# The role of linker DNA in chromatin fibers

Brouwer, T.B.

# Chapter 4

# Rigid Base Pair Monte Carlo Simulations of 1-Start and 2-Start Chromatin Fiber Unfolding by Force

The organization of chromatin in 30-nm fibers remains a topic of debate. Here we quantify the mechanical properties of the linker DNA and evaluate the impact of these properties on chromatin fiber folding. We extended a rigid base pair DNA model to include (un-)wrapping of nucleosomal DNA and (un-) stacking of nucleosomes in 1-start and 2-start chromatin fibers. Monte Carlo simulations that mimic single-molecule force spectroscopy experiments of folded nucleosomal arrays reveal different stages of unfolding as a function of force and are largely consistent with a 2-start folding for 167 and 1-start folding for 197 nucleosome repeat length fibers. Nucleosome stacking appears better accommodated in 1-start than in 2-start conformations, and we suggest that this difference can compensate for the increased energy for bending the linker DNA. Overall, these simulations capture the dynamic structure of chromatin fibers while maintaining realistic, physical properties of the linker DNA.

This chapter is based on: de Jong B. E., Brouwer T. B., Kaczmarczyk A., Visscher B., and van Noort S. J. T.: Rigid Basepair Monte Carlo Simulations of One-Start and Two-Start Chromatin Fiber Unfolding by Force, **Biophysical journal**, 115.10, 2018: 1848-1859.

## 4.1 Introduction

The folding of chromatin fibers remains elusive despite more than three decades of research [1–7]. It is an important topic in biology because transcription regulation, and perhaps all processes involving DNA, are regulated by chromatin [8]. Post-translational modifications (PTMs) of both the histone tails and the globular parts of the histones have proven to be the hallmarks of activity on DNA that is wrapped around histone cores [9]. Some of these PTMs can be linked to structural features of the chromatin. For example, the acetylation of H3K56, which is indicative of transcriptionally active chromatin, increases the unwrapping of nucleosomal DNA [10]. Acetylation of H4K16, also prominent in transcriptionally active euchromatin, has been shown to affect fiber folding [11–13]. Other modifications, like methylation of H3K9, induce the transition from euchromatin to more condensed heterochromatin, mediated by Heterochromatin Protein 1 (HP1)[14]. Next to direct structural changes to the chromatin, PTMs regulate the binding of remodeling factors that add or remove other PTMs or reposition, assemble and/or disassemble nucleosomes. Although significant progress has been made in the structural biology of single nucleosomes and their complexes with other factors [7], a consensus on the structure of higher-order folded chromatin fibers is dearly lacking.

High-resolution crystallography [15] and, more recently, single-particle Electron Microscopy [16] of reconstituted fibers clearly points to a 2-start organization. Attractive interactions between odd and even nucleosomes in such fibers induce the formation of a double superhelix of stacked nucleosomes. Attempts to resolve the structure of native chromatin fibers by EM, super-resolution light microscopy or diffraction techniques, however, generally yield less order [17–19], if any, suggesting the absence of higher-order structure *in vivo*. Since the positions of the nucleosomes on the DNA are far from regular *in vivo*, it is hard to avoid disorder induced by variations in the lengths of the linker DNA in natively assembled chromatin fibers. To avoid such positional disorder of nucleosomes *in vitro*, chromatin fibers are typically reconstituted on tandem repeats of the strong Widom 601 nucleosome positioning sequence, which allows for perfect control of the nucleosome repeat length (NRL) [20]. Systematic variation of the linker length in steps of 10 base pairs and subsequent reconstitution, folding and EM inspection of highly condensed fibers with linker histone H5 yielded a diameter of 33 nm which was independent of linker length [20]. This observation is incompatible with a 2-start helix and suggests a 1-start structure for NRLs between 177 and 207 base pairs. Unfortunately, both the contrast and the relative disorder of the nucleosomes has prevented elucidation of the path of the DNA in most of the larger NRL fibers, leaving a robust structural interpretation in the open.

Single-molecule force spectroscopy provides a unique method to probe and manipulate the extension of folded chromatin fibers under physiological

conditions and can therefore contribute to a better structural understanding of the 30-nm chromatin fiber. In previous studies, we have reported force-extension data of 167 and 197 NRL fibers [21–23], which feature distinctive low-force unfolding characteristics that we have interpreted as resulting from 2-start and 1-start folding, respectively. However, this interpretation was challenged, suggesting that the data could equally well be explained by the gradual unwrapping of nucleosomal fibers without stacking interactions [24]. We observed that the force-extension curves with and without linker histones were indistinguishable up to the overstretching transition, which increased to forces above 5 pN when linker histones were present [22], raising the question how condensed chromatin fibers are organized at forces below the overstretching transition. Conceptually, one can envision that the force-extension curves can be modeled by the unfolding of three alternative structures: non-stacked, 1-start stacked or 2-start stacked chromatin fibers. Therefore, a more quantitative, physics-driven structural model, rather than qualitative arguing is required to interpret the force-extension data.

A large variety of early models for chromatin structure focused on experimentally observed geometric constraints to capture the high level of condensation achieved by chromatin [25–31]. Many of these models, however, cannot easily be adapted to include forced unfolding of chromatin fibers, though topological restrictions might exclude some of them. Full-atom Molecular Dynamics (MD) on the other hand would give a detailed structural insight, but the timescale of the simulations differs many orders of magnitude with the experimental results. Coarse-grained Monte Carlo (MC) models have successfully reproduced several features of the experimental force-extension data. Rippe *et al.* modeled chromatin fibers consisting of nucleosome- and DNA beads, in combination with a nucleosome-nucleosome interaction potential [32]. The typical chromatin overstretching plateau at 3-5 pN was attributed to breaking nucleosome-nucleosome interactions. Collepardo-Guevara and Schlick presented a model containing more flexible linker DNA, in combination with electrostatic interactions arising from charges that were distributed along the coarse-grained DNA and histone proteins [28]. Their model could also reproduce the typical low-force overstretching plateau when nucleosome-nucleosome stacking interactions, typically in 2-start fibers, were broken. In the current issue, Zhurkin *et al.* employ a rigid base pair model in combination with geometrical constraints to compute the DNA trajectory of the linker DNA in 2-start fibers with high quantitative agreement. Again, the low-force overstretching plateau is reproduced, but now in more detail. It is proposed that gradual unwrapping of nucleosomal DNA can account for the linear stretching region, between 1 and 3 pN, before the overstretching plateau sets in.

Although it may seem that current coarse-grain models suffice to explain experimental force-extension curves, there are several features in the collective dataset that cannot be captured in the above described 2-start geometry and

that are better explained in a 1-start geometry [21, 23, 33]. In this paper, we first reiterate our experimental findings on pulling individual folded chromatin fibers and highlight the distinctive features between 167 NRL fibers that fold in a prototypical 2-start geometry, and 197 NRL fibers that were suggested to fold in a 1-start fiber. We have shown that both 1-start and 2-start geometries fit the same four-state statistical physics model, with very similar dimensions of the states and transition energies. Although these fitted parameters roughly match structural data of the nucleosome, they do not provide insight into the precise energetics and structural transitions at the microscopic level. Therefore, we adapt a rigid base pair Monte Carlo framework to simultaneously include un- and re-wrapping of nucleosomal DNA, bending and twisting of linker DNA, as well as (un-)stacking of nucleosomes in various fiber geometries. Doing so highlights the well-known mechanical properties of DNA as a major parameter in chromatin fiber folding.

We will use the same step-parameter framework for nucleosomal DNA wrapping and nucleosome stacking as for the base pairs in bare DNA and reproduce force-extension curves for different fiber compositions and folding geometries. Though the kinetics of unstacking could not be reproduced in our simulations, we obtained fair agreement with the experimental results in the low force regime where the fibers are fully folded. This allowed for a detailed analysis of the energetics of the linker DNA. Thus, we aim for a quantitative description of the processes during force-induced fiber unfolding and a structural understanding of the loosely organized chromatin fibers before nucleosomes start to unstack and unwrap.

## 4.2 Results

### 4.2.1 Comparison of experimental unfolding of chromatin fibers with different nucleosome repeat lengths

We reconstituted nucleosomes on 7040 base pairs DNA containing $30 \times 167$ and 7125 base pairs DNA with $25 \times 197$ repeats of the Widom 601 nucleosome positioning sequence to create chromatin fibers with the approximately the same DNA contour length. Figure 4.1 shows experimental force-extension curves of a 167 NRL fiber (red) and a 197 NRL fiber (green). Both curves feature three large transitions in extension: at 65 pN the DNA overstretches into a 1.6 times longer overstretched structure. At 15 – 30 pN the nucleosomes yield a single turn of DNA, visible in 33 distinct 25 nm steps for the 167 NRL fiber and 32 steps for the 197 NRL fiber. Both transitions are similar for 167 and 197 NRL and therefore independent on linker length.

Between 3 and 5 pN, a chromatin overstretching transition occurs that is different for both fibers. The slope of the 167 NRL fiber is almost four times larger, the maximum extension is two times less and the transition of

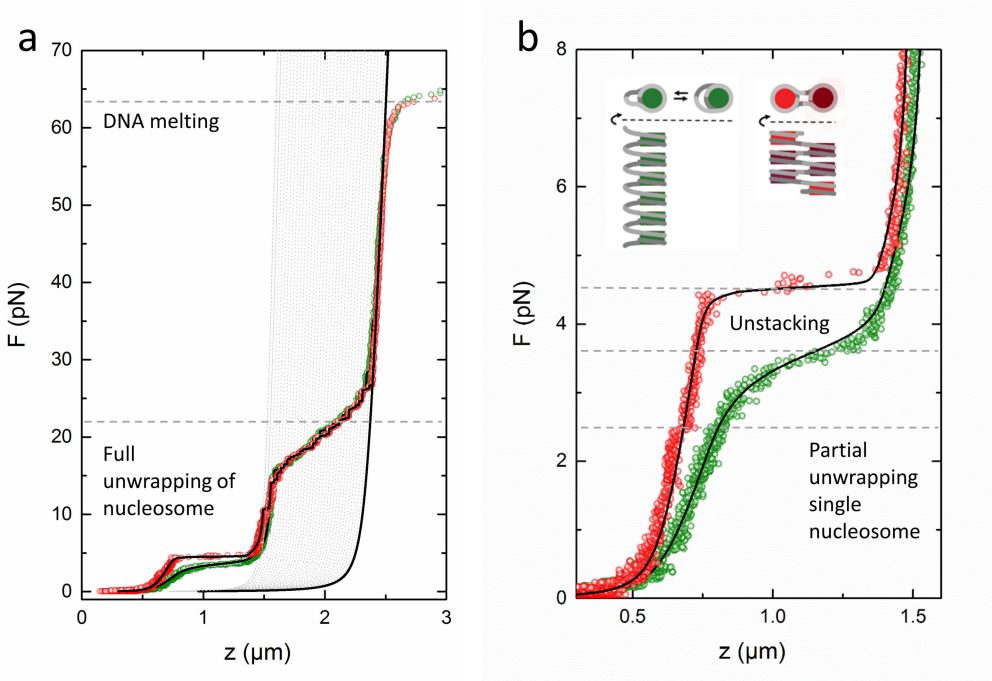


**Figure 4.1**
**Experimental force-extension curves of individual folded 167 NRL (red) and 197 NRL (green) chromatin fibers show a difference in stretching and unfolding at forces below 5 pN.** a) At 65 pN the curves exceed the extension of a Worm Like Chain (WLC) as the DNA overstretches. Multiple 25 nm steps between 10 and 30 pN indicate the unwrapping of the last turn of DNA form each nucleosome. Grey dotted lines correspond to WLC curves that are each 80 base pairs shorter, corresponding to DNA released from a single nucleosome. b) Zoom in of the low force transitions. The 167 NRL fiber unstacks at 4.5 pN in a sharp transition. The 197 NRL fiber unstacks more gradually at 3.5 pN. In addition, a small, 25 nm step is visible in the red curve at 2.5 pN that we attribute to the unwrapping of a single, isolated nucleosome. The inset schematically shows the structural difference between a 1-start (red) and a 2-start (green) fiber. Black lines show a fit to a statistical mechanics model that captures the low force transitions, as described in [23] and a WLC. The fitted parameters and standard errors of the 167 NRL fiber, reconstituted on a 7040 base pairs DNA substrate containing $30 \times 167$ repeats of the 601 sequence were: number of nucleosomes = 30, number of tetrasomes = 3, stiffness = $0.97 \pm 0.03$ pN/um, $\Delta G_1 = 20.0 \pm 0.1\ k_BT$, $\Delta G_2 = 5.7 \pm 0.1\ k_BT$. The fitted parameters and standard errors of the 197 NRL fiber, reconstituted on a 7125 base pairs DNA substrate containing $25 \times 197$ repeats of the 601 sequence were: number of nucleosomes = 24, number of tetrasomes = 8, stiffness = $0.31 \pm 0.01$ pN/um, $\Delta G_1 = 20.0 \pm 0.1\ k_BT$, $\Delta G_2 = 5.8 \pm 0.1\ k_BT$.

the 167 NRL fiber is at a higher force and is sharper than that of the 197 NRL fiber. These characteristic differences reproduce our previous reports [21, 23, 33]. It is noteworthy to mention that the chromatin overstretching plateau occurs at 1 pN larger force than the unwrapping of a single nucleosome under identical experimental conditions, indicating that the force to rupture interactions between the nucleosomes exceeds that of the unwrapping of DNA around the histone core in single nucleosomes [34]. In fact, close inspection of the 167 NRL curve in Figure 4.1b reveals a small step in extension at 2.5 pN, suggesting that a single nucleosome partially unwraps at this force. The detailed difference in structural transitions at low force reflects the (un-)folding of the chromatin fibers and suggests a different geometry of the two fibers before unfolding.

A statistical physics model that we developed previously [23], in which we distinguish four states of chromatin folding, was fitted to the experimental data and is represented by the black lines in Figures 4.1a and b. With increasing force, the nucleosomes undergo a transition from a stacked state, via a partially unwrapped state and a single-turn wrapped state to a fully unwrapped state [23]. Transitions between each of the states are accompanied by a free energy cost, which is compensated by the additional work that is done through the increased extension of each subsequently unfolded state. The stiffness and unstacking energy per nucleosome were independent of the number of nucleosomes in the fiber (*see* Supplementary Figure 4.S1). More details and the observation that small compositional differences between fibers can dominate the force-extension curve can be found elsewhere [23]. Here we want to stress that all fitted parameters are highly reproducible between curves and can be summarized in the above described trends.

The higher stiffness, smaller maximum extension and cooperative unfolding of the 167 NRL fiber are all consistent with having two stacks of nucleosomes, like in the 2-start fiber. This can easily be explained in terms of a simple toy-model. When the connecting linker DNA is ignored, it is clear that splitting a single stack of nucleosomes into two stacks of half the number of nucleosomes, quadruples the stiffness and halves the maximum extension. As the force would be divided between the two parallel stacks, each stacked nucleosome pair would be exposed to less force and thus yield a higher rupture force. The sharp overstretching transition can be understood when including the linker DNA in the 2-start helix. In that case, the internal nucleosomes are further stabilized by their neighbors in the other stack, while the end-nucleosomes miss such stabilization, resulting in a cooperative rupturing of the fiber that starts with the end-nucleosomes. The absence of cooperativity in 197 NRL fibers is therefore another indication of a 1-start helix. We recently substantiated this interpretation by covalently linking the H4 tails to the acidic patch of their stacking partner, which yielded an increased rupture force, but unchanged stiffnesses for both types of fibers as compared to the non-crosslinked fibers [33].

Thus, single-molecule force spectroscopy reveals qualitative and quantitative differences in stretching and unfolding of the higher-order structures of different NRLs that can be rationalized by an altered stacking order of the nucleosomes.

Though the DNA topology in 167 and 197 NRL fibers may be different, the local stacking interaction between two interacting nucleosomes should be similar. From the extension of the fibers at low forces, it is clear that the stacking of nucleosomes in a folded fiber cannot be closely packed such as in nucleosome crystals [15]. We measure a maximum extension per nucleosome of 10 nm, which is significantly more than the 6 nm height of the globular part of the nucleosome [35]. This gap can just be spanned by the H4 histone tails that mediate nucleosome stacking. In the high-resolution cryo-EM structure [16] and recent force spectroscopy data at very low salt [36], nucleosomes appear to be organized in tetramers with pairs of more closely packed nucleosomes alternated by pairs of nucleosomes featuring larger gaps. Under more physiological conditions, recent single-pair Förster Resonance Energy Transfer (spFRET) measurements suggest a highly dynamic organization in which stacking pairs may rapidly change [37]. Nevertheless, we expect that nucleosome stacking, mediated by very flexible, unstructured histone tails that transiently interact, is locally very similar in the two geometries. This implies that linker DNA largely defines the global structure of the chromatin fiber.

### 4.2.2 Extension of a rigid base pair model with nucleosomes

To model the linker DNA in folded chromatin fibers we extended the HelixMC package [38] to include nucleosomal DNA wrapping and nucleosome stacking. The HelixMC package was based on the work of Olson and coworkers, who captured the mechanical properties of DNA in six sequence-dependent kinematic parameters [39]. Here we include two extra layers of kinematic parameters that define DNA wrapping around the histone core and stacking of nucleosomes in the fiber. We used the DNA step parameters of the 1kx5 crystal structure of the nucleosome [35] to replace the step parameters in a random pose of DNA. The frames of 14 base pairs with the lowest B-factor were each assigned an energy of up to 2.5 $k_{\mathrm{B}}T$ to account for deviations from the crystal structure, *see* Figure 4.2.

Figure 4.3 shows the results of Monte Carlo simulations of nucleosomal arrays consisting of 8 non-interacting nucleosomes reconstituted on an 1800 base pairs DNA substrate. In 50,000 steps, the force was ramped linearly from 0.1 pN to 10 pN and back. In 1000 preceding steps, the nucleosomes were forced in a stacked conformation (for proper comparison with other simulations, *see* below), but since there is no nucleosome-nucleosome interaction potential in the remainder of the simulation, the nucleosomes behave largely independent in the remaining trajectory. These initial 1000 steps were discarded. Snapshots of the initial structure up to the structure at 10 pN are shown below the

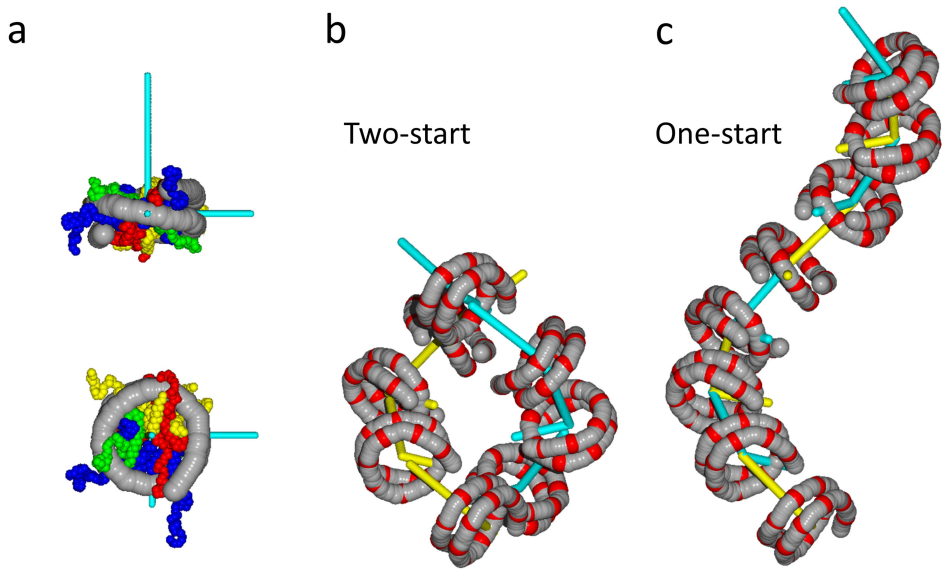


**Figure 4.2**
**Two additional layers of step parameters define the folded chromatin fiber.** a) A nucleosome frame (cyan) was created based on the coordinates of the 14 base pairs that are fixed in the nucleosome (colored red in panels b and c), as determined from the local minima of the average B-factors of the DNA coordinates in nucleosome structure 1kx5 [35]. The center of mass of the central eighth fixed base pair forms the origin of the frame. The $x$-axis points to the dyad, defined by the base pair that is halfway between the central two fixed base pairs. The $z$-axis is parallel to the vector connecting the center of mass of the first seven fixed base pairs with the center of mass of the last seven fixed base pairs. b) A 2-start helix is constructed by arranging odd nucleosomes (yellow frames) and even nucleosomes (cyan frames) in a helical structure. The $y$-vectors of the nucleosome frames were omitted for clarity. c) Stacking odd nucleosomes on even nucleosomes defines a 1-start helix. Note that the relative orientation of the nucleosomes is highly similar in both structures and the parametrization of the 6 kinematic parameters of these structures is used to impose fiber folding in the MC simulations.

force-extension curve. Movies of the structures corresponding to all traces are available as supplemental data. Both force-extension curves feature an unwrapping transition at 2.5 - 3 pN and converge to a wormlike chain (WLC) with a contour length of $1800\breve{}8 \times 80 = 1160$ base pairs, corresponding to a single wrap of DNA around each nucleosome, in close agreement with experimental data [34]. Interestingly, the nucleosomes in the 167 NRL array unwrap at higher forces than those in the 197 NRL array, probably because unwrapping is somewhat held up by steric hindrance of the nearby nucleosomes.

Overall, these simulations largely reproduce the experimental behavior of 8 times a single nucleosome, which features the unwrapping of the first turn of DNA at 2.5 pN [34]. The increased condensation of the $8 \times 197$ NRL fiber

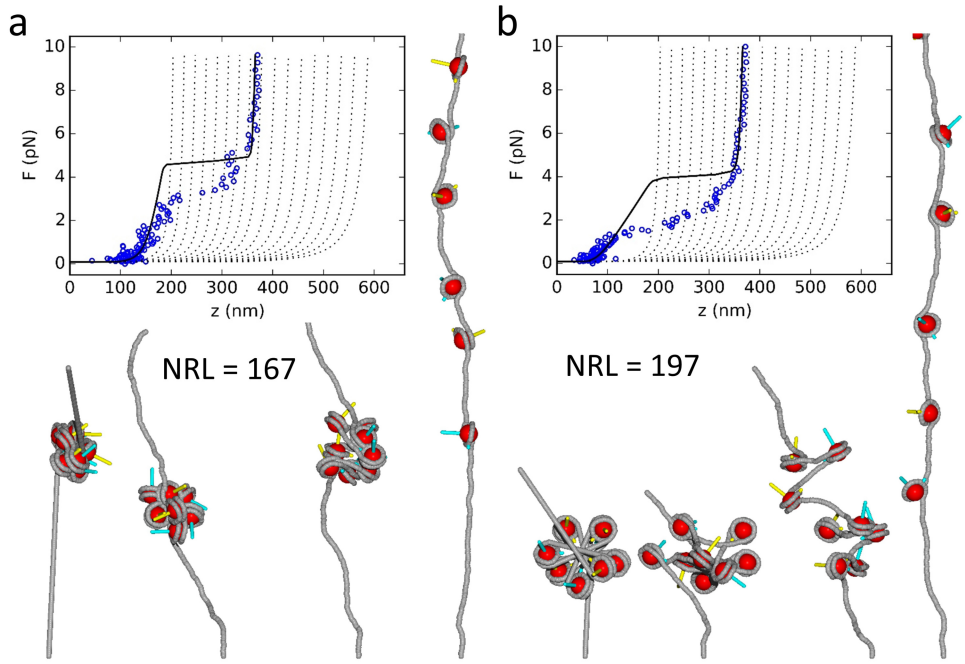


**Figure 4.3**
**MC simulations of arrays of non-interacting nucleosomes that readily unfold when the force is increased up to 10 pN.** a) A simulated force-extension curve of an 1800 base pairs 167 NRL fiber containing 8 nucleosomes. b) Simulated force-extension curve of an 1800 base pairs 167 NRL fiber containing 8 nucleosomes. The nucleosomes release their first turn of DNA at 2.5 pN. Black lines show the statistical mechanics model for fiber unfolding using parameters fitted to experimental curves for both fibers. Grey dotted lines are plotted for reference and represent WLC curves in which the contour length is increased for both unwrapping events of each nucleosome. The snapshots below the curve show the structures of the initial conformation (left) and conformations at 0.1, 1 and 10 pN. Note that the top part of the tallest structures is clipped from the figure.

at low forces relative to a WLC with a contour length that is reduced by $8 \times 147$ base pairs should be attributed to the kink in the DNA trajectory that is induced by the nucleosomes [40]. Strikingly, at forces below 1 pN, the simulated force-extension curves overlap with the statistical physics model based on the experimental data, which is plotted in black, suggesting that this level of condensation can be achieved without attractive nucleosome-nucleosome interactions. However, the significantly higher level of the overstretching plateau in the experimental data suggests that we need nucleosome-nucleosome stacking interactions to obtain a better match.

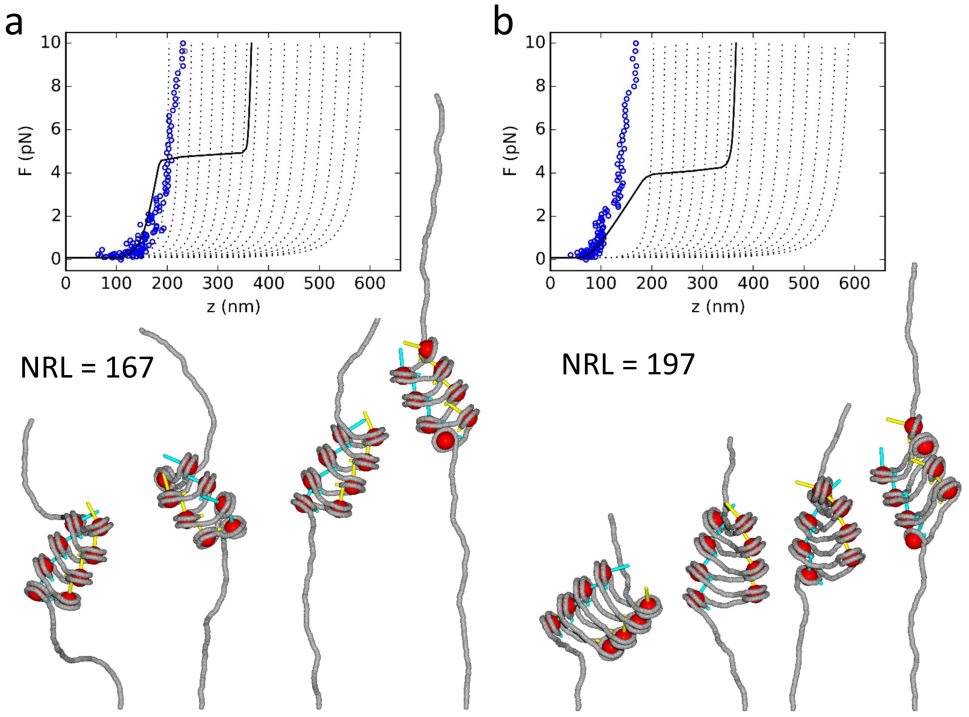


**Figure 4.4**
**MC simulations of the stretching of 2-start fibers show a high stiffness.** a) A simulated force-extension curve of a 2-start 167 NRL fiber containing 8 nucleosomes. b) The force-extension curve of a 2-start 197 NRL fiber. The snapshots below the curves show the structures of the fiber at 0.1, 1, 2 and 10 pN.

### 4.2.3 Nucleosome stacking in 2-start fibers

The linker DNA in 2-start fibers does not bend severely to accommodate nucleosome stacking. The structures in Figure 4.4 show nevertheless that a balance between optimal stacking and minimizing the bending of the linker DNA is found in the simulations. Optimal stacking would place the ends of the colored nucleosome frame $z$-vectors in the center of the next nucleosome, which is clearly not the case for both fibers. In the simulations, we do not observe the typical overstretching transition at 3 – 5 pN, even when the force is increased up to 10 pN. The stacked nucleosomes appear kinetically trapped, as we estimate that the work associated with unwrapping and unstacking at 10 pN to be about twice the used stacking energy. Only the outer two nucleosomes feature DNA unwrapping, but the stacking is maintained throughout fibers. This is consistent with a cooperative transition for overstretching the 2-start fibers. However, because detailed balance was only achieved when nucleosomes

remain stacked, we refrain from a kinetic interpretation at higher forces.

The extension of the fibers agrees with the experimental extension at low forces but deviates at forces exceeding 1 pN. In the 167 NRL fiber, a larger extension is found in the simulations that is caused by the partial unwrapping of the outer nucleosomes as depicted in the snapshots below the curve. It never exceeds the equivalent of 60 base pairs of unwrapped DNA. Partial unwrapping of the outer nucleosomes also occurs in the 197 NRL fiber, but this is not sufficient to reach the experimentally observed extensions for this type of fibers. Some extra extension, corresponding to roughly the linker length, is gained at higher forces by shearing the two stacks of nucleosomes. Both the unwrapping at the ends and the shearing of the stacks do not scale with the number of nucleosomes though, so we expect that these effects become negligible in simulations of larger arrays of nucleosomes, as used in the experiments. The simulated force-extension curve of a 2-start 167 NRL fiber would overlap with the experimental curves up to the overstretching plateau, when end-effects are corrected for, but the stiffness of the simulated 2-start 197 NRL fiber is significantly larger than measured experimentally, suggesting that this fiber folds in a different manner.

### 4.2.4 Nucleosome stacking in 1-start fibers

The force-extension curves in Figure 4.5 started with fiber structures that were forced into a 1-start helix before ramping up the force. Although the linker DNA features big curvatures, especially in the case of the 167 NRL fiber, the nucleosomes largely remain trapped in the stacked state over the entire force range. In both cases the stacking seems to be better optimized than for the 2-start fibers, as the $z$-vectors of the nucleosome frames are much closer to the centers of the next nucleosomes. At larger forces, the stacking gets compromised, as the helical stack of nucleosomes is extended. The end nucleosomes are prone to partial unwrapping, like in the 2-start structures. In the 197 NRL fiber, we observed a rare event of a stacking defect at about 1 pN, which leads to further unwrapping and the release of extensive lengths of linker DNA at larger forces. Overall, 1-start 197 NRL fibers were more fragile than 1-start 167 NRL fibers, despite the larger bending stress in the latter. The larger work associated with the unstacking in 197 NRL fibers makes them more susceptible to force.

Compared to the experimental model, the 1-start 167 NRL fiber was more extended at forces larger than 1 pN. For the 1-start 197 NRL fiber there is a better match. However, the stiffness of the simulated fiber exceeded that of the experimental model. Though the simulations do not perfectly capture the force-extension behavior in this case, the stiffness is roughly 4 times smaller than that of the 2-start fiber. Further relaxing of the stacking step parameters may yield a better agreement with the experimental data, but the ratio of stiffnesses confirms the toy-model comparison of 1-start versus 2-start stretching.

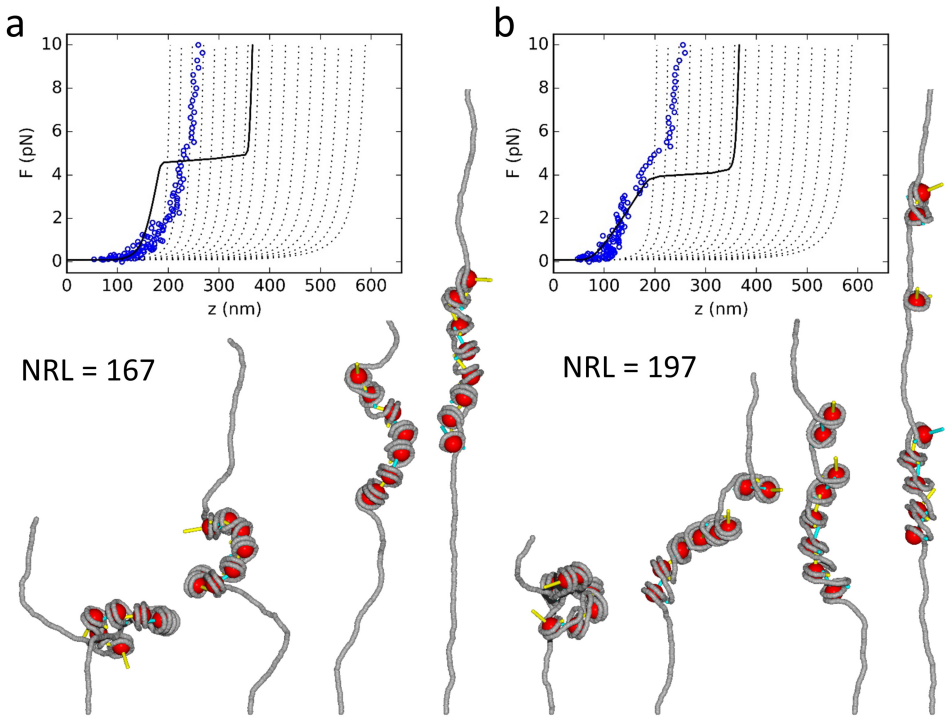


**Figure 4.5**
**MC simulations of the stretching of 1-start fibers feature a lower stiffness.** a) A simulated force-extension curve of the 1-start 167 NRL fiber containing 8 nucleosomes. b) The force-extension curve of a 1-start 197 NRL fiber. The snapshots below the curves show the structures of the fiber at 0.1, 1, 2 and 10 pN.

### 4.2.5 MC simulations of larger chromatin fibers

To better compare the MC simulations with our experimental data we simulated fibers containing 15 nucleosomes at forces increasing up to 3 pN and subsequent release, back to 0.1 pN. Since this is well below the chromatin fiber overstretching transition, that we attributed to nucleosome unstacking, the corresponding large increase in extension was fully absent, as shown in Figure 4.6. The extension of the 167 NRL 2-start simulated data (Figure 4.6a) showed over the entire force range good agreement with the modeled curve, plotted in black. The 197 NRL fiber was more condensed (Figure 4.6b), both in the simulations and in the model curve. However, the simulated data featured a higher stiffness than the model based on experimental data.

Whereas the extension of the 167 NRL fiber was slightly less than the extension that can be expected for non-interacting nucleosomes, the 197 fiber was more than 100 nm shorter over most of the force range. At very low forces

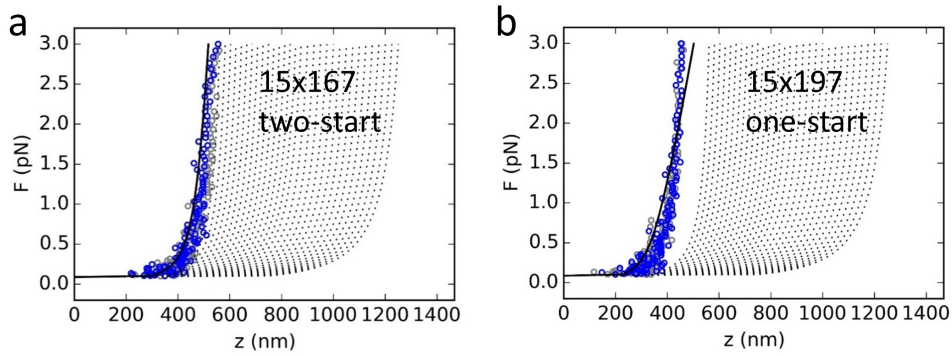


**Figure 4.6**
**MC simulations of 4000 base pairs fibers containing 15 nucleosomes reproduce experimental data up to the unstacking transition.** a) A 167 NRL fiber in a 2-start conformation. b) A 197 NRL fiber in a 1-start conformation. Blue dots represent pulling curves, grey dots the subsequent release curve and black line the statistical physics model with a stiffness corresponding to the stiffness that was fitted to the experimental data.

however, the difference is negligible. In a separate simulation, we tracked the extension of these fibers in absence of force (data not shown). We obtained an extensions $z = 0.29 \pm 0.09$ µm, $0.28 \pm 0.09$ µm and $0.26 \pm 0.09$ µm (mean ± s.d.) for 167 NRL without stacking, 167 NRL fibers in a 2-start structure and 197 NRL fibers in a 1-start structure. Note that the extension of the fibers was dominated in this case by the large DNA handles (1535 base pairs and 1145 base pairs for $15 \times 167$ NRL and $15 \times 197$ NRL DNA substrates), similar to DNA substrates used in the experiments.

The overlap of the pull and release simulations shows that the MC simulation is fully reversible. The condition of detailed balance was corroborated by analysis of the autocorrelation of the extension and the unwrapping energy at zero force, shown in Supplementary Figure 4.S2. Correlation in the extension could be observed up to 30 steps in all fiber conformations. The unwrapping was somewhat slower, with vanishing correlations after approximately 200 MC steps for stacked nucleosomes and approximately 75 MC steps for non-interacting nucleosomes.

Like in the shorter fibers, we observed only a small amount of unwrapping in the 15 nucleosome fibers. In absence of force, the unwrapping energy per nucleosome was $0.7 \pm 0.5$ $k_\mathrm{B}T$, $0.6 \pm 0.3$ $k_\mathrm{B}T$ and $0.3 \pm 0.3$ $k_\mathrm{B}T$ (mean ± s.d.) for 167 NRL without stacking, 167 NRL in a 2-start structure and 197 NRL in a 1-start structure respectively. The energy penalty for the release of 10 base pairs of DNA was set 2.5 $k_\mathrm{B}T$, so at any time in the simulation only roughly a quarter of the nucleosomes featured the release of one out of fourteen histone DNA interactions. The small variations in unwrapping energy

| NRL (fiber) | 167 (1-start) | 167 (2-start) | 197 (1-start) | 197 (2-start) |
|---|---|---|---|---|
| **Shift** | 0.5 ± 0.4 | -0.3 ± 0.3 | -0.7 ± 0.3 | -0.6 ± 0.4 |
| **Slide** | 1.3 ± 0.4 | 0.1 ± 0.2 | 0.3 ± 0.1 | 0.4 ± 0.5 |
| **Rise** | 1.6 ± 0.3 | 0.0 ± 0.4 | -0.7 ± 0.4 | -0.4 ± 0.7 |
| **Tilt** | 4.9 ± 0.6 | 1.3 ± 0.2 | 3.8 ± 0.7 | -1.8 ± 0.7 |
| **Roll** | 10.3 ± 0.1 | 2.5 ± 0.7 | 8.7 ± 0.6 | -1.3 ± 1.0 |
| **Twist** | 6.1 ± 0.8 | -0.6 ± 0.2 | 1.0 ± 1.2 | -1.6 ± 1.6 |
| **$\sum$ step** | **24.7 ± 2.3** | **3.1 ± 0.5** | **12.4 ± 1.8** | **-5.3 ± 2.3** |
| **Wrap** | 0.6 ± 0.4 | -0.3 ± 0.1 | -0.4 ± 0.2 | -0.6 ± 0.1 |
| **Stack** | -22.7 ± 0.2 | -11.3 ± 0.5 | -19.8 ± 1.7 | -11.8 ± 0.4 |
| **Total** | **2.6 ± 2.8** | **-8.6 ± 0.7** | **-7.8 ± 0.4** | **-17.6 ± 2.0** |

**Table 4.1**
**Quantification of the energy changes upon stacking of nucleosomes in 1-start and 2-start fibers.** Stacked nucleosomes in 2-start 167 NRL fibers and 1-start 197 NRL fibers can both reduce their energy with approximately 8 $k_BT$ when nucleosomes stack. For comparison, the average base pair step energies were calculated between $F = 0.1$ and 1.5 pN and summed between pairs of neighboring dyads. The corresponding energies of the non-stacking fibers (shown in Figure 4.3) were subtracted. All energies are expressed in units of $k_BT$. Errors represent standard deviations obtained from 4 separate simulations. $\sum$step is the sum of the energies that correspond to the 6 degrees of freedom of the base pair and quantifies the penalty for linker DNA bending and twisting. Together with the wrapping and the stacking energy we calculated the total change in energy upon nucleosome stacking in this framework.

confirmed the absence of significant transient unwrapping. Thus, though our model only imposed wrapping of 132 out of 147 base pairs of the nucleosome, further unwrapping of nucleosomal DNA did not significantly contribute to stabilization of the stacking interactions.

### 4.2.6 Evaluation of the energies associated with fiber folding

The MC simulations not only result in 3D structures of the chromatin fibers with force-extension curves that compare well with the experimental data, they also allow for a detailed comparison of the energies associated with the dynamic structures. In Table 4.1 we summarized the energy contributions of the folded fibers relative to fibers with the same NRL but without stacking interactions. To cancel the large variations between static structures, we averaged the data between 0.1 and 1.5 pN, in which force range all nucleosomes were stably stacked. We calculated the energies from the step parameters and the stiffness matrices. To exclude end-effects we only considered DNA sections from dyad to dyad.

As expected, the 1-start fibers feature a large energetic penalty for bending the linker DNA which is exemplified in the energy differences defined by roll of the base pairs: 10.3 $k_BT$ and 8.7 $k_BT$ for 167 NRL and 197 NRL fibers. The 1-

start 167 NRL fiber also has excessive energies for tilt and twist. For the 2-start 167 NRL fiber it appears that excluded volume effects and unfavorable twisting of the linker DNA to accommodate stacking do not favor linker conformations relative to those of non-stacking nucleosomes. Only in the 2-start 197 NRL fiber is the energy of the linker DNA reduced as compared to non-stacking nucleosomes.

Previously, we suggested that partial unwrapping of the nucleosomal DNA may relieve some of the bending of the linker DNA in 1-start fibers [23]. In the current simulations, it appears that this is not the case: the averaged energy associated with unwrapping is within 1 $k_{\mathrm{B}}T$ for all fibers. Note, however, that we fixed the histone-DNA contacts in the nucleosome at 14 frames, which constrains only 132 base pairs. The outer 15 base pairs were therefore free to move and this allowed for substantial reduction of the bending of the linker DNA. The small values and fluctuations of the wrapping energy, relative to 2.5 $k_{\mathrm{B}}T$ per contact point, indicate that the central 132 base pairs of nucleosomal DNA remain firmly wrapped around the histone core at low forces.

Surprisingly, we find the largest variation in the stacking energy. 1-start fibers appear to be much better positioned to optimize stacking between nucleosomes than 2-start fibers. This is not caused by steric hindrance of the linker DNA, as both structures can fold in the same relative orientation of stacked pairs of nucleosomes without clashes of the linker DNA. It rather seems that the linker DNA of a pair of nucleosomes in a 2-start conformation positions the nucleosome in between in an unfavorable orientation for stacking with its next neighbor. Excluded volume effects between the nucleosomes, which was included in the MC simulations, may complement to this effect. Though stacking is better in 1-start fibers, none of the fibers geometries position the nucleosomes in a perfect orientation, yielding a realized stacking energy that is reduced 3 to 13 $k_{\mathrm{B}}T$ less than the set stacking energy.

When all energy contributions are summed, 2-start 167 NRL fibers and 1-start 197 NRL fibers are both about 8 $k_{\mathrm{B}}T$ per nucleosome lower in energy than fibers consisting of non-interacting nucleosomes. The 2-start 197 NRL fiber appears to be most stable, reducing the energy another 10 $k_{\mathrm{B}}T$. The current simulations, however, do not include entropy calculations and therefore miss an important part of the free energy that would be required for a full comparison. Moreover, the simple harmonic stacking potential may not properly weigh non-optimal stacking of nucleosomes. This may bias the stacking energy towards 2-start conformations. It is clear though that each fiber configuration needs to balance optimal nucleosome stacking with minimal linker DNA deformation and that small changes in linker length have a large impact on the energies and the resulting fiber structures.

## 4.3 Discussion

Single-molecule force spectroscopy of folded chromatin fibers provides a unique way to probe chromatin conformations and dynamics under physiological conditions. Though the extensions of single fibers can be measured with nanometer accuracy and manipulated with pN precision, the resulting force-extension curves cannot unequivocally discriminate between different folding topologies. In this paper, we presented a framework of rigid base pair MC simulations supplemented by nucleosome wrapping and nucleosome stacking step parameters that provides quantitative insight into the energy contributions that define folded chromatin fibers. We simulated two frequently studied NRLs and evaluated the force-extension relation and the energy contributions of three different nucleosome stacking options. The results largely agree with previous interpretations of the force-extension curves and uniquely quantify and visualize the delicate balance between optimal nucleosome stacking and minimal deformation of the linker DNA. The simulations disclose the very dynamic structure of the fibers that can best be seen in the movies (*see* supplemental information online). The relatively disordered structures, despite strong stacking interactions and the stiff and stressed linker DNA, may be inhibitive for experimental techniques that rely on averaging, which would partially explain the reported absence of chromatin higher-order structure [41].

For 167 NRL fibers, both the force-extension curves and energy calculations favor the well-known 2-start conformation. This NRL appears to be optimal for a 2-start helix as the energy required for twisting the linker DNA is only 0.6 $k_{\mathrm{B}}T$ and can hardly be reduced. Addition of a single base pair will increase the twist by 36 degrees and thus increase the twist energy, which may prevent stacking of nucleosomes in larger NRLs. Future simulations and experiments will need to confirm this hypothesis. Despite the ideal linker length, the stacking is far from optimal. The simulations did not yield the tetra-nucleosomal patches that were observed in a crystal structure [15], EM reconstruction [16], force-spectroscopy [36], and FRET experiments [37]. Such an arrangement would provide an alternative compromise for balancing linker bending and nucleosome stacking. Nevertheless, the quantitative agreement with experimental force-extension data shows that fibers can be stretched up to 5 nm per nucleosome while remaining fully stacked and wrapped.

In experiments, the 197 NRL fibers yield larger extensions before overstretching and this trend is only reproduced in the case of the 1-start helix. Though the simulations do not fully catch the experimentally obtained large extension and low stiffness, it is clear that the 1-start conformation is a better match than the 2-start conformation, which cannot reach such large extensions. Nevertheless, the 2-start conformation yields the lowest energy for 197 NRL fibers, since the linker DNA is relatively straight. However, the experimentally observed extensions can only be achieved when nucleosomes unstack, which

costs more energy than the alternative scenario in which a 1-start is formed with the same extension. Contrary to popular belief, the energy calculations show that the 1-start helix does not present an inhibitively large penalty for linker DNA bending. The ability to find a better stacking conformation largely compensates the 12.4 $k_BT$ that is required for bending the linker DNA. This may make the 1-start helix the preferred structure for 197 NRL fibers. The simulations show that the experimentally obtained force-extension curves of 167 NRL fibers can be fully reproduced by stretching stacked nucleosomes in a 2-start conformation. The experimental force-extension curves of 197 NRL fibers give a better, though not perfect, match with a 1-start conformation. Yet, the values of the stacking energy of fibers with this repeat length appear to favor a 2-start conformation. We attribute this discrepancy to the simplified calculation of the stacking energy, for which we used a harmonic potential. A potential that would penalize large deviations stronger may rectify this. Indeed, when stacking is mediated by the H4 tails, one would expect a non-harmonic potential. Nevertheless, despite the simplified interaction potential, the simulations show that the force-extension curves of both fibers can be largely reproduced, up to the overstretching transition, without breaking nucleosome-nucleosome interactions.

The step parameter approach appears to work well for nucleosome (un-)wrapping. The energy in each histone-DNA contact point is similar as reported before [42] and the force at which the first unwrapping transition occurs agrees with the experimentally observed force [34]. The unwrapping is reversible and simulated force-extension curves do not show hysteresis, indicating thermodynamic equilibrium. Like in experiments, the second unwrapping transition is not in equilibrium, but the force at which this transition occurs is much higher in the simulations than in the experiments (data not shown). Likewise, the unstacking transition is only rarely observed in the simulations up to 10 pN. Longer simulations may catch the dynamics of the force-extension curves better, but more extensive simulations need to be performed to test whether these transitions can be accurately captured in this MC model or that more efficient sampling strategies are required. Note that the timescale of MC simulations cannot be coupled directly to the experimental pulling rates. The current simulations do, however, nicely match the experiments in which we cross-linked the H4 tail to the acidic patch on the neighboring nucleosomes [33], showcasing their relevance for interpretation of the experimental data.

Optimization of the stiffness matrix that defines the stacking interactions may further improve the simulations. In absence of experimental input, we postulated a simple stiffness matrix with only diagonal terms and implemented a slightly non-linear response. The used stiffness values form a compromise between high flexibility and imposing only very local interactions. The decisive role of the H4 tails in nucleosome stacking suggests that the interactions may be better modeled by a potential that describes the extension of a Freely

Jointed Chain. Other, more refined models, including all-atom simulations [27] can of course provide much more detailed insight. Such models will be able to include the role of electrostatics, histone tails, PTMs and other essential ingredients of chromatin biology. However, because the local geometry between two stacked nucleosomes is the same in all stacking conformations that we simulated, the current approach may yield a fair comparison in which those more local effects remain the same between different geometries. Moreover, coarse-grained structures as presented here may be the input for more refined models.

Awaiting such larger computational efforts, we aim to make the discussion about chromatin fiber structure and the interpretation of force-spectroscopy experiments on these fibers more quantitative with our simulations. The highly dynamic features of the chromatin fibers advocate a shift in the paradigm of chromatin structure: rather than looking for perfectly regular, ordered fibers or fully disordered structures it may be more constructive to describe folded chromatin fibers in terms of interacting pairs of nucleosomes. While the interactions can be stable, maintaining a fixed topology, the structure will still have enough freedom of movement for relatively large shape changes, while locally allowing for transient defects. Such an open, dynamic, but well-organized structure may offer the eukaryotic cell the means to organize its genome in an efficient manner and offers many possibilities for further regulation.

## 4.4 Materials and Methods

### 4.4.1 Experimental force-extension data

All experimental procedures and the model that was fitted to the curves were described in detail elsewhere [23].

### 4.4.2 Rigid base pair Monte Carlo simulations of DNA

We extended the HelixMC package [38] for MC simulations of nucleosome (un-)wrapping and chromatin fiber folding. Random step parameters of non-nucleosomal DNA were drawn from a distribution as previously validated for similar buffer conditions as used in our experiments [38]. Base pair replacement was evaluated using a standard Metropolis scheme, which included a work term for the DNA tether. The force was linearly increased from 0.1 to 10 pN in typically $10^5$ steps. Starting configurations of straight DNA interspersed by DNA folded into nucleosomes (*see* below) were equilibrated in 1000 iterations of all base pairs prior to ramping the force. For generation of force-extension plots, 250 logarithmically distributed points were stored during the simulations, without averaging. To further mimic the experimental conditions, we added a hard-wall potential for both the surface of the flow cell and the bead when

the $z$-coordinate of the base pair was smaller than 0, or larger than that of the last base pair. Hard-wall potentials were set to $10^6$ $k_BT$. This simulation scheme resulted in force-extension curves of DNA that closely matched a WLC with a persistence length of 50 nm and without hysteresis, indicating full equilibrium. Though the resulting thermal fluctuations bear close resemblance with experimental curves, part of the fluctuations in the experimental data are filtered out due to the large viscous drag of the bead at small forces.

### 4.4.3 Nucleosome (un-)wrapping in Monte Carlo simulations

The wrapping of DNA into nucleosomes was implemented using a second layer of step parameters, corresponding to the 14 DNA histone contact points in the nucleosome, as depicted in Figures 4.2b and c. From the 1kx5 pdb file, the average B-factor was computed for all base pairs. 14 fixed base pairs were picked from 14 local minima of the B-factor. Note that this reduces the number of nucleosome-wrapped base pairs to 132 out of 147. Using the frames of these fixed base pairs, 14 step parameters were computed relative to the frame of the dyad base pair. The stiffness matrix for the wrapping parameters has not been described before. We assumed a standard deviation of the step parameters of 1 Å and 0.1 radians to define the diagonal of the stiffness matrix. Outcomes did not change notably for slightly different values. Unwrapping energies were clipped to 2.5 $k_BT$ to match known histone-DNA contact energies [42], yielding nucleosomes that can unwrap from the entry and exit points of the nucleosome. By defining the histone-DNA contact step parameters relative to the dyad base pair frame, we allowed for full unwrapping that can occur in an asymmetric fashion. The boundary of nucleosomal DNA shifts into the nucleosome in forward iterations. Once inside the nucleosomal DNA, step parameters were kept unchanged, up to the first free base pair on the other side of the nucleosome. If the first free base pair was within the range of nucleosomal base pairs, new step parameters were drawn from the corresponding base pair in the crystal structure, rather than from a random pool, and were accepted following the standard Metropolis criterion. Thus, rewrapping of nucleosomal DNA was implemented at the end of the nucleosome. To impose symmetry in the unwrapping mechanism, the iteration direction of the MC computation was reversed in odd and even iterations. The wrapping energy term was complemented by a hard-wall excluded volume term for all nucleosome pairs with center-to-center distances smaller than 5 nm.

### 4.4.4 Nucleosome stacking in Monte Carlo simulations

Nucleosome stacking was parametrized using a third layer of step parameters that describe the reference frames of the nucleosomes, as shown in Figure 4.2a. Similar to Korolev *et al.* [43]. However, we only used DNA coordinates to

obtain the nucleosome reference frame. We define this frame with its origin at the center of mass of the nucleosome, the $x$-axis pointing to the dyad and the $z$-axis pointing along the direction of the nucleosomal DNA superhelix and the $y$-axis perpendicular to the $x$- and $z$-axis. Nucleosome step parameters were defined by positioning the nucleosomes in a left-handed 1-start or 2-start helix, shown in Figures 4.2b and c. The dimensions of the chromatin fiber were chosen to roughly the match geometries obtained from condensed fibers using EM: an outer diameter of 33 nm, a nucleosome line density of 2 nm and 7 nucleosomes per superhelical gyre [20]. Using nucleosomes cast in this manner, nucleosome step parameters were calculated from the nucleosome reference frames of either neighboring nucleosomes, in the case of a 1-start fiber, or next-neighboring nucleosomes for 2-start fibers. Note that this imposes identical step parameters for nucleosomes stacked in both geometries, reflecting the local similarity in the structure of stacked nucleosome pairs.

The stiffness matrix for nucleosome stacking is unknown. For our simulations, we assumed a standard deviation due to thermal fluctuations of 10 Å and 1 radian to define the diagonals of the shift and rotation parameters in the stiffness matrix. This relatively low stiffness, reflecting the high flexibility of the H4 histone tails that have been indicated to physically mediate nucleosome stacking [44], allows for significant rearrangements of the nucleosomes in the fiber, while maintaining a defined global topology. We introduced a moderate nonlinearity of the stiffness by adding an extra term of 0.01 times the square of the stacking energy, resulting in a reduced freedom for movements at large deviations from the preferred position. By clipping the maximal energy for stacking fluctuations to 25 $k_{\mathrm{B}}T$, unstacking of nucleosome pairs was allowed for sufficiently large forces.

## 4.5 Acknowledgements

We thank Helmut Schiessel and Victor Zhurkin for valuable discussions and Chi Pham for DNA preparation. This work is part of the research program VICI with project number 680-47-616, which is (partly) financed by the Netherlands Organisation for Scientific Research (NWO).

## 4.6 Supplementary Material

### 4.6.1 Reconstitution of chromatin fibers

XL-1-blue E.coli competent cells (Agilent Technologies) were transfected with Ampicillin-resistant pUC18 plasmids containing tandem repeats of the Widom-601 sequence with Nucleosome Repeat Lengths (NRL) of 167 or 197 base pairs. After harvesting and purification using a NucleoBond® Xtra Midi kit (Macherey-Nagel), the plasmids were digested with BseYI and BsaI (New England Biolabs) and purified with a Promega Wizard SV Gel & PCR cleanup kit (Promega). Digoxigenin–11–dUTP (Roche) and Biotin–16–dUTP (Roche) were sequentially built into the ends of the DNA array using LC Klenow Fragment (Thermo Fisher Scientific) and purified with a Promega Wizard SV Gel & PCR cleanup kit (Promega). Chromatin fibers were reconstituted overnight with recombinant human histone octamer (EpiCypher) by salt dialysis in Slide-A-Lyzer 10,000 MWCO MINI dialysis tubes (Thermo Scientific) in a titration series of 601-DNA : octamer ratios between 0.8 and 1.5, using a pump to reduce the salt concentration from 5 M NaCl, Tris EDTA (TE) down to TE. Reconstitution quality was checked by native agarose gel electrophoresis.

### 4.6.2 Flow cell preparation

Flow cells were assembled from 2 cover slips, 24 × 40 mm and 24 × 60 mm, #1.5 thickness (Menzel Gläser), sandwiched around a Polydimethylsiloxane (PDMS, Sylgard) mold, which was washed thoroughly with milli-Q water and 2-propanol. The bottom slide was coated with 0.1% nitrocellulose in amyl-acetate (Ladd Research Industries) prior to incubation with 300 µl of 10 ng/µl sheep anti-digoxigenin (Sigma-Aldrich). Next, the flow cell was passivated overnight with 300 µl 4% (w/v) Bovine Serum Albumin (BSA, Sigma-Aldrich) and 2% (w/v) TWEEN 20 (Sigma-Aldrich) in milli-Q water at 4°C. Experiments were done in measurement buffer (MB): 100 mM KCl, 2 mM $MgCl_2$, 10 mM $NaN_3$, 10 mM HEPES pH 7.5 and 0.4% (w/v) BSA. Flow cells were washed with MB prior to measurements. 500 µl of 40 pg/µl chromatin in MB was incubated for 10 min at room temperature. Next, 20 µl of M270 magnetic beads (Invitrogen) were washed three times in TE buffer and diluted 500 times in MB before they were flushed into the flow cell. After 10 min the flow cell was washed again with MB and the tethered chromatin fibers were ready for force spectroscopy.

### 4.6.3 Force spectroscopy on single chromatin fibers using magnetic tweezers

A home-built magnetic tweezers apparatus (MT), consisting of a 25 Mpix Condor camera (CMOS Vision GmbH), a Camera Link PCIe-1433 frame grabber (National Instruments), a 40×, oil, NA = 1.3, Plan Fluor objective (Nikon

Corporation), an infinity-corrected tube lens (ITL200, Thorlabs), a 100 µW 645 nm LED (IMM Photonics GmbH), a multi-axis piezo scanner P-517.3CL (Physik Instrumente), two 5 mm cube magnets N50 (Supermagnete, Webcraft), a hollow shaft Stepper Motor (Casun) and two M-126.2S1 translation stages (Physik Instrumente). By changing the magnet position between 10 and approximately 0.3 mm above the flow cell bottom at a velocity of $\sim 0.5$ mm/s, the force was increased and decreased exponentially between $\sim 0.05$ and 70 pN. Typically $\sim 100$ beads were tracked in real-time with an accuracy better than 5 nm using a custom tracking algorithm based on Fast Fourier Transform cross-correlation of a 100 pixels region of interest around a bead with a computer-generated image featuring cylindrically symmetric pattern of interference signals. Computer-controlled magnet movements and synchronous real-time bead tracking were implemented in LabVIEW (National Instruments).

### 4.6.4 Experimental data analysis

Force-extension traces were selected to discard stuck beads, double tethered beads and other anomalous traces. Drift and offset were subtracted from the curves, such that the extension at forces between 30 and 70 pN followed an extensible Worm Like Chain with a persistence length of 50 nm, a contour length equal to the length of the used DNA substrate and a stretch modulus of 900 pN. Force-extension data of the fiber at forces below 30 pN were fit to a statistical mechanics model, following Meng *et al.* [23]. The force dependent extension of the chromatin tether was described by the sum of the extension of bare DNA handles, a fraction of stacked nucleosomes, a fraction of partially unwrapped nucleosomes, a fraction of singly-wrapped nucleosomes and a fraction of fully unwrapped nucleosomes. The extension of the stacked fraction was described by a Hookean spring with a rest length of 12 Å per nucleosome for 197 NRL fibers and 6 Å for 167 NRL fibers. The extension of the other nucleosomal conformations followed that of the DNA, in which the contour length equaled the NRL for the fully unwrapped nucleosome, NRL minus 56 base pairs for the patially wrapped nucleosome and 5 nm less for the singly-wrapped nucleosome.

The transition between singly wrapped nucleosomes and fully unwrapped nucleosomes was not reversible at the time scale of the experiment as indicated by the experimentally observed hysteresis (data not shown). The unfolding could nevertheless be reversed after $\sim 5$ s delay at low force between successive pulling curves. These transitions resulted in a characteristic 25 nm steps in extension which directly yielded the number of nucleosomes and tetrasomes in the fiber. All transitions below 10 pN featured no hysteresis and were fitted to an equilibrium model. The free energy of each of the conformations was obtained from integrating the extension of each conformation over the force and was supplemented by a stacking energy $\Delta G_1$, a partial unwrapping energy $\Delta G_2$ and a work term. The free energy contributions of each of the nucleosome

conformations, the number of each of the nucleosome conformations and a possible degeneracy term, which describes whether nucleosome pairs unstack independently or cooperatively, were used to calculate the Boltzmann weighted average of the extension of the fiber as a function of force.

Fitting this model, using the above described standard mechanical parameters of DNA, yielded for every force-extension curve the number of nucleosomes, the number of tetrasomes, the stiffness of the fiber per nucleosome, rupture energy $\Delta G_1$, and interaction energy $\Delta G_2$. The amount of unwrapping $L_{\text{unwrapp}} = 56$ base pairs and a degeneracy factor $D = 0$ for 197 NRL fibers and $D = 1$ for 167 NRL fibers were fixed in all fits and corresponded to previously fitted values [23]. All analysis of experimental data was implemented in LabVIEW (National Instruments) and this software is available upon request.

### 4.6.5 Base pair step parameters

For MC simulations we made use of several modules from the HelixMC package (`http://helixmc.readthedocs.io/index.html`) described in [38]. Base pair step parameters were draw from `helixmc.random_step.RandomStepBase()` using HelixMC file `\data\DNA_gau.npy`. The sequence averaged parameters, as provided in this module, were:

| | **Shift** (Å) | **Slide** (Å) | **Rise** (Å) | **Tilt** (°) | **Roll** (°) | **Twist** (°) |
|---|---|---|---|---|---|---|
| **mean** | 0.00 | -0.32 | 3.30 | -0.05 | 1.60 | 35.21 |
| **s.d.** | 0.57 | 0.86 | 0.23 | 3.56 | 5.17 | 6.24 |

**Table 4.S1**
***Mean and standard deviation of the DNA base step parameters used in HelixMC.***

The base pair step parameters and the use of them in MC simulations to reconstruct experimental force-extension data of bare DNA at room temperature and physiological conditions were validated previously [38].

### 4.6.6 Nucleosome wrapping step parameters

Nucleosome base pair step parameters were extracted from the 1kx4 crystal structure [35] by uploading the pdb file to the web 3DNA website (`http://w3dna.rutgers.edu/`). From the 14 local minima of the average B-factor in the pdb-file we extracted 14 fixed base pairs: 7, 17, 28, 38, 48, 59, 70, 80, 90, 101, 111, 121, 132, 142. The dyad was defined as the mean base pair index of these 14 fixed locations: base pair 75. A starting conformation of the chromatin fiber was generated by replacing the base pair step parameters of a zero-temperature DNA conformation, by the those of the nucleosome. The dyads of each of the nucleosomes defined the NRL of the fiber.

To maintain DNA wrapping around the histone core, we defined step parameters for DNA wrapping relative to the reference frame of the dyad base pair in a similar fashion as for the DNA itself. This yielded the following step parameters for the 14 fixed base pairs relative to the dyad frame:

| **base pair** | **Shift** (Å) | **Slide** (Å) | **Rise** (Å) | **Tilt** (°) | **Roll** (°) | **Twist** (°) |
|---|---|---|---|---|---|---|
| 7 | 18.8 | -13.8 | 33.7 | 37 | -50 | 158 |
| 17 | 17.7 | -12.5 | 58.1 | 60 | -79 | 143 |
| 28 | 2.2 | -10 | 79.1 | -77 | 122 | -161 |
| 38 | 26.5 | -43.9 | -63.8 | -101 | 129 | 93 |
| 48 | 12.8 | -9.8 | -70.3 | -86 | 102 | 138 |
| 59 | 4 | -7.7 | -49.5 | -35 | 61 | 157 |
| 70 | -0.5 | 2.7 | -16 | 3 | -28 | -164 |
| 80 | 1.2 | -4.4 | 16.4 | -11 | 9 | -170 |
| 90 | 2.2 | -10 | 46.5 | -30 | 56 | -169 |
| 101 | 2 | -10.7 | 68.2 | -31 | 112 | -163 |
| 111 | 9.8 | -24 | 78.2 | -48 | 150 | -149 |
| 121 | 7.8 | -21.2 | -76 | -59 | 139 | 129 |
| 132 | 2.3 | -26.9 | -58.3 | 41 | -108 | -157 |
| 142 | 2 | 18.2 | -39.2 | -27 | 71 | 174 |
| **s.d.** | **1** | **1** | **1** | **5** | **5** | **5** |

**Table 4.S2**
***Mean and standard deviation of the used DNA wrapping step parameters for each fixed base pair.*** *The mean values were calculated from the 1kx4 crystal structure of the nucleosome. The flexibility, represented by the standard deviations, was guessed.*

The unwrapping energy $E_{\text{unwrap}}$ was calculated based on a standard deviation of 1 Å for shift, slide and rise and 5 degrees for tilt, roll and twist:

$$E_{\text{unwrap}} = \sum_j \min\left(\sum_i \frac{1}{2}\frac{k_{\text{B}}T}{\sigma_{x_i}^2}(x_i - \bar{x_l})^2, E_{\text{unwrap,max}}\right), \tag{4.1}$$

with $x_i$, $\bar{x_l}$, $\sigma_{x_i}^2$ the actual, mean, and standard deviation of step parameter $i$ for fixed base pair $j$ and thermal energy $k_{\text{B}}T$. The unwrapping energy was clipped for each fixed base pair at $E_{\text{unwrap,max}}$, which was typically set to 2.5 $k_{\text{B}}T$. Because of this clipping, the results were not sensitive to the precise values of the standard deviations of the wrapping step parameters, as long as they were sufficiently small. Note that the interaction potential is harmonic in all 6 degrees of freedom and that we did not include correlations between degrees of freedom in this simple model.

### 4.6.7 Nucleosome stacking step parameters

Nucleosome stacking was implemented in a similar fashion as base pairs in the rigid base pair model. From the global dimensions of chromatin fibers [20], *i.e.* a diameter of 330 Å and a nucleosome line density of 20 Å and assuming a left-handed fiber we calculated the following step parameters for nucleosome stacking:

| | **Shift** (Å) | **Slide** (Å) | **Rise** (Å) | **Tilt** (°) | **Roll** (°) | **Twist** (°) |
|---|---|---|---|---|---|---|
| **mean** | 44.1 | -4.3 | 85.4 | 5 | 55 | 11 |
| **s.d.** | 1 | 1 | 1 | 5 | 5 | 5 |

**Table 4.S3**
***Mean and standard deviation of the used nucleosome stacking step parameters.*** *The mean values were calculated based on the dimensions of folded chromatin fibers and a left-handed chirality. The flexibility, represented by the standard deviations, was guessed.*

We assumed a similar rigidity of the stacking interaction as for wrapping, yielding a similar energy term for each stacked nucleosome:

$$E_{\text{unstack}} = \min\left(\sum_i \frac{1}{2}\frac{k_{\text{B}}T}{\sigma_{x_i}^2}(x_i - \bar{x_l})^2, E_{\text{unstack,max}}\right). \tag{4.2}$$

Simulations of non-interacting nucleosomes were done with $E_{\text{unstack,max}} = 0$. For 1-start and 2-start fibers we evaluated the nucleosome stacking parameters only for neighbor or next-neighbor nucleosomes respectively. Note again that the interaction potential is harmonic in all 6 degrees of freedom and that we did not include correlations between degrees of freedom in this simple model.

### 4.6.8 Excluded volume effects

Next to attractive interactions between histones and DNA and between nucleosomes we included hard-wall repulsive interactions $E_{\text{rep}} = 10^6\ k_{\text{B}}T$ to mimic the experimental conditions:

$$E_{\text{surface/bead}} = E_{\text{rep}} \tag{4.3}$$

if any of the base pair $z$-coordinate is smaller than 0 or larger than the $z$-coordinate of the last base pair $z_{\text{max}}$. Overlap between nucleosomes was also penalized:

$$E_{\text{nuc.overlap}} = E_{\text{rep}} \tag{4.4}$$

if the center of mass of any pair of nucleosomes was smaller than 55 Å.

### 4.6.9 MC initialization

A chromatin fiber was initialized by creation of a DNA pose, using `helixmc.pose.HelixPose()`.The start conformation, having straight (linker) DNA, represents the zero-temperature structure in absence of nucleosome-nucleosome interactions. Folded fibers were created by setting the stacking energy to infinite and running 1000 MC iterations, as described below. Nucleosome stacking was achieved rapidly, typically within 100 iterations. The initial 1000 conformations were discarded.

### 4.6.10 MC iterations

After initialization, the stacking energy was clipped to $E_{\text{unstack,max}}$. Iterating through every base pair, starting at the first, the step parameters were tentatively replaced by random parameters using `helixmc.random_step.RandomStepSimple()`, drawing from the HelixMC file `\data\DNA_gau.npy`. The change was evaluated and accepted using a standard Metropolis criterion:

$$P(\text{accept}) = \begin{cases} 1 & \Delta E \leq 0 \\ e^{-\Delta E / k_{\text{B}} T} & \Delta E > 0 \end{cases} \tag{4.5}$$

with

$$\Delta E = \Delta E_{\text{unwrap}} + \Delta E_{\text{unstack}} + \Delta E_{\text{surface/bead}} + \Delta E_{\text{nuc.overlap}} - F z_{\text{max}} \tag{4.6}$$

The latter term represents the work and depends on force f and the $z$-coordinate of the last base pair $z_{\text{max}}$. Rewrapping of the nucleosomal DNA was achieved by replacing the step parameters of the first free nucleosomal base pair by the corresponding step parameter of the nucleosome crystal structure, rather than from a random draw, and was evaluated using Equation 4.5 and 4.6. In each iteration the direction of base pair replacements was inverted, resulting in symmetric unwrapping/rewrapping of the nucleosomal DNA.

### 4.6.11 Simulation settings

All force ramps were performed using 50,000 steps in which the force was ramped up and down linearly, up to a preset maximum force. For force-extension curves and 3D animations of the structures, 250 snapshots were sampled in a logarithmic fashion, resulting in an even distribution of the points along the force-extension curve. Typically, a single run of an 1800 base pairs DNA containing 8 nucleosome fiber took 2-3 days on an Intel® Xeon® CPU E5-26660 v4 @ 2.00GHz PC running Windows 10. Source code is available from Github: `https://github.com/JvN2/ChromatinMC`.

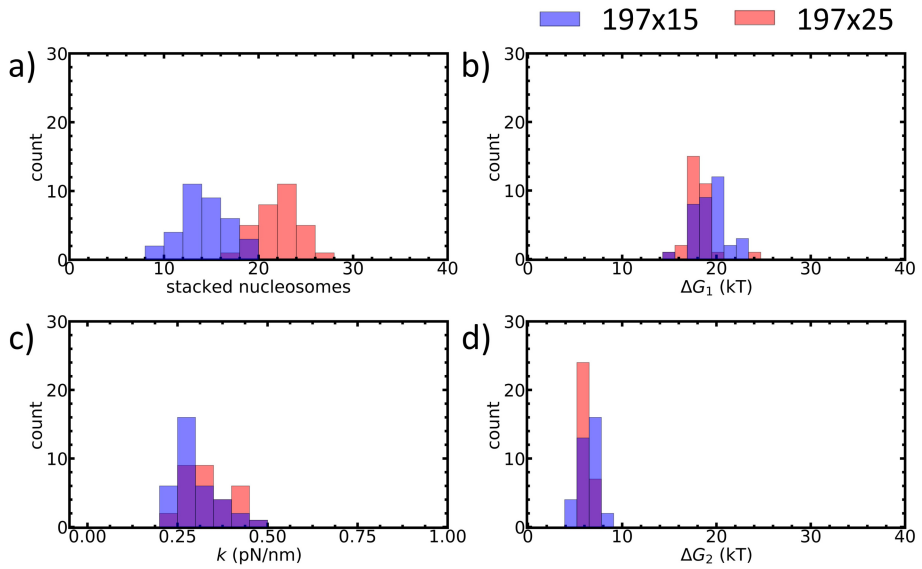


**Figure 4.S1**
**Experimental force-extension curves can be captured in 3 parameters that are independent of the number of nucleosomes.** Force-extension curves of more than 30 different fibers having either 15 (blue) or 25 (red) 197 NRL repeats of the 601 nucleosome positioning sequence were fitted to the statistical physics model described above with 4 free fitting parameters: a) the number of nucleosomes, b) the stacking energy, c) the fiber stiffness and d) the unwrapping energy up to the singly wrapped nucleosome. Both types of fibers were, on average, slightly under-saturated, but feature a narrow and overlapping distribution in stacking energy, stiffness and residual unwrapping energy, showing that the model parameters fully describe fiber unfolding at forces up to 8 pN.

### 4.6.12 Supplementary experiments

Supplemental Figure 4.S1 shows that fits of individual force-extension curves to the statistical physics model described above yield reproducible parameters that are shared between different NRLs, despite some variation in composition, Like in the experimental data, we analyzed individual force-extension curves separately and refrained from averaging in time or between simulations. All simulations were, however, executed in triplo and featured similar results. Supplemental Figure 4.S2 shows the autocorrelation times of the extension and the unwrapping energy of a single time trace at constant force for three different chromatin structures. In all cases, correlation is lost after 200 steps, and for the extension 30 steps were sufficient, showing that simulation results were in equilibrium and detailed balance was achieved.

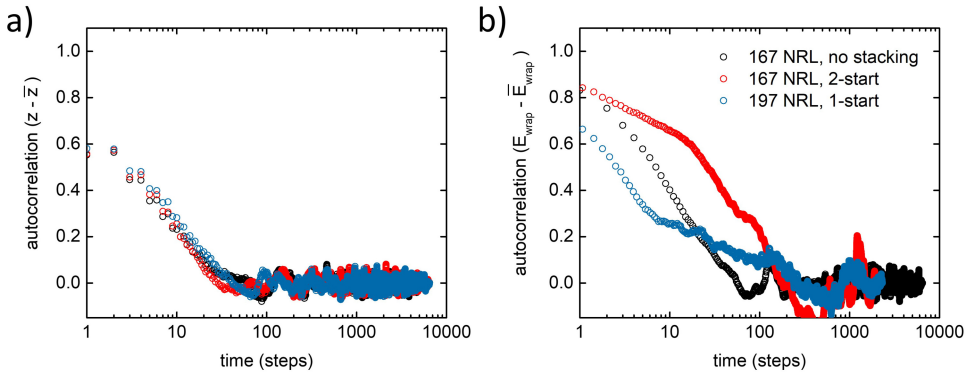


**Figure 4.S2**

**Normalized autocorrelation times for fiber folding.** a) The correlation of the extension at f = 0 pN of three typical chromatin fibers is lost after 50 steps in the MC simulations on 4000 base pairs chromatin fibers containing 15 nucleosomes. The extensions were $z = 0.29 \pm 0.09$ µm, $0.28 \pm 0.09$ µm and $0.26 \pm 0.09$ µm (mean ± s.d.) for 167 NRL without stacking (black), 167 NRL in a 2-start structure (red) and 197 NRL in a 1-start structure (blue). b) The corresponding autocorrelation of the unwrapping energy shows that DNA unwrapping is more persistent, but correlations disappear after 70 (no stacking) to 200 steps. The unwrapping energies were $E_{\mathrm{unwrap}} = 0.7 \pm 0.5\ k_{\mathrm{B}}T$, $0.6 \pm 0.3\ k_{\mathrm{B}}T$ and $0.3 \pm 0.3\ k_{\mathrm{B}}T$ (mean ± s.d.) for 167 NRL without stacking (black), 167 NRL in a 2-start structure (red) and 197 NRL in a 1-start structure (blue).

# Bibliography IV


[1] Jeffrey C Hansen et al. "The 10-nm chromatin fiber and its relationship to interphase chromosome organization". In: *Biochemical Society Transactions* 46.1 (2018), pp. 67–76.

[2] Margot P Scheffer, Mikhail Eltsov, and Achilleas S Frangakis. "Evidence for short-range helical order in the 30-nm chromatin fibers of erythrocyte nuclei". In: *Proceedings of the National Academy of Sciences* 108.41 (2011), pp. 16992–16997.

[3] Andrew Travers. "The 30-nm fiber redux". In: *Science (New York, NY)* 344.6182 (2014), pp. 370–372. ISSN: 0036-8075. DOI: 10.1126/science.1253852.

[4] Eden Fussner, Reagan W Ching, and David P Bazett-Jones. "Living without 30 nm chromatin fibers". In: *Trends in biochemical sciences* 36.1 (2011), pp. 1–6.

[5] Chenyi Wu, John E. McGeehan, and Andrew Travers. "A metastable structure for the compact 30-nm chromatin fibre". In: *FEBS Letters* 590.7 (2016), pp. 935–942. ISSN: 18733468. DOI: 10.1002/1873-3468.12128.

[6] Guohong Li and Danny Reinberg. "Chromatin higher-order structures and gene regulation". In: *Current opinion in genetics & development* 21.2 (2011), pp. 175–186.

[7] Robert K. McGinty and Song Tan. "Nucleosome structure and function". In: *Chemical Reviews* 115.6 (2015), pp. 2255–2273. ISSN: 15206890. DOI: 10.1021/cr500373h.

[8] Swaminathan Venkatesh and Jerry L Workman. "Histone exchange, chromatin structure and the regulation of transcription." In: *Nature Reviews. Molecular Cell Biology* 16.3 (2015), pp. 178–189. ISSN: 1471-0080. DOI: 10.1038/nrm3941.

[9] Aaron D Goldberg, C David Allis, and Emily Bernstein. "Epigenetics: a landscape takes shape". In: *Cell* 128.4 (2007), pp. 635–638.

[10] Heinz Neumann et al. "A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation". In: *Molecular cell* 36.1 (2009), pp. 153–163.

[11] Abdollah Allahverdi et al. "The effects of histone H4 tail acetylations on cation-induced chromatin folding and self-association". In: *Nucleic acids research* 39.5 (2010), pp. 1680–1691.

[12] Ying Liu et al. "Influence of histone tails and H4 tail acetylations on nucleosome–nucleosome interactions". In: *Journal of molecular biology* 414.5 (2011), pp. 749–764.

[13] Qinming Chen et al. "Regulation of nucleosome stacking and chromatin compaction by the histone H4 N-terminal tail–H2A acidic patch interaction". In: *Journal of molecular biology* 429.13 (2017), pp. 2075–2092.

[14] Monika Lachner et al. "Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins". In: *Nature* 410.6824 (2001), p. 116.

[15] Thomas Schalch et al. "X-ray structure of a tetranucleosome and its implications for the chromatin fibre". In: *Nature* 436.7047 (2005), pp. 138–141. ISSN: 0028-0836. DOI: 10.1038/nature03686.

[16] F. Song et al. "Cryo-EM Study of the Chromatin Fiber Reveals a Double Helix Twisted by Tetranucleosomal Units". In: *Science* 344.6182 (2014), pp. 376–380. ISSN: 0036-8075. DOI: 10.1126/science.1251413.

[17] Mikhail Eltsov et al. "Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes in situ". In: *Proceedings of the National Academy of Sciences* 105.50 (2008), pp. 19732–19737.

[18] Maria Aurelia Ricci et al. "Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo". In: *Cell* 160.6 (2015), pp. 1145–1158.

[19] Kazuhiro Maeshima et al. "Chromatin as dynamic 10-nm fibers". In: *Chromosoma* 123.3 (2014), pp. 225–237.

[20] Philip J J Robinson et al. "EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure." In: *Proceedings of the National Academy of Sciences of the United States of America* 103.17 (2006), pp. 6506–11. ISSN: 0027-8424. DOI: 10.1073/pnas.0601212103.

[21] Maarten Kruithof et al. "Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30-nm chromatin fiber." In: *Nature structural & molecular biology* 16.5 (2009), pp. 534–40. ISSN: 1545-9985. DOI: 10.1038/nsmb.1590.

[22] Fan-Tso Chien and John van Noort. "10 Years of Tension on Chromatin: Results From Single Molecule Force Spectroscopy." In: *Current pharmaceutical biotechnology* 10.5 (2009), pp. 474–85. ISSN: 1873-4316. DOI: 10.2174/138920109788922128.

[23] He Meng, Kurt Andresen, and John van Noort. "Quantitative analysis of single-molecule force spectroscopy on folded chromatin fibers". In: *Nucleic Acids Research* 43.7 (Apr. 2015), pp. 3578–3590. ISSN: 13624962. DOI: 10.1093/nar/gkv215.

[24] Jean Marc Victor et al. "Pulling chromatin apart: Unstacking or Unwrapping?" In: *BMC biophysics* 5.1 (2012), p. 21.

[25] Maria Barbi et al. "On the topology of chromatin fibres". In: *Interface focus* 2.5 (2012), pp. 546–554.

[26] Annick Lesne and J-M Victor. "Chromatin fiber functional organization: some plausible models". In: *The European Physical Journal E* 19.3 (2006), pp. 279–290.

[27] Hua Wong, Jean-Marc Victor, and Julien Mozziconacci. "An all-atom model of the chromatin fiber containing linker histones reveals a versatile structure tuned by the nucleosomal repeat length". In: *PloS one* 2.9 (2007), e877.

[28] Rosana Collepardo-Guevara and Tamar Schlick. "Chromatin fiber polymorphism triggered by variations of DNA linker lengths". In: *Proceedings of the National Academy of Sciences* 111.22 (2014), pp. 8061–8066.

[29] Rosana Collepardo-Guevara and Tamar Schlick. "The effect of linker histone's nucleosome binding affinity on chromatin unfolding mechanisms". In: *Biophysical journal* 101.7 (2011), pp. 1670–1680.

[30] Tamar Schlick, Jeff Hayes, and Sergei Grigoryev. "Toward convergence of experimental studies and theoretical modeling of the chromatin fiber". In: *Journal of Biological Chemistry* 287.8 (2012), pp. 5183–5191.

[31] Helmut Schiessel. "The physics of chromatin". In: *Journal of Physics: Condensed Matter* 15.19 (2003), R699.

[32] Nick Kepper et al. "Force spectroscopy of chromatin fibers: extracting energetics and structural information from Monte Carlo simulations". In: *Biopolymers* 95.7 (2011), pp. 435–447.

[33] Artur Kaczmarczyk et al. "Single-molecule force spectroscopy on histone H4 tail-cross-linked chromatin reveals fiber folding". In: *Journal of Biological Chemistry* 292.42 (2017), pp. 17506–17513.

[34] M. Kruithof et al. "Subpiconewton dynamic force spectroscopy using magnetic tweezers." In: *Biophysical journal* 94.6 (2008), pp. 2343–2348. ISSN: 00063495. DOI: 10.1529/biophysj.107.121673.

[35] K Luger et al. "Crystal structure of the nucleosome core particle at 2.8 A resolution." In: *Nature* 389.6648 (1997), pp. 251–260. ISSN: 0028-0836. DOI: 10.1038/38444. arXiv: NIHMS150003.

[36] Wei Li et al. "FACT remodels the tetranucleosomal unit of chromatin fibers for gene transcription". In: *Molecular cell* 64.1 (2016), pp. 120–133.

[37] Sinan Kilic et al. "Single-molecule FRET reveals multiscale chromatin dynamics modulated by HP1$\alpha$". In: *Nature communications* 9.1 (2018), p. 235.

[38] Fang-Chieh Chou, Jan Lipfert, and Rhiju Das. "Blind predictions of DNA and RNA tweezers experiments with force and torque". In: *PLoS computational biology* 10.8 (2014), e1003756.

[39] Bernard D Coleman, Wilma K Olson, and David Swigon. "Theory of sequence-dependent DNA elasticity". In: *The Journal of chemical physics* 118.15 (2003), pp. 7127–7140.

[40] Igor M. Kulić et al. "Equation of state of looped DNA". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 75.1 (2007). ISSN: 15393755. DOI: 10.1103/PhysRevE.75.011913. arXiv: 0509003 [q-bio].

[41] Yoshinori Nishino et al. “Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure”. In: *The EMBO journal* 31.7 (2012), pp. 1644–1653.

[42] WJA Koopmans et al. “Single-pair FRET microscopy reveals mononucleosome dynamics”. In: *Journal of fluorescence* 17.6 (2007), pp. 785–795.

[43] Nikolay Korolev, Alexander P Lyubartsev, and Lars Nordenskiöld. “A systematic analysis of nucleosome core particle and nucleosome-nucleosome stacking structure”. In: *Scientific reports* 8.1 (2018), p. 1543.

[44] Masatoshi Wakamori et al. “Intra-and inter-nucleosomal interactions of the histone H4 tail revealed with a human nucleosome core particle with genetically-incorporated H4 tetra-acetylation”. In: *Scientific reports* 5 (2015), p. 17204.