + \sigma_\beta}{2} \right) \right]^2, \tag{1}$$

where $\mathbf{r}^{(i)}_{\alpha}$ and $\mathbf{r}^{(j)}_{\beta}$ are position vectors of ith or j^{th} beads of types α and β; $k^{spring}_{(\alpha,\beta)}$ is the corresponding spring constant. (i) For DNA-DNA bead interaction $\alpha = d$, $\beta = d$, $j = i + 1$ and i ranges from 1 to $N_d - 1$. (ii) For histone core-DNA interaction $\alpha = d$, $\beta = h$, i ranges from 1 to N_h (where N_h is total number of nucleosomes) and j takes 14 values, for each i, representing the identity of the corresponding histone-bound DNA beads. (iii) For tail-tail bead interaction $\alpha = t$, $\beta = t$, $j = i + 1$ and $i = N_t - 1$ where N_t is number of tail bead. (iv) For DNA and linker histone interaction $\alpha = l$, $\beta = d$, i ranges from 1 to N_l (where N_l is total number of linker histones), j stands for 3 DNA bead positions at the entry site, exit site and the central DNA bead (wrapped around histone) along the dyad axis. The interaction energy of the non-histone proteins(NHPs) is also calculated using equation 1. An NHP binds at any linker region between two linker DNA beads α and β. If there are ν NHPs, each having size spanning 3 beads, i is the count of the NHPs varying from 1 to ν and $j = i + 3$, and ν/N_h is called protein density. The total spring energy (U^{spring}) is obtained by summing over with appropriate nearest neighbour bead pairs α, β. Apart from the nearest neighbour spring interaction described above, any two beads interact via the standard Lennard-Jones (LJ) potential (U^{LJ}) such that

$$U^{LJ} = \sum_{\alpha,\beta} \sum_{i,j} 4\epsilon_{(\alpha,\beta)} \left[\frac{\left(\frac{\sigma_\alpha+\sigma_\beta}{2}\right)^{12}}{(|\mathbf{r}^{(i)}_{\alpha} - \mathbf{r}^{(j)}_{\beta}|)^{12}} - \frac{\left(\frac{\sigma_\alpha+\sigma_\beta}{2}\right)^{6}}{(|\mathbf{r}^{(i)}_{\alpha} - \mathbf{r}^{(j)}_{\beta}|)^{6}} \right], \tag{2}$$

when $|\mathbf{r}^{(i)}_{\alpha} - \mathbf{r}^{(j)}_{\beta}| < 2.5 \left(\frac{\sigma_\alpha+\sigma_\beta}{2}\right)$, and $U^{LJ} = 0$ otherwise. That is, the energy is zero when $|\mathbf{r}^{(i)}_{\alpha} - \mathbf{r}^{(j)}_{\beta}| \geq 2.5 \left(\frac{\sigma_\alpha+\sigma_\beta}{2}\right)$. Here $\epsilon_{(\alpha,\beta)}$ is the corresponding potential well depth. All beads interact with each other using screened electrostatic potential. We use the standard Debye-Huckel potential and compute this energy of chromatin (21) as follows:

$$U^{electro} = \frac{C}{\epsilon^w} \sum_{\alpha,\beta} \sum_{i,j} \left[\frac{q_\alpha q_\beta}{(|\mathbf{r}^{(i)}_{\alpha} - \mathbf{r}^{(j)}_{\beta}|)} \mathrm{e}^{-\kappa(|\mathbf{r}^{(i)}_{\alpha} - \mathbf{r}^{(j)}_{\beta}|)} \right], \tag{3}$$

where q_α and q_β are effective charges of beads, κ is the inverse of debye length (1nm^{-1}), ϵ^w is the dielectric constant (set to 80 assuming water-like medium) and C is a constant as per the screened Coulomb electrostatic potential energy formula (45). Total charges on wrapped DNA are estimated as $-296e$ (46). 14 DNA beads wrap around core histone, so charge value for one DNA bead $q_d = -296/14 = -21.14e$. The core histone bead (without tails) has a charge of $q_h = 52e$ (46). We take charge on each histone-tail bead $q_t = 2e$ such that the total charge on one nucleosome (wrap DNA+core histone+histone tails) is maintained as $-156e$ (46). Charges for histone H1 are taken $q_l = 13.88$ (21).

The bending energy (U^{bend}) of the DNA chain is given by

$$U^{bend} = \frac{k^{bend}}{\sigma_d} \sum_{i=1}^{N_d-1} (1 - \cos\theta^{(i)}) \tag{4}$$

where k^{bend} is the bending stiffness of DNA and $\theta^{(i)}$ is the angle between two nearby bonds in the bead-spring model.

The total energy of the chromatin in this model is given by $U^{tot} = U^{spring} + U^{electro} + U^{LJ} + U^{bend}$. This system was simulated using Molecular Dynamics simulation package LAMMPS (47) and we obtained chromatin configurations (3D positions of all beads) as a function of time.

bioRxiv preprint doi: https://doi.org/10.1101/453126; this version posted October 26, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Parameters used in simulations

In our model, 4 types of beads sizes(diameters) and masses are denoted by m_α and σ_α respectively, where $\alpha \in \{d, h, t, l\}$; here d represents DNA bead, h represents core-histone bead, t represents histone tail bead and l represents the linker histone(H1) bead. To perform Molecular Dynamics using LAMMPS (47), mass of DNA bead is taken $m_d = 6000$gm/mole (1). Other masses are calculated by relation $m_\alpha = \frac{m_d\,\sigma_\alpha^3}{\sigma_d^3}$. For stability of all types beads, we took same and high spring constant ($k^{spring}_{(\alpha,\beta)} = k^{spring}$) values. For this, we took minimum from σ_α which is $\sigma_t = 15.6A^o$. Then $k^{spring} = \frac{100k_BT}{\sigma_t^2}$, where $1k_BT = 0.6$ kcal/mol (21). Calculated spring constant is $k^{spring} = 0.17$ kcal/mol/A^2. We calculated time-step(Δt) for BD simulations using formula $\Delta t = \frac{1}{2\pi}\sqrt{\frac{m_{(\alpha,\beta)}}{k^{spring}}}$; where $m_{(\alpha,\beta)}$ is the effective mass, given by $m_{(\alpha,\beta)} = \frac{m_\alpha m_\beta}{m_\alpha + m_\beta}$ and (m_α, m_β) is appropriate combinations of (m_d, m_h, m_t, m_l). For small Δt, we took minimum of $m_{(\alpha,\beta)}$ from the combinations. Calculated time-step is $\Delta t = 359.5fs$.

3 Results

Below we present the results from our study, using computer simulations, where we examined how the chromatin organization alters as we vary different interaction potential energies (electrostatic, LJ), DNA-bending due to non-histone proteins, and the length of the interacting domain. We will discuss how heterogeneity in interaction potential energies will lead to heterogeneous chromatin configurations.

Role of different interactions in chromatin packaging

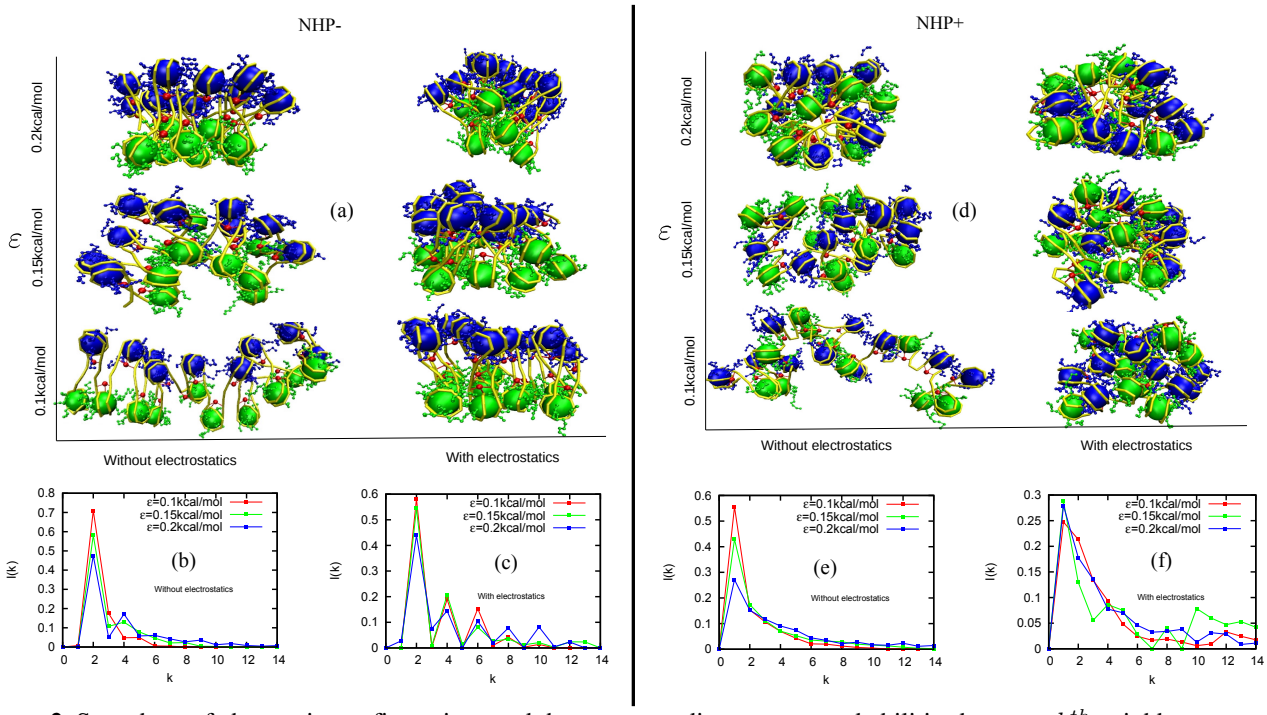

Figure 2. Snapshots of chromatin configurations and the corresponding contact probabilities between k^{th} neighbour nucleosomes (I(k)) in the absence (NHP-), and presence (NHP+) of DNA-bending non-histone proteins (NHPs) that bind along linker DNA regions. The results are presented for four cases: case 1: Simulations with no non-histone proteins and no electrostatic interactions (left column of (a), and (b)); case 2: No non-histone proteins but with electrostatic interactions (right column of (a), and (c)); case 3: with non-histone proteins and no electrostatic interactions(left column of (d), and (e)).case 4: with non-histone proteins and with electrostatic interactions (right column of (d), and (f)). All the results are presented for 3 different strengths of Lennard-Jones (ϵ) interaction potentials (see text for details).

bioRxiv preprint doi: https://doi.org/10.1101/453126; this version posted October 26, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Nucleosomes in a chromatin interact with its neighbours via different interaction potential energies. However the interaction energies and their strengths vary depending on the precise state of the chromatin. For example, acetylation of histones remove positive charges from the proteins leading to an altered electrostatic potential energy landscape (48). Similarly, interactions of many proteins (for example, HP1) can bring about inter-nucleosomal interactions (49). To mimic these interactions, we introduce electrostatic interaction energy and Lennard-Jones interaction energy into our model.

First we present results from a simulation of a chromatin made of 20 nucleosomes, accounting for histone tails, and H1 linker histones, but not considering the DNA-bending non-histone proteins (NHP-). 3D configurations of such a chromatin for different values of interaction potential strengths are shown in Fig. 2(a). In the left panel within Fig. 2(a), we have chromatin for different LJ interactions strengths but with no electrostatic interactions. In the right panel, chromatin configurations for exactly the same condition but with electrostatic interactions are shown. When electrostatic interaction is present, the chromatin is more compact. The positive charges of the histone tails attract with negatively charged parts of the chromatin. This leads to a compact structure.

Based on nucleosome structure data, we know that each tail-bead, in our coarse-grained model, has one or more sites that can be acetylated. Acetylation will reduce electrostatic interaction in the tails. Our "without electrostatics" results are comparable to a chromatin with highly acetylated nucleosomes; we get open structures when there is no electrostatic interactions and relatively more compact structures with electrostatic interactions analogous to chromatin with acetylated and de-acetylated nucleosomes. This is consistent with earlier findings (5, 42).

If we go along the the column from bottom to top (Fig. 2(a)), we find chromatin organization for different values of ϵ (LJ interaction strength). For small values of ϵ, chromatin is more open while for larger values of LJ interaction strengths, the chromatin is compact. Here we introduced LJ interaction to mimic different types of interactions mediated by various proteins (like HP1) or other interactions that may bring together nucleosomes (49).

We then computed $I(k)$—the probability that a nucleosome is contact with its kth neighbour— from the above-discussed 3D chromatin configurations (see Figs. 2(b) & (c)). We find that, irrespective of electrostatic interactions and strength of the LJ interactions the most prominent peak is at $k = 2$. Note that when we change the strength of the LJ, in the absence of electrostatics, the height of the prominent peak varies. This is indeed related to the compact packaging of zig-zag structure. In the open form (low LJ), the zig-zag is more loose, hence only next-neighbor nucleosomes (k=2) mainly interact; far away nucleosomes ($k > 2$) are not contributing to $I(k)$. However, for large LJ, the chromatin is compact and far away nucleosomes have higher chance of being close and hence I(k) values for $k > 2$ are larger. Since total probability is 1($\sum_k I(k) = 1$), this reduces the peak value of I(2) for large LJ values. Similarly, when electrostatic interactions are present, the compactness of the chromatin is reflected in higher $k > 2$ peaks.

Even compact chromatin structures are irregular in presence of DNA-bending non-histone proteins

In this subsection, we want to understand the interplay between DNA-bending due to non-histone proteins and electrostatic/LJ interactions in deciding the 3D folding of chromatin. We did similar simulations as above, but introducing NHPs that bind in the linker region and bend DNA, assuming NHP density of 0.5 per linker region. Typical snapshots of chromatin 3D configurations are shown in Fig. 2(d) with (right panel) and without (left panel) electrostatic interactions; also see the LJ variation along the vertical axis. Without electrostatics, for small LJ, the chromatin is in open configuration as this resembles acetylated (charge neutralized) chromatin with no major inter-nucleosomal interactions. As we increase the LJ interaction (high ϵ), the chromatin becomes more compact. When electrostatic interactions are switched on (right panel), we got relatively compact structures for all values of ϵ.

The important point to note is that in this figure (Fig. 2(d)), as opposed to the earlier case in Fig. 2(a), different color beads (odd and even numbered beads with blue and green colors) are well mixed, suggesting a irregular organization of chromatin. The presence of NHPs have created an irregular chromatin even with electrostatic interactions and high LJ interactions.

To quantify this, we computed the contact probability $I(k)$ corresponding to the configurations given above. We find that, in the presence of NHPs, irrespective of electrostatic interactions and strength of the LJ interactions the most prominent peak is at $k = 1$ (Fig. 2(e)). Here, it is interesting to see that even for high compact form of chromatin, structures are irregular in presence of NHPs. Note that when we change the strength of the LJ, with no electrostatic interaction, the height of the prominent peak (peak at $k = 1$) varies. This is indeed related to the compact packaging of irregular structure. In the open form (low LJ), the irregular structure is more loose, hence only neighbor nucleosomes (k=1) mainly interact; far away nucleosomes ($k > 2$) are contributing relatively less to $I(k)$. However, for large LJ, the chromatin is compact and far away nucleosomes have higher chance of being close and hence I(k) values for $k > 1$ are higher. Since total probability is 1($\sum_k I(k) = 1$), this reduces the peak value of $I(1)$ for large LJ values. In Fig. 2(f), with electrostatic interactions, the structure is highly compact and irregular; hence the $I(k)$ peaks are roughly the same for all values of LJ interaction strengths.

bioRxiv preprint doi: https://doi.org/10.1101/453126; this version posted October 26, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Regular vs. irregular chromatin: packing density, fiber width and mean cluster size

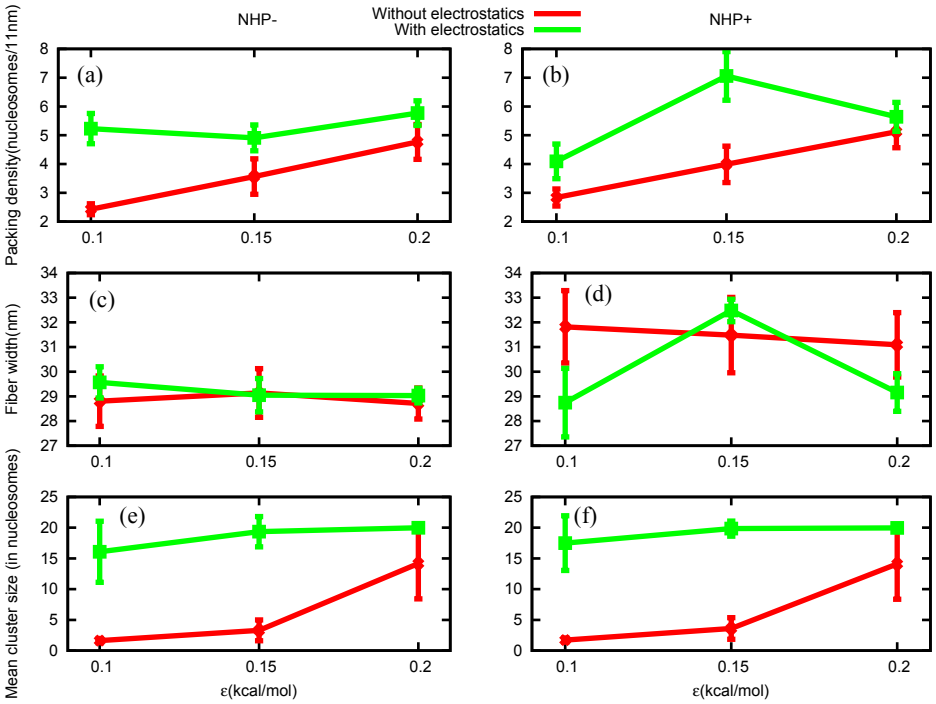

Figure 3. Packing density, fiber width and mean cluster size are plotted for different LJ strengths (ϵ), in the absence of NHPs (NHP-, left side) and presence of NHPs (NHP+, right side). (a),(b) Without electrostatic interaction, the packing density increases as we vary the strength of the LJ (ϵ) potential (red curves). However, with electrostatic interaction, packing density is ≈ 6 nucleosomes/11nm (green curve) having less variation as a function of the LJ parameter. (c),(d) Chromatin fiber diameter(width) with and without electrostatic for different LJ strengths (ϵ). With (green curve) and without (red curve) electrostatic interaction, the fiber diameter is ≈ 30 nm (constant) as we vary the strength of the LJ (ϵ) potential. (e),(f) Without electrostatic interaction, mean cluster size increases on increasing of the LJ (ϵ) parameter (red curve), but with electrostatic interaction it remains constant(green curve). In all the subfigures, the vertical bars represent standard deviation.

We computed packing density to know compaction of different chromatin structures for the above cases. Packing density measures roughly the number of nucleosomes packed in every 11nm of effective chromatin length. First we consider chromatin organization with no NHPs; these are regular ordered chromatin structures as seen in the above results (Fig. 2(a)). In the absence of NHPs and in the presence of electrostatic interactions, the packing density is ≈ 6 nucleosomes/11nm (see Fig. 3(a), green curve) as expected for zig-zag or regular chromatin (11, 16, 17, 23–27). The results show that the chromatin is highly packed even for a very small LJ interaction strength($\epsilon = 0.1$kcal/mol) and remains nearly the same for larger LJ interaction strengths ($\epsilon = 0.2$kcal/mol). In the absence of electrostatic interactions (Fig. 3(a), red curve), packing density is much smaller ($\approx 2-3$ nucleosomes/11nm) than the case with electrostatic interactions suggesting that the electrostatics play an important role in packaging. On increasing the strength of the LJ interactions from $\epsilon = 0.1$kcal/mol to $\epsilon = 0.2$kcal/mol, we find the the packing density slightly increases, but remains smaller than that of the chromatin with electrostatic interactions.

With DNA-bending NHPs, we got a packing density of 4 to 7 nucleosomes/11nm, in the presence of electrostatic interactions (Fig. 3(b)). This implies that, even though our model with NHPs do not give raise to regular/ordered structure, the packing/compaction remains similar to that of what is observed in *in vitro* experiments. Our theory predicts that DNA-bending non-histone proteins will make chromatin irregular but will have similar compaction as observed in the case without NHPs (11, 16, 17, 23–27). The findings for the case with no electrostatic interactions are also similar to the case without NHPs —a lower packing density that increases as a function of LJ interaction strengths.

In Fig. 3(c) and (d), we present our results for fiber width in the absence and presence of NHPs, with (green curve) and without (red curve) electrostatics. Interestingly, for all the cases, the width is ≈ 30nm. This width is computed according to a simple definition based on polymer physics ideas. We know (from Fig. 2(b)) that in the presence of NHPs, chromatin is irregular; we also know that in the absence of electrostatic interactions, the chromatin is relatively more open. However, in all

bioRxiv preprint doi: https://doi.org/10.1101/453126; this version posted October 26, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

these cases the width is approximately 30nm. This shows that the width as a quantity, as defined here, cannot easily distinguish between regular, irregular, open and compact chromatin. However, note that the packing density in the above figure could distinguish between open and compact chromatins (red and green curves).

Since both compact and open structures might have similar width, to improve the quantification, we computed how nucleosomes are clustered near one another. Any two nucleosomes that are closer than 2.5 times its diameter ($2.5\sigma_h$) are considered to be in the same cluster. We find that the open chromatin is not just one cluster but many small clusters having a few nucleosomes in each—see Figs. 3(e) & (f) (red curve) where the mean cluster size is small ($\approx 2-5$) and hence the mean cluster number will be large $(4-10)$. On the other hand, the compact chromatin is nearly one cluster — ≈ 20 nucleosomes are part of the same cluster (see Fig. 3(e),(f), green curve). This suggests that open chromatin, even through it appears like a loose zig-zag visually, may not appear as a single entity (punctate) in experiments (like cryo-EM) where the density and cluster size can affect the measurement.

Role of chromatin length in packing density and fiber width

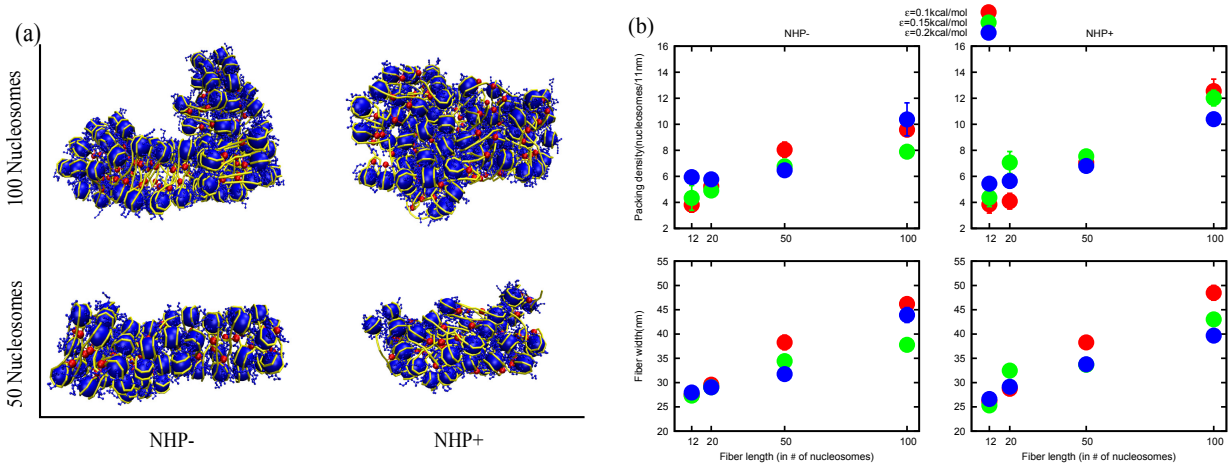

Figure 4. (a) Snapshots of simulation results of 50 nucleosomes and 100 nucleosomes in the absence and presence of NHPs (left to right). (b) Packing density and fiber width on varying chromatin fiber length in the absence and presence of NHPs for different LJ interaction strengths (ϵ). Packing densities and fiber widths increase as we increase the chromatin length.

So far, we studied packing density, fiber width and cluster size for a chromatin having 20 nucleosomes. We found that, even though the NHPs make the chromatin irregular, a compact chromatin fiber has a packing density$\approx$ 6 nucleosomes/11nm, and a width $\approx$ 30nm. How do packing density, width of the fiber and other parameters vary as we change the length of the polymer? In this section we simulated chromatin fibers of many different lengths starting with chromatin having 12 nucleosomes going upto 100 nucleosomes.

In Fig. 4(a) snapshots of chromatin structure for polymers of having 50 nucleosomes and 100 nucleosomes are shown in the absence and presence of NHPs. Packing densities and fiber widths for different lengths (12, 20, 50 and 100 nucleosomes) are shown in Fig. 4(b). In the absence (left panel) and presence (right panel) of NHPs packing densities and fiber widths increase on increasing the polymer length from 12 nucleosomes to 100 nucleosomes. The longer the polymer, the more the width and higher the packing density. For polymers having 20 nucleosomes or higher, the width is 30nm or above; the packing density is above 6 nucleosomes per 11nm. This result suggests that, if we pack chromatin even in an irregular manner, with high compaction, one will get structures having width 30nm or higher. However, in the context of *in vivo* chromatin, where recent experiments do not find chromatin of width 30nm or higher, this presents a puzzle. Why do *in vivo* experiments do not find chromatin having width 30nm? The recent *in vivo* experiment finds chromatin of widths only in the range 5nm–24nm (37).

Spatial variation in histone modifications and interaction potentials

To address the above-mentioned puzzle, where one gets 30nm wide chromatin if we have highly interacting 20 nucleosomes or higher, we turn our attention to the histone modifications. We know that histone modifications vary spatially (along the contour

bioRxiv preprint doi: https://doi.org/10.1101/453126; this version posted October 26, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of a chromatin) and this can influence the nature of effective interactions between different parts of chromatin (48). Presence of certain methylations (like H3K9me3) could lead to recruitment of certain proteins (like HP1) and induce high attractive interactions between chromatin segments (49). The absence/presence of acetylations could also affect the local electrostatic potentials leading to heterogeneous interactions. All these will affect the packaging and clustering of nucleosomes.

To understand this heterogeneity in interactions, as an example, we examined the extend of H3K9me3 modifications along chromatin fiber length. Using ChIP-seq data available in public databases, we asked the following question. What is the length of the typical contiguous patch of chromatin having H3K9me3 modification?(human T-cell in ref. (50)). We defined any two modification peaks as a part of the same contiguous patch if the separation between the peaks is less than 1000bp. Using this definition, we computed mean lengths and standard deviations in lengths of contiguous patches of chromatin having H3K9me3 modifications for 22 human chromosomes (see Fig. 5(a)). The mean length of H3K9me3 modification patch across all chromosomes is ≈ 1600 bp and the gap length between two contiguous patches is ≈ 2200 bp. If you convert in terms of nucleosomes, this is equivalent to chromatin having ≈ 8 nucleosomes with modifications and ≈ 12 nucleosomes without.

To mimic this spatial variation in modifications, we simulated 20 nucleosomes where we introduced electrostatic interactions into one fraction (8 nucleosomes) and the other fraction (12 nucleosomes) was left with no electrostatic interactions. The LJ potential is kept uniform ($\epsilon = 0.1$kcal/mol) and the linker regions are bent with NHPs having a density of 0.5. The resulting chromatin structure is shown in Fig. 5(b). The 8 nucleosomes having electrostatic interactions form one cluster, while the other 12 nucleosomes divide into many small clusters. Then we calculated the width of the single cluster of 8 nucleosomes and got 22.6 ± 1.5nm which is comparable recent experiments results where they found fiber width 24nm (37). The other small clusters will have widths varying from a single-nucleosome width ($\approx 5 - 10$nm) to two or three nucleosome widths (10-20nm). Note that here the chromatin is irregular due to the presence of NHPs.

This suggests that heterogeneity in interactions resulting from spatial variation of histone modifications may determine the width of chromatin fibers *in vivo*.

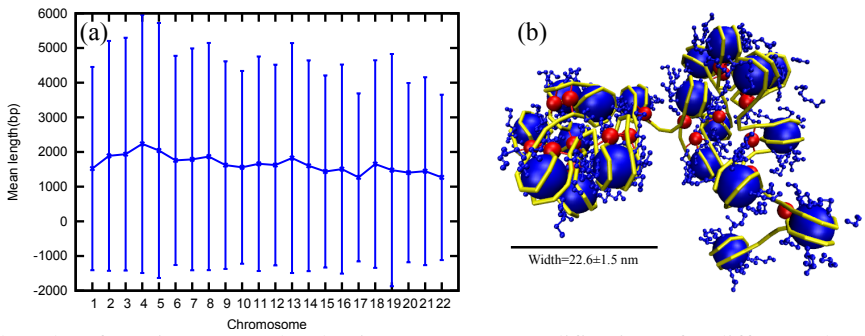

Figure 5. (a) Mean lengths of contiguous patches having H3K9me3 modifications, for different chromosomes (Human cell). The corresponding standard deviations are shown as vertical bars. (b) A snapshot of the chromatin structure simulated with electrostatic interactions in the first 8 nucleosomes, and with no electrostatic interaction in the remaining 12 nucleosomes.

Conclusion

In this work we examined the width, packing density and clustering properties of chromatin that is irregular due to the binding of DNA-bending non-histone proteins. We find that the DNA-bending non-histone proteins will make chromatin irregular; however, the resulting structure can have similar packing density and width as that of the chromatin without NHP (regular). We examined how different factors such as the length of the chromatin and the nature/strength of interactions determine the width, packing density and clustering properties of chromatin. In our simulations, we explicitly accounted for histone tails, and electrostatic interaction and varied the length of chromatin across several nucleosomes (length of a typical gene to many genes). We showed that electrostatic interactions make the chromatin more compact and a chromatin simulated without electrostatic interactions are more open. This is consistent with the fact that reduction in net positive charges (hence reduction in electrostatic interactions) due to acetylation of histone tails lead to more open chromatin. We examined how the nucleosomes are clustered in a highly packed chromatin and in an open chromatin. We calculated mean cluster size of simulated structure and showed that open chromatin is essentially many small clusters of nucleosomes; they may not appear as a single thick fiber in many experiments. We varied fiber length and calculated packing densities and fiber widths and showed that both quantities increase with fiber length. Then we addressed the resulting puzzle that how one does not observe highly packed chromatin fibers of

bioRxiv preprint doi: https://doi.org/10.1101/453126; this version posted October 26, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

width 30nm or above *in vivo*? We argued that one of the missing components here could be the heterogeneity in interactions resulting from histone modifications. We simulated chromatin configurations considering this heterogeneity in interactions and showed that heterochromatin structure could have typical width less than 30 nm.

The other factor that might affect the local structure and the width is the ATP-dependent activity and dynamics of nucleosomes and proteins. In this work, we have assumed that nucleosomes and NHPs are static, and do not change their organization with time. However, remodeling machines and cellular processes like gene-expression can disassemble/ re-organize nucleosomes which may affect the properties of chromatin in the lengthscale of a gene. However, our model will be a reasonable description for inactive regions—regions where there are least remodeling activity and gene expression. It is believed that such heterochromatic regions will have 30nm wide chromatin structures. However, our work suggest that even there, depending on the nature of histone modifications, one may get chromatin widths smaller than 30nm. We have assumed that electrostatic interactions obey a Debye-Huckel potential; however, fluctuations of counter-ions and other charged constituents are not accounted in this description of the interactions. These are some of the limitations of our model.

Suggestion for new experiments to test our prediction: Experimentally, our findings can be tested in a few different ways. One can possibly perform *in vitro* chromatin reconstitution experiments with NHPs and measure the width, packing density and cluster sizes . Our prediction is that the chromatin will be irregular, but having a width of $\approx$ 30nm, if the length is appropriately chosen as we show in our results; however depending on the nature of modifications, one may not find chromatin as one cluster but many small clusters of nucleosomes. This may be repeated for many different lengths and measure quantities as a function of length. Our work also suggests that one should experiment with heterogeneity in interactions; this may be introduced by mutating lysine to some other appropriate amino acids in the tail region in a fraction of histones. This will bring heterogeneity in electrostatic interactions, and according to our predictions this can lead to alterations in the width and packing density of chromatin.

To conclude, in this work we simulated chromatin in the lengthscale a few genes, accounting for various factors. Our results show the importance, heterogeneity in interactions and the role of non-histone proteins. This work should be considered as a step in the direction towards a more complete model to study chromatin states and dynamics of chromatin accounting realistic details like protein-binding and interactions due to histone modifications. We hope that this work will lead to further experimentation and computation.

References

1. Alberts, B., 2014. Molecular Biology of The Cell. Garland Science, Taylor and Francis Group, New York, 6 edition.
2. Kornberg, R. D., 1974. Chromatin Structure: A Repeating Unit of Histones and DNA. Science 184:868–871.
3. Luger, K., A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond, 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389:251–260.
4. 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.
5. Sun, J., Q. Zhang, and T. Schlick, 2005. Electrostatic mechanism of nucleosomal array folding revealed by computer simulation. Proc. Natl. Acad. Sci. U. S. A. 102:8180–8185.
6. Arya, G., Q. Zhang, and T. Schlick, 2006. Flexible histone tails in a new mesoscopic oligonucleosome model. Biophysical journal 91:133–150.
7. Richmond, T. J., and C. A. Davey, 2003. The structure of DNA in the nucleosome core. Nature 423:145–150.
8. Erler, J., R. Zhang, L. Petridis, X. Cheng, and J. C. Smith, 2014. The Role of Histone Tails in the Nucleosome : A Computational Study. Biophys J 107:2911–2922.
9. Biswas, M., K. Voltz, J. C. Smith, and J. Langowski, 2011. Role of histone tails in structural stability of the nucleosome. PLoS Comput. Biol. 7:2–13.
10. Arya, G., and T. Schlick, 2006. Role of histone tails in chromatin folding revealed by a mesoscopic oligonucleosome model. Proc. Natl. Acad. Sci. 103:16236–16241.
11. Graziano, V., S. E. Gerchman, D. K. Schneider, and V. Ramakrishnan, 1994. Histone H1 is located in the interior of the chromatin 30-nm filament. Nature 368:351–354.
12. Robinson, P. J., and D. Rhodes, 2006. Structure of the 30nm chromatin fibre: A key role for the linker histone. Curr. Opin. Struct. Biol. 16:336–343.
13. Collepardo-Guevara, R., and T. Schlick, 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. 40:8803–8817.
14. Maeshima, K., R. Imai, S. Tamura, and T. Nozaki, 2014. Chromatin as dynamic 10-nm fibers. Chromosoma 123:225–237.
15. Bednar, J., R. A. Horowitz, S. A. Grigoryev, L. M. Carruthers, J. C. Hansen, A. J. Koster, and C. L. Woodcock, 1998. Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. Proceedings of the National Academy of Sciences 95:14173–14178.
16. Song, F., P. Chen, D. Sun, M. Wang, L. Dong, D. Liang, R.-M. Xu, P. Zhu, and G. Li, 2014. Cryo-EM study of the chromatin fiber

bioRxiv preprint doi: https://doi.org/10.1101/453126; this version posted October 26, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

reveals a double helix twisted by tetranucleosomal units. Science 344:376–380.
17. Widom, J., and A. Klug, 1985. Structure of the 300A chromatin filament: X-ray diffraction from oriented samples. Cell 43:207–213.
18. Finch, J. T., and A. Klug, 1976. Solenoidal model for superstructure in chromatin. Proc. Natl. Acad. Sci. U. S. A. 73:1897–1901.
19. Wedemann, G., and J. Langowski, 2002. Computer simulation of the 30-nanometer chromatin fiber. Biophys. J. 82:2847–2459.
20. Langowski, J., 2006. Polymer chain models of DNA and chromatin. Eur. Phys. J. E. Soft Matter 19:241–249.
21. Perisic, 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.
22. Schiessel, H., W. M. Gelbart, and R. Bruinsma, 2001. DNA Folding: Structural and Mechanical Properties of the Two-Angle Model for Chromatin. Biophysical Journal 80:1940–1956.
23. Finch, J. T., and a. Klug, 1976. Solenoidal model for superstructure in chromatin. Proceedings of the National Academy of Sciences of the United States of America 73:1897–901.
24. Langmore, J. P., and J. R. Paulson, 1983. Low angle x-ray diffraction studies of chromatin structure in vivo and in isolated nuclei and metaphase chromosomes. J. Cell Biol. 96:1120–1131.
25. Pearson, E. C., P. J. Butler, and J. O. Thomas, 1983. Higher-order structure of nucleosome oligomers from short-repeat chromatin. EMBO J. 2:1367–1372.
26. Gerchman, S. E., and V. Ramakrishnan, 1987. Chromatin higher-order structure studied by neutron scattering and scanning transmission electron microscopy. Proc. Natl. Acad. Sci. U. S. A. 84:7802–7806.
27. Ghirlando, R., and G. Felsenfeld, 2008. Hydrodynamic studies on defined heterochromatin fragments support a 30 nm fiber having 6 nucleosomes per turn. J. Mol. Biol. 376:1417–1425.
28. Robinson, P. J. J., L. Fairall, V. A. T. Huynh, and D. Rhodes, 2006. EM measurements define the dimensions of the 30-nm chromatin fiber: Evidence for a compact, interdigitated structure. Proc. Natl. Acad. Sci. U. S. A. 103:6506–6511.
29. Belmont, A., 2006. A variable topology for the 30-nm chromatin fibre. Curr. Opin. Cell Biol. 8:8–13.
30. Routh, A., S. Sandin, and D. Rhodes, 2008. Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. Proc. Natl. Acad. Sci. 105:8872–8877.
31. Bassett, A., S. Cooper, C. Wu, and A. Travers, 2009. The folding and unfolding of eukaryotic chromatin. Curr. Opin. Genet. Dev. 19:159–165.
32. Szerlong, H. J., and J. C. Hansen, 2011. Nucleosome distribution and linker DNA: connecting nuclear function to dynamic chromatin structure. Biochem. Cell Biol. 89:24–34.
33. Eltsov, M., K. M. Maclellan, K. Maeshima, A. S. Frangakis, and J. Dubochet, 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. U. S. A. 105:19732–19737.
34. Nishino, Y., M. Eltsov, Y. Joti, K. Ito, H. Takata, Y. Takahashi, S. Hihara, A. S. Frangakis, N. Imamoto, T. Ishikawa, and K. Maeshima, 2012. Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. EMBO J. 31:1644–1653.
35. Bouchet-Marquis, C., J. Dubochet, and S. Fakan, 2006. Cryoelectron microscopy of vitrified sections: a new challenge for the analysis of functional nuclear architecture. Histochem. Cell Biol. 125:43–51.
36. Maeshima, K., and M. Eltsov, 2008. Packaging the Genome: the Structure of Mitotic Chromosomes. J. Biochem. 143:145–153.
37. Ou, H. D., S. Phan, T. J. Deerinck, A. Thor, M. H. Ellisman, and C. C. O'Shea, 2017. ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. Science 357.
38. Lieberman-Aiden, E., N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, and J. Dekker, 2009. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science 326:289–293.
39. Bickmore, W. A., and B. Van Steensel, 2013. Genome architecture: Domain organization of interphase chromosomes. Cell 152:1270–1284.
40. Grigoryev, S. A., G. Arya, S. Correll, C. L. Woodcock, and T. Schlick, 2009. Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. Proc. Natl. Acad. Sci. U. S. A. 106:13317–13322.
41. Bajpai, G., I. Jain, M. M. Inamdar, D. Das, and R. Padinhateeri, 2017. Binding of DNA-bending non-histone proteins destabilizes regular 30-nm chromatin structure. PLOS Computational Biology 13:1–19.
42. Collepardo-Guevara, R., and T. Schlick, 2014. Chromatin fiber polymorphism triggered by variations of DNA linker lengths. Proc. Natl. Acad. Sci. U. S. A. 111:8061–8066.
43. Grigoryev, S. A., 2018. Chromatin Higher-Order Folding: A Perspective with Linker DNA Angles. Biophys. J. 114:2290–2297.
44. Dootz, R., A. C. Toma, and T. Pfohl, 2011. Structural and dynamic properties of linker histone H1 binding to DNA. Biomicrofluidics 5:24104.
45. Arya, G., and T. Schlick, 2009. A Tale of Tails: How Histone Tails Mediate Chromatin Compaction in Different Salt and Linker Histone Environments. J. Phys. Chem. A 113:4045–4059.
46. Fan, Y., N. Korolev, A. P. Lyubartsev, and L. Nordenskiöld, 2013. An Advanced Coarse-Grained Nucleosome Core Particle Model for Computer Simulations of Nucleosome-Nucleosome Interactions under Varying Ionic Conditions. PLoS One 8.
47. Plimpton, S., 1995. Fast Parallel Algorithms for Short-Range Molecular Dynamics. Journal of Computational Physics 117:1 – 19.
48. Bannister, A. J., and T. Kouzarides, 2011. Regulation of chromatin by histone modifications. Cell Res. 21:381–395.

bioRxiv preprint doi: https://doi.org/10.1101/453126; this version posted October 26, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

49. Machida, S., Y. Takizawa, M. Ishimaru, Y. Sugita, S. Sekine, J.-I. Nakayama, M. Wolf, and H. Kurumizaka, 2018. Structural Basis of Heterochromatin Formation by Human HP1. Mol. Cell 69:385–397.e8.
50. Barski, A., S. Cuddapah, K. Cui, T.-Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev, and K. Zhao, 2007. High-Resolution Profiling of Histone Methylations in the Human Genome. Cell 129:823 – 837.