The role of linker DNA in chromatin fibers

Brouwer, T.B.

Summary

The genetic information of all living organisms is contained in their DNA. Every human cell, for instance, contains approximately 2 meters of DNA. This massive length needs to be condensed thousands of times to physically fit in the cell nucleus. Simultaneously, the genetic information in the DNA must be selectively accessible to sustain cellular development or maintenance. Cells modify the degree of DNA compaction by epigenetics, which largely determines what genes are read out and which genes are transcriptionally silent.

Despite decades of research into this mechanism, there is no consensus on how cells realize the various degrees of DNA compaction *in vivo*. Eukaryotes, such as humans, fold DNA around histone octamers to form structures called nucleosomes. Subsequently, the nucleosomes interact to form higher-order structures such as compact chromatin fibers. Several structural models for these fibers have been described. The 1-start solenoid fiber and the 2-start zig-zag fiber are most popular. The debate on the prevalence of such organized structures *in vivo* is, however, ongoing since unequivocal structural data is lacking. In our group, we have shown before that the length of the linker DNA, the DNA between adjacent nucleosomes, controls the mechanical properties of chromatin fibers, hinting at more than one higher-order fiber structure. A deeper understanding of the chromatin fiber and its compaction mechanism is important because the dysfunction of such regulation results in various medical conditions. For instance, the epigenetic disorder type 1 diabetes, fragile X syndrome, or various cancers have been related to malfunctioning of DNA reorganization.

The study of the structure and dynamics of the proteins that organize chromatin has benefited tremendously from the development of single-molecule force spectroscopy techniques like magnetic tweezers, that allow for real-time, nanometer accuracy measurements of the compaction of DNA, and allow manipulation with pico-Newton scale forces. In addition to applying force to single DNA molecules, magnetic tweezers have the unique ability to exert twist on the DNA molecule. Twist plays an important role when DNA organizing proteins form or release wraps, loops and bends in DNA. In Chapter 2, we introduce our custom-built multiplexed magnetic tweezers apparatus, which can measure hundreds of molecules simultaneously. We further describe the

experimental procedure to prepare DNA substrates, assemble flow cells, tether DNA to a magnetic bead inside these flow cells, and manipulate and record the extension of such DNA tethers with magnetic tweezers. Last, we explain how mechanical parameters of nucleo-protein filaments can be extracted by fitting a statistical mechanics model to the data. Such multiplexed experiments yield a comprehensive description of the mechanical properties of chromatin fibers, which informs us of their higher-order structure.

Many single-molecule force spectroscopy techniques rely on nanometric tracking of micrometer-sized beads to obtain quantitative information about the mechanical properties of biomolecules. The three-dimensional position of the microbeads can be resolved by holographic analysis of the diffraction pattern in wide-field imaging. However, most popular bead-tracking methods are computationally intensive and not optimized for multiplexed experiments. In Chapter 3, we introduce 3D phasor tracking, a fast and robust bead tracking algorithm with nanometric localization accuracy in a z-range of over 10 µm. The algorithm is based on a 2D cross-correlation between the diffraction pattern and several computer-generated reference images using Fast Fourier Transforms, and yields a processing rate of up to 10,000 regions of interest per second. We implemented the technique in our magnetic tweezers setup, introduced in Chapter 2, and were able to track the 3D position of over 100 beads in real-time on a generic CPU. 3D phasor tracking can improve biophysical bead tracking applications, especially where high throughput is required. The development of this lightweight tracking method was instrumental for the rest of the work described in this thesis.

In Chapter 4 we report on our studies of the mechanical properties of chromatin fibers with Monte Carlo simulations. 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. With our simulations, we reproduced different stages of chromatin fiber unfolding as a function of force. The results are largely consistent with a 2-start folding for short linker DNA and a 1-start folding for long linker DNA. Because the simulations provide the three-dimensional structures of chromatin fibers with base pair accuracy, while maintaining realistic physical properties of linker DNA, they facilitate the interpretation of force-extension data obtained by force spectroscopy. Furthermore, they relate the parameters that are extracted by fitting the statistical physics model in a more detailed, structural way.

With the multiplexed magnetic tweezers setup (introduced in Chapter 2) equipped with the newly developed tracking algorithm (introduced in Chapter 3), we performed a thorough investigation of the unfolding pathway of regular chromatin fibers in Chapter 5. We quantified the effect of increasing the linker length on the mechanical properties of chromatin fibers with single base pair steps and complemented these experiments with rigid base pair Monte Carlo simulations (introduced in Chapter 4). Both the computational and

experimental results reveal a periodic variation of the folding energy with linker length, due to the limited flexibility of the linker DNA. The most stable stacking interactions were observed for linker lengths of multiples of 10 base pairs, which corresponds to the helical periodicity of B-DNA. We analyzed nucleosome stacking in both 1- and 2-start topologies and show that the length of the linker DNA determines the preferred stacking topology. Short linkers facilitate the formation of a 2-start fiber, whereas long linkers drive the chromatin fibers in a 1-start conformation. Moreover, we present evidence that the sequence of the linker DNA also modulates nucleosome stacking. The prevalence and balance between different fiber structures that depend on both linker length and linker sequence suggests a pathway for epigenetic regulation in the cell.

Where Chapters 3-5 detail the folding of eukaryotic chromatin, and how the physical organization of the genome plays a key role in gene regulation, the last chapter of this thesis is dedicated to the genomic organization in Archaea, our evolutionary predecessor. Recent X-ray crystallography studies of *Methanothermus fervidus* histone HMfB bound to 80 base pair DNA fragments suggested the formation of hypernucleosomes consisting of DNA wrapped around an 'endless' histone-protein core. However, if and how such a hypernucleosome structure assembles on a long DNA substrate and which interactions provide for its stability, remained unclear. Electrostatic interactions may drive ionic modulation of such a hypernucleosome structure which has been proposed to regulate gene expression. In Chapter 6, we describe micromanipulation studies of complexes of the *M. fervidus* histones HMfA and HMfB with single DNA tethers. The experimental procedure introduced in Chapter 2 could also be applied to archaeal chromatin fibers, though we needed to adjust the statistical mechanics model that captures the events that we observed in the experimental data. Hypernucleosome assembly resulted from the cooperative binding of archaeal histones to DNA and was facilitated by stacking interactions between neighboring histone dimers. Furthermore, rotational force spectroscopy demonstrated that the HMfB-DNA complex has a left-handed chirality, but torque can drive it in a right-handed conformation. The structure of the hypernucleosome thus depends on stacking interactions, torque, and force. This suggests an important role for archaeal histones as a transcription regulator, responsive to environmental changes. Understanding these ancient cellular processes may provide additional insight into the more complicated eukaryotic DNA compaction mechanisms.

The work presented in this thesis advances the fundamental knowledge of chromatin compaction. In particular, the force spectroscopy experiments on chromatin fibers in Chapter 5 provide compelling evidence that more than one higher-order structure exists *in vitro*. The linker length-dependent preference for different topologies which we found in both experimental and simulated force-extension curves could explain the more disordered structure of chromatin *in vivo*, where linker length and sequence vary. In the future, these experiments

should be expanded to investigate at what length the switch between 2-start and 1-start fibers takes place and what factors influence this balance. There is, however, still no experimental evidence for regular higher-order structures *in vivo*. Since linker lengths are quantized in the cell and long linkers are prevalent (Widom – PNAS, 1992), we predict that small patches of these structured fibers will be identified soon. Importantly, the detailed dependence of the mechanical properties of the fiber with linker length, as presented here, further substantiates the very existence of 1-start structures.

Another noteworthy finding in Chapter 5 is that nucleosome stacking in 1-start fibers appeared not to be mediated by the histone H4 tail. This is in apparent contradiction to earlier experiments from our group (Kaczmarczyk *et al.* – JBC, 2017) which showed that chromatin fibers with cross-linked H4-tails fixed both the mechanical properties and higher-order structure these fibers. It may hint at more variations in the higher-order structure of chromatin and ways to regulate it. In addition, we found that the sequence of the linker DNA affects the stability of the chromatin fiber to a surprisingly large extent. Both histone-tail mediated nucleosome stacking and linker DNA sequence effects need to be further evaluated by means of simulation and experiment. Our Monte Carlo model, with some adjustments, can be used to compare the various stacking mechanisms within chromatin fibers and can be further expanded to include sequence effects.

Ultimately, we would like to characterize chromatin that includes the same composition and nucleosomal occupancies as found in the cell. Our parallelized approach will facilitate the investigation of these heterogeneous samples. Such experiments may further elucidate the mechanism of DNA organization *in vivo*. In our group, we have already taken the first steps in this direction. We have probed nucleosome stacking characteristics of native chromatin extracted from yeast by means of a custom-developed purification technique to isolate chromosomal loci in their native state (Hermans *et al.* – Scientific Reports, 2017). This work opens up a large new area of investigation into epigenetic regulation by the cell. In conclusion, we hope that the research described in this thesis constitutes the first step of further research into epigenetic conditions, perhaps outside of the biophysical discipline alone, which in turn may provide the cornerstone on which improved treatments of such conditions can be founded.