

Universiteit
Leiden
The Netherlands

# The role of linker DNA in chromatin fibers

Brouwer, T.B.

**Citation**
Brouwer, T. B. (2020, November 4). *The role of linker DNA in chromatin fibers*. *Casimir PhD Series*. Retrieved from https://hdl.handle.net/1887/138082

| | |
|---|---|
| Version: | Publisher's Version |
| License: | Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden |
| Downloaded from: | https://hdl.handle.net/1887/138082 |

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden




The handle http://hdl.handle.net/1887/138082 holds various files of this Leiden University dissertation.

**Author**: Brouwer, T.B.
**Title**: The role of linker DNA in chromatin fibers
**Issue Date**: 2020-11-04

# The Role of Linker DNA
## in Chromatin Fibers

# The Role of Linker DNA
## in Chromatin Fibers

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C. J. J. M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 4 november 2020
klokke 11:15 uur

door


Thomas Bauke Brouwer

geboren te Amsterdam, Nederland
in 1987

| | |
|---|---|
| Promotores: | Prof. dr. ir. S. J. T. van Noort<br>Prof. dr. T. Schmidt |
| Promotiecommissie: | dr. F. Müller-Planitz (Biomedizinisches Zentrum, München, Duitsland)<br>Prof. dr. ir. G. J. L. Wuite (Vrije Universiteit, Amsterdam)<br>Prof. dr. N. H. Dekker (Technische Universiteit, Delft)<br>dr. H. van Ingen (Universiteit Utrecht, Utrecht)<br>Prof. dr. H. Schiessel<br>Prof. dr. E. R. Eliel |

© 2020 Thomas B. Brouwer. All rights reserved.

Cover: *The Unknown Pleasures of Stretching Chromatin.* The cover depicts an artistic rendering of force-spectroscopy data, extension in time, measured while stretching many unique chromatin fibers. The curves were median-filtered, randomly amplified in extension, occasionally reversed in time, and some artificial noise was added to emulate the iconic cover of *Unknown Pleasures*, the debut record of *Joy Division*. Cover design: Valérie Nusman and Thomas B. Brouwer.

Casimir PhD Series, Delft-Leiden, 2020-26
ISBN 978-90-8593-452-3
An electronic version of this thesis can be found at
`https://openaccess.leidenuniv.nl`

This work was supported by the Netherlands Organisation for Scientific Research (NWO/OCW), as part of the Frontiers of Nanoscience program, and by the NWO-VICI Research Program Project 680-47-616.

*To Sandra (and the little one)*

"I can appreciate the beauty of a flower. At the same time, I see much more about the flower than (the artist) sees. I could imagine the cells in there, the complicated actions inside, which also have a beauty. It only adds."
*Richard P. Feynman*

# Contents


**1 Introduction** **1**
- 1.1 The Tree of Life . . . 3
- 1.2 DNA . . . 5
- 1.3 Genetics and epigenetics . . . 9
- 1.4 Chromatin . . . 11
  - 1.4.1 Nucleosomes . . . 12
  - 1.4.2 Chromatin fibers . . . 14
- 1.5 Force spectroscopy experiments . . . 18
- 1.6 Particle tracking . . . 19
  - 1.6.1 Stretching DNA . . . 21
  - 1.6.2 Stretching chromatin . . . 22
- 1.7 Rigid base pair Monte Carlo simulations . . . 24
- 1.8 Outline of this thesis . . . 26

**2 Methods for Unravelling DNA organization with single-molecule Force Spectroscopy using Magnetic Tweezers** **39**
- 2.1 Introduction . . . 41
  - 2.1.1 Chromatin organization by DNA-binding proteins . . . 41
  - 2.1.2 Magnetic tweezers . . . 42
  - 2.1.3 Measurement schemes . . . 42
  - 2.1.4 Data analysis . . . 43
- 2.2 Materials . . . 45
  - 2.2.1 Stock solutions . . . 45
  - 2.2.2 Isolation of DNA plasmids . . . 45
  - 2.2.3 Digestion and labeling of DNA . . . 45
  - 2.2.4 Flow cell assembly . . . 46
  - 2.2.5 Bead-tether assembly . . . 47
  - 2.2.6 Microscope . . . 48
- 2.3 Methods . . . 48
  - 2.3.1 Isolation of DNA plasmids . . . 48
  - 2.3.2 Digestion and labeling of torsionally free DNA constructs . . . 49
  - 2.3.3 Digestion and labeling of torsionally constrained DNA constructs . . . 50
  - 2.3.4 Flow cell assembly . . . 54
  - 2.3.5 Bead-tether assembly . . . 55
  - 2.3.6 Initial bead selection and height calibration . . . 56
  - 2.3.7 Force calibration . . . 57
  - 2.3.8 Force spectroscopy experiments . . . 58
  - 2.3.9 Rotational spectroscopy experiments . . . 59
  - 2.3.10 Data analysis of force spectroscopy experiments . . . 61
  - 2.3.11 Data analysis of rotational spectroscopy experiments . . . 62
- 2.4 Notes . . . 63

**3 Multiplexed Nanometric 3D Tracking of Microbeads Using an FFT-Phasor Algorithm 81**
3.1 Introduction . . . 83
3.2 Results . . . 86
3.2.1 Tracking of super-paramagnetic beads using Lorenz-Mie scattering theory 86
3.2.2 3D Phasor Tracking . . . 87
3.2.3 Performance of 3DTP . . . 90
3.2.4 Robustness of 3DTP . . . 94
3.3 Discussion . . . 97
3.4 Materials and Methods . . . 101
3.4.1 Magnetic tweezers setup . . . 101
3.4.2 Software . . . 101
3.4.3 Spectral analysis of tracking accuracy . . . 101
3.4.4 Lorentz–Mie scattering theory . . . 101
3.4.5 LUT-tracking . . . 102
3.4.6 Robustness simulations . . . 103
3.4.7 Characterization of mechanical vibrations and drift . . . 103
3.4.8 Quantification of the step size of DNA unwrapping . . . 104
3.5 Acknowledgments . . . 104
3.6 Supplementary Material . . . 105

**4 Rigid Base Pair Monte Carlo Simulations of 1-Start and 2-Start Chromatin Fiber Unfolding by Force 115**
4.1 Introduction . . . 117
4.2 Results . . . 119
4.2.1 Experimental unfolding of chromatin fibers with different NRLs . . . 119
4.2.2 Extension of a rigid base pair model with nucleosomes . . . 122
4.2.3 Nucleosome stacking in 2-start fibers . . . 125
4.2.4 Nucleosome stacking in 1-start fibers . . . 126
4.2.5 MC simulations of larger chromatin fibers . . . 127
4.2.6 Evaluation of the energies associated with fiber folding . . . 129
4.3 Discussion . . . 131
4.4 Materials and Methods . . . 133
4.4.1 Experimental force-extension data . . . 133
4.4.2 Rigid base pair Monte Carlo simulations of DNA . . . 133
4.4.3 Nucleosome (un-)wrapping in Monte Carlo simulations . . . 134
4.4.4 Nucleosome stacking in Monte Carlo simulations . . . 134
4.5 Acknowledgements . . . 135
4.6 Supplementary Material . . . 136
4.6.1 Reconstitution of chromatin fibers . . . 136
4.6.2 Flow cell preparation . . . 136
4.6.3 Force spectroscopy on single chromatin fibers using magnetic tweezers 136
4.6.4 Experimental data analysis . . . 137
4.6.5 Base pair step parameters . . . 138
4.6.6 Nucleosome wrapping step parameters . . . 138
4.6.7 Nucleosome stacking step parameters . . . 140
4.6.8 Excluded volume effects . . . 140
4.6.9 MC initialization . . . 141
4.6.10 MC iterations . . . 141
4.6.11 Simulation settings . . . 141
4.6.12 Supplementary experiments . . . 142

**5 A Critical Role for Linker DNA in Higher-Order Folding of Chromatin Fibers 149**
5.1 Introduction . . . 151
5.2 Results . . . 153
5.2.1 rbMC simulations yield the detailed conformation of linker DNA . . . 153
5.2.2 Twist as a decisive factor in fiber folding . . . 157
5.2.3 A 10 base pair periodic dependence of fiber folding parameters on NRL 159
5.2.4 H4 tails are essential for 2-start folding, but not for 1-start folding . . 164
5.2.5 Not only linker length, but also linker sequence modulates fiber folding 166
5.3 Discussion . . . 168
5.4 Materials and Methods . . . 171
5.4.1 rbMC simulations . . . 171
5.4.2 DNA substrates . . . 172
5.4.3 Chromatin reconstitution . . . 172
5.4.4 Force spectroscopy on single chromatin fibers . . . 172
5.4.5 Statistical mechanics model . . . 173
5.5 Acknowledgments . . . 173
5.6 Supplementary Material . . . 173
5.6.1 rbMC simulations . . . 173
5.6.2 DNA substrates . . . 174
5.6.3 Magnetic tweezers setup . . . 175
5.6.4 Statistical mechanics model . . . 175
5.6.5 Fitting procedure . . . 178
5.6.6 Fitted parameters of figures . . . 179

**6 Mechanical and Structural Properties of Archaeal Hypernucleosomes 191**
6.1 Introduction . . . 193
6.2 Results . . . 195
6.2.1 Both HMfA and HMfB cooperatively compact DNA . . . 195
6.2.2 HMfB organizes long DNA tethers into endless hypernucleosomes . . . 197
6.2.3 Stacking interactions mediate hypernucleosome formation . . . 200
6.2.4 The hypernucleosome forms a left-handed structure on DNA . . . 200
6.2.5 Unstacked HMfB-DNA complexes impose a right-handed DNA structure 202
6.3 Discussion . . . 203
6.4 Materials and Methods . . . 205
6.4.1 DNA substrate preparation . . . 205
6.4.2 Tethered Particle Motion . . . 206
6.4.3 Force spectroscopy and rotational spectroscopy . . . 206
6.4.4 Quantitative modeling of stretching the hypernucleosome . . . 206
6.5 Acknowledgments . . . 209
6.6 Supplementary Material . . . 210
6.6.1 Sequence of the TPM substrate . . . 210
6.6.2 Sequence of the torsionally free substrate . . . 210
6.6.3 Sequence of the torsionally constrained substrate . . . 211

**Summary 224**

**Samenvatting 229**

**List of Publications 232**

**Curriculum Vitae 234**

# Chapter 1

# Introduction

The hereditary information of a cell, its genetics, is contained within the sequence of base pairs of its DNA. Every human cell, for instance, contains approximately 2 meters of DNA when it is fully stretched out. This massive length needs to be compacted thousands of times, by folding it around nucleosomes which stack into dense chromatin fibers, to fit inside the cellular nucleus. Simultaneously, the useful information in the DNA needs to be dynamically accessible for the cell to sustain itself, respond to its surroundings, or adapt during tissue growth or prenatal development. How the cell realizes such a juxtaposition, strong compaction versus accessibility, is remarkable.

The level of chromatin compaction regulates which genes are actively read out, and which genes are transcriptionally silent. This epigenetic regulation can be triggered by a complicated interplay of post-translational protein modifications [1–3], transcription factors [4–6], gene repressors [7, 8], or environmental queues [9–11]. Despite decades of research into the nature of the chromatin fibers, there is still no consensus on its structure; 1-start, 2-start, or disorganized. Furthermore, the debate on the prevalence of such organized structures *in vivo* is ongoing to this day.

In this thesis, we used single-molecule force spectroscopy to study how the length of the linker DNA, the DNA between adjacent nucleosomes, impacts the structure of the folded chromatin fibers. Using a statistical mechanics model, we could extract their mechanical properties and quantify the energetic interactions between stacked nucleosomes. Single-molecule experiments have the advantage that they measure each tether individually, and are not subjected to ensemble averaging that can obscure unique properties. The bottleneck of single-molecule techniques, however, is the throughput. Here, nevertheless, we stretched more than 1000 chromatin fibers. This could be accomplished by the development of

a multiplexed Magnetic Tweezers apparatus, which could measure hundreds of molecules at once. Since multiplexing leads to an exponential increase in data throughput we developed a lightweight image analysis algorithm to process this data. In addition, we created a rigid base pair Monte Carlo model, to enrich our experimental data with simulations, which can provide structural insight that goes beyond interpretation of the structure through only the parameters *force* and *extension*. This combination of techniques provided a thorough examination of chromatin condensation and epigenetic regulation of Eukaryotes.

Furthermore, we investigated how archaeal histones compact DNA. Their mechanisms of compaction are suggestive of primordial bacterial-like transcription regulatory mechanisms preceding eukaryotic regulatory mechanisms. These experiments are highly relevant as Archaea are one of the main pillars of life, and form the link between humans and the last universal common ancestor in life on earth.

## 1.1 The Tree of Life

It is overwhelming to fathom the complexity, coexistence, and codependency of the many aspects of life that exist in a symbiotic relationship in the ecosystems on earth. Their staggering biodiversity can be vastly different, even though they frequently exist within proximity of each other. Life is everywhere: even places that seem hostile, such as in hydrothermal vents or deeply acidic environments, harbor life.

As Richard P. Feynman once mentioned, one can look at a flower and appreciate it as an artist, but the beauty of life is not limited to that [12]. Modern science grants us the opportunities to look inside the flower, inside the cells that constitute it, and appreciate the complicated intracellular mechanisms that work together in perfect harmony to sustain life.

The complexity of life originated 3.5 billion years ago from a simple, single cell. Every generation added layers of complexity, driven by evolution, adaptations to an ever-changing environment. With time, the world inside cells became increasingly complex. Our current understanding of evolution arose from the combination of two different ideas, by two different people, who walked the earth together but supposedly never met: Charles Darwin and Gregor Mendel.

1. **Darwin** realized that evolution depends on the existence of heritable variability within a species to generate the differences between ancestral and descendant populations, as stated in *On the Origin of Species*, his groundbreaking book from 1859 which is considered to be the foundation of evolutionary biology [13]. The book was highly controversial because of its volatile ideas which argued against creationist values. In addition, the theory was problematic for the scientific community since there was a lack of direct evidence for natural selection at the time, and genetics was not well understood.

2. **Mendel** investigated inheritance by cross-pollinating pea plants and discovered that crossbreeding of animals and plants could favor certain desirable traits [14]. He coined the terms *recessive* and *dominant*, and established many of the rules of heredity, which are now known as the laws of Mendelian inheritance. He published his ideas before a proper understanding of the cellular basis of sexual reproduction was achieved.

For many years, it was not obvious that Mendel's studies of heredity had any relevance to Darwin's theory of evolution. It would take nearly 60 years for their ideas to be combined in the modern theory of evolution. Today, the most compelling evidence for evolution can be found in the fact that even though highly-developed multicellular organisms such as humans are vastly different and infinitely more complex than simple, unicellular organisms, we share innumerable mechanisms that sustain (our) life.

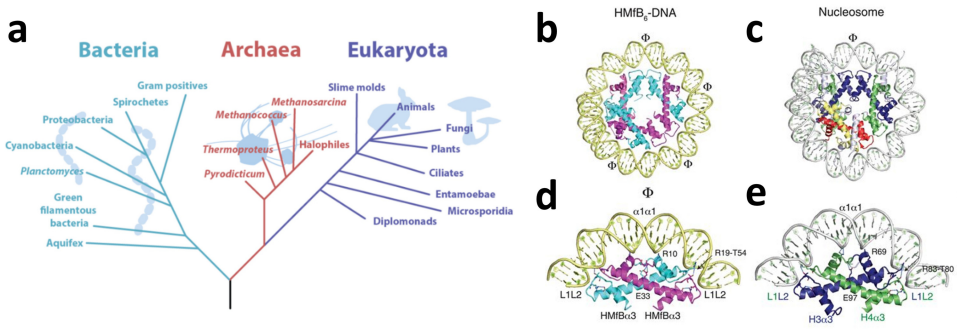



**Figure 1.1**
**Nucleosome folding in humans is highly analogous to DNA compaction in Archaea.** a) The phylogenetic tree of life shows that Archaea form a third branch between Bacteria and Eukaryota. Figure reprinted from Wikimedia Commons. A comparison of three archaeal histone HMfB homodimers (b) and a hexasome (one H2A-H2B dimer and histone tails are removed) (c) shows that Archaea express a form of histones that are structurally homologous to eukaryotic human histones. The similarity of the histone fold of an HMfB homodimer (d) and an H3-H4 heterodimer (e) is remarkable. Figure adapted from [15].

By studying how these mechanisms increase in complexity between species we can retrace our evolutionary path and organize it in the tree of life, which shows how all kingdoms of life are related, including living and extinct species. Figure 1.1a depicts part of the phylogenetic tree of life, dividing the living world into three domains: Bacteria, Archaea, and Eukaryota. The trunk of the tree links the three domains together, indicating that all life originates from a common ancestor.

According to most definitions, the cell is the smallest common denominator of life. Some cells are complete organisms in themselves, where others are part of a multicellular organism. Cells can respond to environmental queues, metabolize, reproduce and make copies of themselves. Frequently, unicellular organisms live with multicellular organisms in a complex symbiosis. Cells can look vastly different from each other. Imagine, for instance, the brick-like structure of a plant cell versus the complicated structure of a nerve cell with its long protrusions. These are both cells, however, they are unrecognizable from each other at first glance. Nevertheless, all cells share distinct universal features and are made up of the same major classes of organic molecules: nucleic acids, proteins, carbohydrates, and lipids.

There are two categories of cells: **prokaryotes** and **eukaryotes**. Prokaryotic cells are the simplest form of life. They are formed by a membrane surrounding the cytoplasm and contain no nucleus or organelles. Prokaryotes are colloquially known as Bacteria. Regardless of their simplicity, they constitute a huge fraction of life of Earth. Eukaryotic cells are more complicated,

and contain membrane-bound organelles next to a nucleus. They can exist as unicellular organisms, or multicellular organisms such as fungi, plants, animals. Eukaryotes can only live with oxygen, while some prokaryotes are able to live using different substances, hence, prokaryotes can be found in almost every environment on Earth.

Part of the complexity of eukaryotes seemed to have arisen from when a larger prokaryotic cell engulfed a smaller one and both continued to live in a symbiotic relationship [16]. The engulfed cell thrived within the larger one as an organelle. In particular chloroplasts and mitochondria are thought to have arisen in this manner, as they still retain a separate genome and separate individually.

The microbiologist Carl Woese argued in 1977, based on the analysis of ribosomal RNA, that the domain of prokaryotes should be further divided into Bacteria and Archaea (as shown in Figure 1.1a) [17–19]. Archaea are also prokaryotic unicellular organisms (without compartmentation), but their genetic makeup is vastly different from Bacteria. Archaea possess a unique evolutionary history and diverse metabolisms, which allow them to feed on inorganic matter and to withstand extreme environments such as high temperatures, extremely salty condition, or places without oxygen.

Even though unicellular Archaea are evolutionary quite distant from eukaryotes we can observe remarkable similarities in cellular mechanisms, for instance, in genomic compaction and organization. Eukaryotic cells compact three-quarters of their genome by wrapping their DNA around a histone octamer to form a nucleosome. Figures 1.1b and 1.1c illustrate that nucleosome folding in humans is highly analogous to DNA compaction in Archaea. Figures 1.1d and 1.1e furthermore show that the structure of the building blocks is virtually identical. These similarities, between two species that are strongly separated on the phylogenetic tree of life, form compelling evidence that all life on earth today originated from a common ancestor.

## 1.2 DNA

If life on earth can be expressed as a tree, the root of the tree is surely formed by DNA. DeoxyriboNucleic Acid (DNA) is the most important molecule shared by our common ancestor and virtually all lifeforms today. Even viruses, a peculiar form of life, carry genetic material. DNA contains the blueprint for all life: it dictates growth, development, function, and regulates reproduction. DNA is unique to each individual, although all cells in a multicellular body share the same DNA.

Many researchers need to be credited for the discovery and understanding of DNA. DNA was first isolated in 1869, by the Swiss physician Friedrich Miescher in the nuclei of white blood cells [20]. He came across a substance that was

high in phosphorus content and was resistant against proteolysis. He coined the substance *nuclein*. The implications of the discovery were long underestimated. For years, scientists continued to believe that proteins held all of our genetic content, as they could not fathom that nuclein was complex enough to contain all of the information.

Several decades later, the American researcher Phoebus Levene, specialized in chemistry of biomolecules, discovered the three major components of a single nucleotide in 1909: phosphate, sugar, and a nucleobase [21]. He furthermore identified that there were 4 unique nucleobases: adenine (A), thymine (T), cytosine (C), and guanine (G) which were apparently randomly alternating. The biochemist Erwin Chargaff was subsequently credited with two major discoveries. First, he discovered that the total amount of purines (A + G) and the total amount of pyrimidines (C + T) were equal, commonly referred to as Chargaff's rule. Second, he found that the nucleotide composition differed among different species – revealing its hidden purpose [22, 23]. These discoveries paved the way for a revolutionary publication that saw the light in the early fifties.

James Watson and Francis Crick famously resolved the structure of DNA

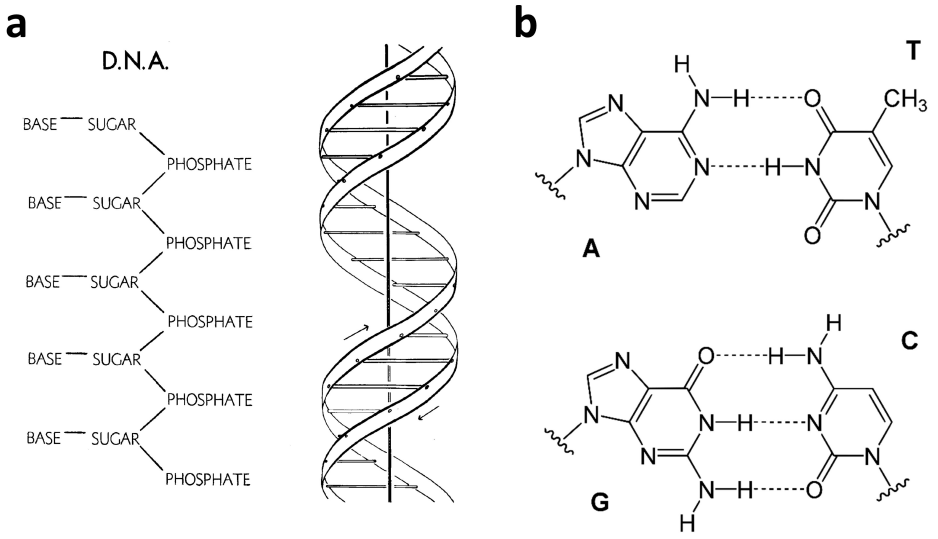



**Figure 1.2**
**The structure of DNA.** a) The chemical structure of DNA, as originally published by Watson and Crick in 1953 [24], consists of a sugar-phosphate backbone and nucleobases that contain the genetic code. Figure reprinted from [24], b) Schematic representation of base pairing through hydrogen bonds (dashed lines). Adenine (A) pairs with Thymine (T) by means of two hydrogen bonds, and Guanine (G) pairs with Cytosine (C) by means of three hydrogen bonds. Figure reprinted from Wikimedia Commons.

[24]. They relied heavily on the work of others, not only on Miescher, Levene, and Chargaff, but particularly the X-ray crystallography work by Rosalind Franklin and Maurice Wilkins inspired them to imagine the double-helical structural model of DNA [25–28]. The research by Linus Pauling on molecular distances and bond angles was also crucial to their work [29–31]. A major insight was to realize that DNA consisted of two complementary antiparallel strands. Watson and Crick put together all the pieces of the puzzle that resulted in the structure of DNA.

The structure that was resolved was of B-DNA, the most common form of DNA. Figure 1.2a depicts the original structure as it was published in 1953 [24]. This iconic figure describes the archetypical double-stranded helix, 2 nanometer (nm) in diameter. Watson and Crick identified five main characteristics of the DNA structure:

1. Each strand is composed of a rigid sugar-phosphate backbone decorated with a variable sequence of nucleobases.
2. The sugar-phosphate backbone has a polarity: the orientation of the strand is 5' to 3' (or reversed), which refers to the $5^{\text{th}}$ and the $3^{\text{rd}}$ carbon atom in the sugar pentose to which the phosphate attaches.
3. The two strands are of complementary sequence.
4. The helix has a 10.4 base pairs pitch, with asymmetric minor and major grooves. The asymmetry allows for certain sequences to be recognized by protein or restriction enzymes. The outer edges of the nucleobases are exposed: available for hydrogen bonding, for instance interacting with (regulatory) proteins.
5. B-DNA is right-handed. Other forms of DNA are left-handed, for instance, A-DNA. An obscure form of DNA is Z-DNA, which is right-handed and possesses different mechanical and structural properties.

Since its publication in 1953, there have only been minor changes to the iconic structural model of DNA.

The genetic code is composed of a sequence of nucleobases. Figure 1.2b illustrates each of the nucleobases and their complementary pairing: A is always paired with T by means of two hydrogen bonds, while C is always paired with G by means of three hydrogen bonds. The different number of hydrogen bonds indicates that the stability of base pairs is not equal: the bond strength between C and G is approximately 0.75 $k_{\text{B}}T$ (-21 kcal/mol) where the bond strength between A and T is 0.5 $k_{\text{B}}T$ (-13 kcal/mo) at physiological conditions[1] [33].

[1] $k_{\text{B}}T$ is the product of the Boltzmann constant and the temperature, which at room temperature equals $4.11 \times 10^{-21}$ Joule. $k_{\text{B}}T$ is frequently used as a unit of energy in biophysics. For example, the interaction energy between nucleosomes (25 $k_{\text{B}}T$) or between Archaeal histone proteins (2 $k_{\text{B}}T$) can be expressed in $k_{\text{B}}T$. Furthermore, the equipartition theorem states that each degree-of-freedom in an object contains 0.5 $k_{\text{B}}T$ in thermal equilibrium [32].

The strands of DNA are complementary. This degeneracy ensures that the genetic information is backed up on the antiparallel strand. Watson and Crick had a foresight that turned out to be crucial in the understanding of genetics, inheritance, and evolution: "*It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material*" [24]. They understood that DNA contains all genetic information, and two copies of this information are stored on both strands, sustaining an efficient copying mechanism. What they nevertheless could not yet comprehend is how all this information is (selectively) accessed.

Information flows from the nucleus of the cell by means of RiboNucleic Acid (RNA): a biopolymer similar to DNA. Both RNA and DNA are nucleic acids, composed of a sugar-phosphate backbone decorated with nucleobases. RNA consists of a single strand, however, and in RNA the thymine is replaced by uracil (U). Unlike DNA, an RNA is stabilized by intramolecular interactions, resulting in unique higher-order structures.

The flow of information in the cell follows two distinct steps: transcription and translation.

1. Transcription: The genetic code is read out from the DNA by protein-complexes such as RNA polymerase. During transcription, a transcript is composed of pre-messenger RNA (pre-mRNA). The pre-mRNA structure is spliced to get rid of redundant information, and the mRNA transcript is excreted from the cell nucleus.

2. Translation: The mRNA transcript is translated into protein by a huge protein-RNA complex, the ribosome. Aided by various chaperone proteins, the newly synthesized protein is folded into its desired, functional shape.

The flow of information in the cell is frequently referenced as *The Central Dogma of Molecular Biology*, a term coined by Crick [34]. The dogma describes that there are two directions in the information flow: from DNA to DNA (DNA replication), and from DNA to RNA (transcription). Interestingly, RNA viruses follow an alternative route, from RNA to RNA (copying themselves), and from RNA to protein (translation). In both cases, these flows never seemed to go the other way around, hence the stated dogma.

The discovery of reverse transcriptase in 1970, which allows information to flow from RNA to DNA, challenged the dogma [35]. The DNA double-strand break repair mechanisms homologous recombination and non-homologous end joining undermined the dogma further. In the case of homologous recombination, the damaged genetic information of the cell is rewritten from a substrate that is (preferably) the sister chromatid [36]. In the case of non-homologous end joining, the damaged single-stranded overhangs form the template, although there are even models advocating for direct RNA-templated DNA repair [37]. In addition, retroviruses use RNA to alter the DNA sequence of their host [38].

In short, there is mounting evidence that alterations of DNA based on RNA sequence is not merely restricted to exceptional cases, but is a more general phenomenon in nature.

For a long time, Crick's essential argument still held: there was no evidence that the information in a DNA sequence could be rewritten from the protein level [39]. In recent years, however, this notion was challenged with the discovery of the CRISPR-Cas9 system, essentially a prokaryotic immune system that directs controlled changes in DNA sequence guided by an RNA template. This system was successfully used to manipulate the human genome in 2013 [40–42]. Another mechanism that regulates the expression of genes is RNA interference (RNAi). RNAi is a process where small RNA molecules (microRNA and small interfering RNA) degrade mRNA, resulting in post-transcriptional gene silencing [43]. RNAi is a natural process that cells employ to destroy RNA-based viruses. However, these non-coding RNAs can also modulate transcription suggesting alternative mechanisms of in genetic organization that are not contained in the central dogma of microbiology.

## 1.3 Genetics and epigenetics

**Genetics** is the study of the information that is contained within the genome, genetic variation, and heredity in organisms. Studying the genome may provide information on the physical characteristics of an individual: *e.g.* if someone has blue eyes or brown, is tall or short, or the presence and shape of a molecular motor. Studying the genome does not, however, reveal anything on the availability of the information to the individual cell. For this, one needs to know whether the genes are actually expressed, and this is regulated by epigenetics.

**Epigenetics** is the study of the regulation of genetics and developmental changes caused by modification of gene expression. It can tell us how, where, and when the genetic information is accessed [44]. For example, during cellular interphase when the cell is at rest, the cell only accesses the housekeeping genes: the genes coding for the proteins necessary to maintain the functionality of the cell. The situation changes when, for instance, the external factors change, such as salt conditions, hypoxia, or after DNA damage. The cell has to react and swiftly access the genes that can counteract potentially perilous external factors. More substantial are the complicated epigenetic mechanisms at play during stem cell differentiation at different positions within the body. How does a cell know how to develop when it has the information of the entire genome available?

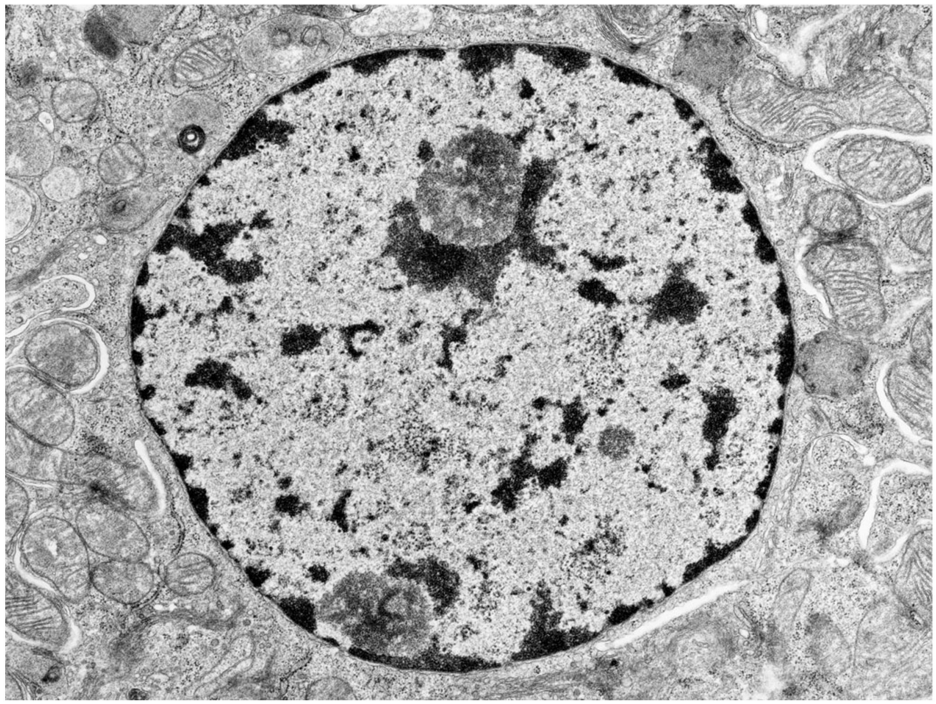


**Figure 1.3**
**Chromatin in a cell nucleus exists in two forms: heterochromatin and euchromatin.** A cell nucleus was imaged with transmission electron microscopy (TEM) and stained, highlighting the packing of the chromatin. Heterochromatin is the most densely packed, yielding darkly stained regions within the nuclear envelope. Euchromatin is less densely packed and actively transcribed, yielding the lightly stained regions. Figure reprinted from [45].

Cells modify the structure in which DNA is organized to keep the genes repressed or active. In the nucleus, the DNA is globally observed to be in one of two states: in tightly packed, gene-poor transcriptionally silent **heterochromatin**; or in loosely packed, gene-rich, transcriptionally active **euchromatin** [46]. Figure 1.3 depicts the nucleus of a cell where this bimodal behavior of DNA can be observed. Heterochromatin is highly condensed and appears as the darkly stained regions. Euchromatin is loosely condensed and is not readily stainable. In general, the density of DNA packing is indicative of the frequency of transcription.

Epigenetic regulation within a multicellular organism varies at different locations in the body. Figure 1.4 depicts epigenetic regulation of a part of the genome during embryonic development. The cells in developing skin tissue are

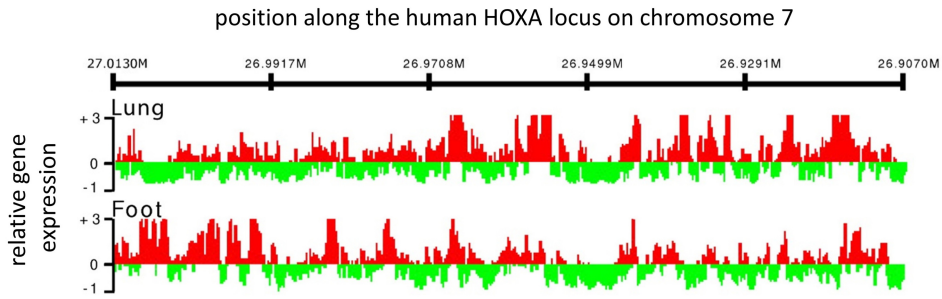



**Figure 1.4**
**Epigenetics: gene regulation in action.** The epigenetic activity of the human HOXA cluster on chromosome 7 shows a tight regulation of transcription in distinct parts of the body. The HOXA cluster is part of the Hox genes that specify development of regions of the body. In other words, it contains a plan of the embryo, and ensures that the correct structures form in the correct places of the body. The relative gene expression in the developing cells of the skin is very different than that of developing cells in the lung. This differentiation is thought to be mediated by changes in chromatin. Figure adapted from [47].

transcribing different parts of the genome in comparison to developing cells in the lungs [47].

Dysfunction of epigenetic regulation results in various medical conditions. For instance, the epigenetic disorder type 1 diabetes is a medical condition that is characterized by abnormally high blood sugar levels, when beta cells in the pancreas stop producing insulin [48]. Type 1 diabetes frequently develops during adolescence, demonstrating that the body has the genetic information to produce insulin, but due to a complicated interplay between genetics (heritability), environment (age, obesity, nutrients), and epigenetic factors (DNA/histone methylation); the body stops producing it. This results in the inability to use glucose for energy or to control the amount of sugar in the blood [49, 50]. At the base of this epigenetic condition lies the DNA compaction, which controls the accessibility of the transcription machinery to the DNA, effectively switching genes on or off.

## 1.4 Chromatin

For an organism to survive, DNA needs to be compacted in an organized manner. In every human cell approximately 2 meters of DNA needs to fit in the cell nucleus, which is on average only 6 µm in diameter, while retaining accessibility to sustain cell function. To accomplish this, the DNA is carefully compacted into higher-order structures such as chromatin, as schematically depicted in Figure 1.5.

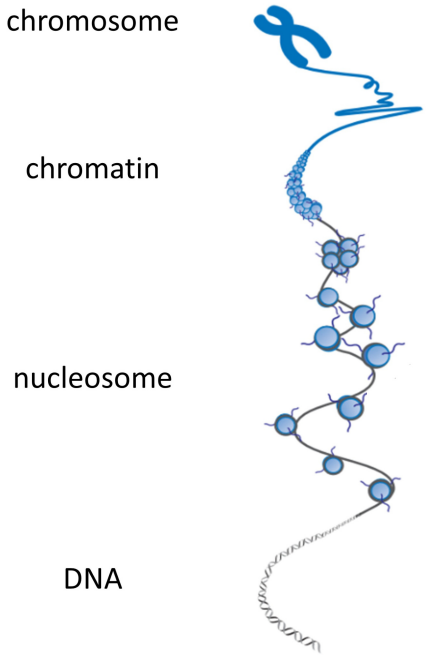



**Figure 1.5**
**The compaction of DNA.** To condense DNA to fit the genetic material in the nucleus, it is compacted in multiple steps. In the first level of compaction, DNA is wrapped around nucleosomes to form a beads-on-a-string conformation. Subsequently, the nucleosomes can interact to form 30-nm chromatin fibers. Ultimately, chromatin fibers can interact and condense to form chromosomes during mitosis, the highest level of compaction. Figure reprinted from [51].

In the first level of compaction, DNA wraps around a histone octamer to form a nucleosome. In the second level, nucleosomes interact to form dynamic chromatin fibers. In the third level, chromatin fibers interact and aggregate to form even higher-order structures to achieve further compaction. Despite several decades of research, there is still a lot of ambiguity about the structure or mechanical properties of these complexes. We know that during mitosis the DNA enters the highest form of compaction, where it is condensed 10,000 times into chromosomes. These characteristic structures are visible under the light microscope, however, many details on a molecular level are missing [52].

### 1.4.1 Nucleosomes

The fundamental unit of chromatin is the nucleosome. The nucleosome forms when two turns of DNA wrap in a left-handed superhelix around a histone core. The core is made up of four pairs of histones: histone H2A combines with H2B,

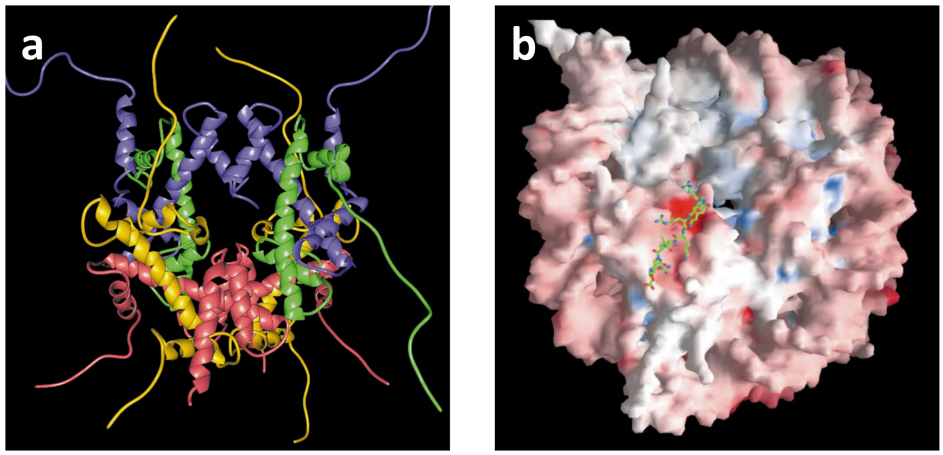



**Figure 1.6**
**The structure of the nucleosome.** a) The nucleosome consists of a histone octamer core, built from two types of heterodimers: H2A–H2B (yellow-red) and H3–H4 (blue-green). All histones have tails that can interact with the DNA, adjacent histones, or adjacent nucleosomes. The DNA is omitted in this image. Figure reprinted from [56]. In the crystal structure of the nucleosome core particle at 2.8 Å resolution the acidic patch in the H2A–H2B dimer (bright red) and the H4 tail from an adjacent nucleosome were resolved, indicating a mechanism for nucleosome stacking. Figure reprinted from [57].

and histone H3 combines with histone H4. Two H3-H4 dimers assemble to form a tetramer onto which two H2A-H2B dimers dock to complete the histone octamer. In most organisms, the nucleosome is complemented by an H1 linker histone. Linker histones bind to specific DNA geometries and prefer nucleosomal DNA over free DNA [53]. Incomplete nucleosomes also exist: tetrasomes (two H3-H4 dimers) and hexasomes (two H3-H4 dimers + one H2A-H2B dimer). These incomplete nucleosomes are relevant in the biological context since they, for instance, relieve transcription induced stress [54, 55].

Nucleosomes were first observed on EM images of chromatin by Olins *et al.* in 1974 [58]. In that same year, their structure, composed of a histone octamer core wrapped by approximately 200 base pairs of DNA, was proposed by Kornberg [59]. A high-resolution crystal structure of the nucleosome was subsequently resolved by Luger *et al.* in 1997 [57]. The nucleosome constrains 146 or 147 base pairs and many different crystal structures of a nucleosomes were resolved, all featuring the same 1.65 wraps around the histone core [60]. Figure 1.6a depicts the histone octamer core, with the highly disordered and therefore unresolved histone tails, omitting the DNA. In recent years, crystallography of single nucleosomes has been refined, for instance by the work of Hitoshi Kurumizaka [61, 62]. He furthermore composed a song to commemorate the

$20^{\text{th}}$ anniversary of Luger's landmark publication (scan the QR code to learn more).

In eukaryotes, all histones consist of a globular part and a flexible tail, as becomes clear in Figure 1.6a. Histone tails fulfill a variety of functions *in vivo*, for instance in epigenetic gene regulation. The histone tails can undergo post-translational modifications (PTMs) such as acetylation, methylation, phosphorylation, and ubiquitination that directly or indirectly influence chromatin structure and control gene activity [63]. Acetylation and phosphorylation generally result in transcriptionally active genes, where methylation and ubiquitination are often associated with gene silencing[2] [66].

Histone tails are primarily involved in the mechanical stability of nucleosomes and chromatin [67, 68]. The positively charged tails can interact with the nucleosomal DNA, linker DNA, and acidic patches on the histone core by means of hydrogen bonds. Nearby nucleosomes can stack through interactions of the H4 tail with the H2A-H2B acidic patch of the adjacent nucleosome. This mechanism is indicated in Figure 1.6b [57, 68–73]. The H3, H2A, and H2B tails may also contribute to nucleosome stacking [68]. Stacked nucleosomes can form regular chromatin fibers. At low ionic strength, these fibers exist as a beads-on-a-string configuration. However, in physiological conditions, and especially in the presence of divalent salt, the nucleosome chains are strongly compacted into organized chromatin fibers [74].

### 1.4.2 Chromatin fibers

The structure of the chromatin fiber has a strong effect on the level of compaction and accessibility of the DNA, hence, it is inherently related to epigenetics. There are two main structural models to describe chromatin fibers: solenoid fibers and zig-zag fibers. Solenoid fibers form when neighboring nucleosomes interact to create a single stack of nucleosomes (1-start structures). The solenoid fiber was first described by Thoma *et al.* in 1979 [75]. The original structural model of the solenoid fiber is depicted in Figure 1.7a. Zig-zag fibers form where

[2]Epigenetic regulation is not only controlled by PTMs, but by a complicated interplay of transcription factors, gene repressors, or environmental queues. The DNA itself can be methylated in a promoter region which causes gene silencing. In addition, the cell employs several ATP-dependent remodeling complexes that displace, exchange, or evict histones from the chromatin fiber, for instance, remodelers from the SWI2/SNF2 group or the SWI group [64]. Furthermore, the cell employs several histone chaperones: complexes that control the folding of free histones into nucleosomes, which work together with the remodelers on histone deposition and eviction [65]. Frequently, a combination of mechanisms is involved. For instance, DNA methylation in combination with histones de-acetylation results in highly compacted chromatin and transcriptionally silent genes.

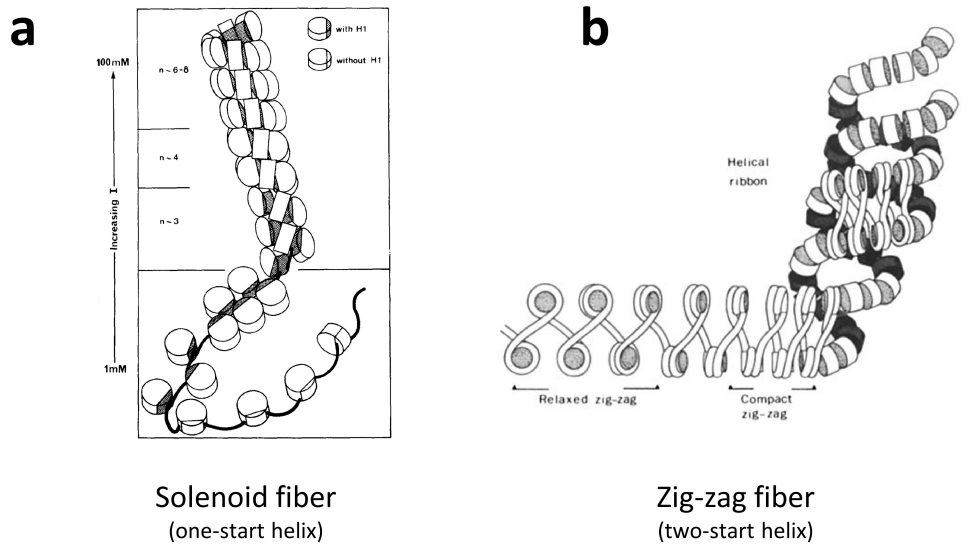



**Figure 1.7**
**Two structural models have been proposed for the chromatin fiber.** a) The solenoid fiber, as described in 1979 by Thoma *et al.* [75], forms a 1-start structure with adjacent nucleosomes stacked together. The presence of the H1 linker histone assists in the formation of the solenoid structure and the ionic strength controls the degree of compaction. Figure reprinted from [75]. b) The zig-zag fiber, as described in 1984 by Woodcock *et al.* [76], forms a 2-start structure with nucleosomes stacking with their next-neighbor. Figure reprinted from [76].

next-neighboring nucleosomes interact to create a double stack of nucleosomes (2-start structures), which rotates around itself (much like the double helix of DNA). The zig-zag fiber was first described by Woodcock *et al.* in 1984 [76]. The original structural model of the zig-zag fiber is depicted in Figure 1.7b. The prevalence of the solenoid fiber, the zig-zag fiber, or the absence of regular higher-order structures is still being debated today.

The experimental evidence for the zig-zag fiber is quite extensive, exemplified by the X-ray structures by Schalch *et al.* [77] and the cryo-EM structures by Song *et al.* [70], revealing the structure of the tetranucleosomal units that constitute zig-zag fibers in great detail. Digestion experiments by Dorigo *et al.* were strong indications for the 2-start structure as well [72]. The solenoid fiber has not been resolved by single-particle cryo-EM reconstruction or X-ray diffraction, perhaps due to the dynamic nature of this structure. Robinson *et al.* used EM to measure the diameter of folded chromatin fibers. The independence of this diameter on linker DNA length indicated a structure of compacted solenoid fibers [78, 79]. In addition, force spectroscopy experiments by our

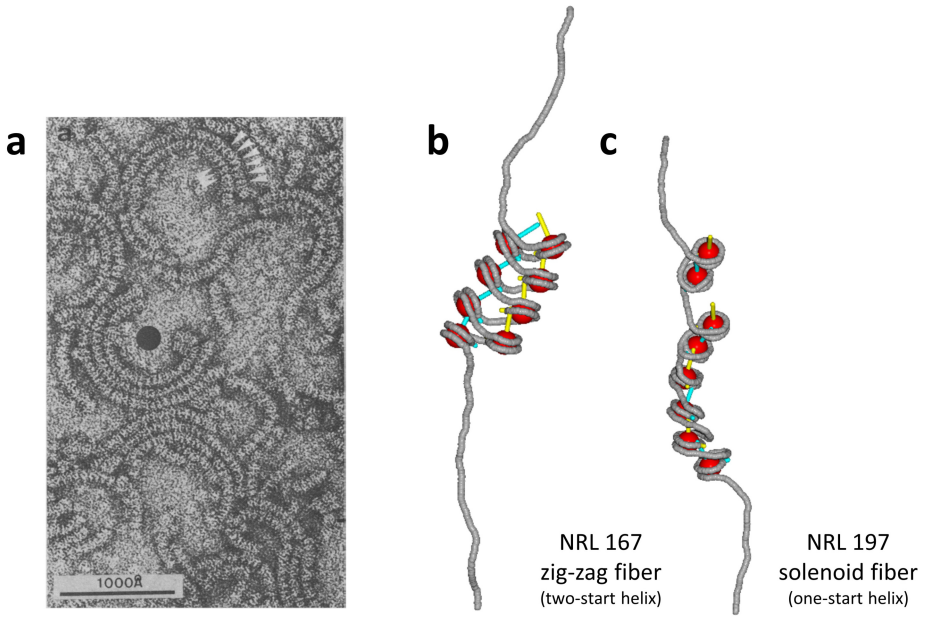



**Figure 1.8**
**Linker DNA plays an important role in the higher-order structure of chromatin fibers.** a) Nucleosome core particles, without linker DNA, tend to form stacks that describe long arcs, as shown by the EM data. Figure reprinted from [83]. b) Chromatin with relatively short linker DNA, for instance NRL 167, forms a zig-zag structure where the linker DNA is approximately straight. c) The relatively long linker DNA in NRL 197 can accommodate a solenoid structure. The chromatin fibers in panel b) and c) depict chromatin fibers in a force spectroscopy experiment, flanked by long handles of bare DNA simulated using rbMC (described in Chapter 4). Two nucleosomes in the solenoid fiber have unstacked because of the exerted force.

group provided evidence for solenoid fiber structures for chromatin fibers with long linker DNA [80, 81]. Cross-linking experiments by Kaczmarczyk *et al.* suggested that these higher-order structures are stacked by means of the H4 tail interacting with an acidic patch on a neighboring nucleosome, arguing in favor of the solenoid fiber [82].

The linker DNA, the DNA between adjacent nucleosomes, plays a crucial role in the higher-order structure of chromatin. By connecting the nucleosomes, linker DNA may impose geometrical constrains that can affect higher-order folding. Dubochet *et al.* showed in 1978 that nucleosome core particles that lack linker DNA stacked into long arcs, as depicted in Figure 1.8a [83]. These experiments show the preferred stacking orientation of nucleosomes in the absence of constrains imposed by linker DNA.

The introduction of nucleosome positioning sequences, such as the Widom 601 sequence [84], made it possible to reconstitute well-defined nucleosomal

arrays, which allowed for a systematic analysis of the effect of linker length on fiber structure. Grigoryev *et al.* used sedimentation analysis and EM to quantify the compaction of chromatin as a function of linker length in 5 base pair steps. The experiments suggested strong compaction into 2-start structures, especially for short linker lengths [85–87]. Previous work by our group revealed that fiber stiffness and unfolding characteristics also depended on linker length, suggesting a 2-start structure for short linker lengths and a 1-start structure for long linker lengths [80–82]. The two alternative higher-order structures are depicted in Figure 1.8b, illustrating the required deformation of the linker DNA to sustain such structures.

*In vivo*, linker length is quantized, often in multiples of 10 base pairs, corresponding to the rotational pitch of DNA [88]. Linker DNA length correlates with the level of compaction. Fibers with mostly $10n$ linker length ($n = 1, 2, \dots$) are strongly condensed and are observed in transcriptionally silent chromatin [85, 89]. In contrast, fibers with predominantly $10n + 5$ linker length are less condensed and often correspond to transcriptionally active genes [90–92].

Next to the debate on the topology of the fiber, there is also no consensus on its occurrence *in vivo* [93]. The group of Kazuhiro Maeshima advocates that the structures that are found *in vitro* do not naturally occur *in vivo*. They imaged vitrified human mitotic cells using cryo-EM in a semi-native state and observed no structures with a diameter larger than 11 nm. They concluded that chromatin must be highly disorganized *in vivo*, which permits a more dynamic and flexible genome organization than would be allowed by regular structures [94–96]. Recent work, however, contradicts their view. Risca *et al.* measured contacts between alternating nucleosomes based on RICC-sequencing, which suggests the manifestation of 2-start structures *in vivo* [97]. Fierz *et al.* argued that if the chromatin fiber exists *in vivo*, 30-nm structures may occur over only a few kilobases or tens of nucleosomes, based on data from EM and chromatin capture techniques [98, 99].

On a scale of tens of kilobases to a few megabases, the mammalian genome is further organized into topologically associating domains (TADs) [100, 101]. These domains occupy discrete positions within the nucleus. Hi-C profiles show that throughout the different phases in the cell cycle and during cell differentiation, different parts of the genome associate with each other, suggesting epigenetic regulation on a much larger scale than that of the chromatin fiber [102]. An important factor in TAD formation and maintenance is the transcriptional repressor CCCTC-Binding factor (CTCF), which brings together remote strands of DNA, forming chromatin loops, and anchors these loops to cellular structures such as the nuclear lamina. In the nucleus, CTCF co-localizes with cohesin, a ring-shaped multi-protein complex that can trap DNA within the ring. The inner diameter of the ring is approximately 40 nm, which can physically accommodate compacted chromatin fibers [103]. CTCF and cohesin define the boundaries between transcriptionally active euchromatin and heterochromatin,

and there is growing evidence that TADs are formed by active extrusion of chromatin loops by cohesin [104]. A future challenge would be to illuminate the relation between the epigenetic regulation on the chromatin fiber level and the TAD level, as at the mechanisms that underlie the formation of TADs are poorly understood.

## 1.5 Force spectroscopy experiments

It appears that the discussion about the structure of chromatin fibers, *in vitro* or *in vivo*, is still ongoing and structural measurements at the level of the chromatin fiber are lacking. In this thesis, I aim to add to this discussion by using single-molecule force spectroscopy on *in vitro* reconstituted chromatin fibers. Single-molecule force spectroscopy is a powerful tool way to study interactions, binding forces, and mechanical parameters of biopolymers. The extension of the molecule-of-interest is characterized as a function of force, which can be probed by a variety of techniques, such as magnetic tweezers (MT), optical tweezers (OT), atomic force microscopy (AFM), or acoustic force spectroscopy (AFS) [105, 106].

With these experiments we were able to resolve the composition of individual chromatin fibers, the forces that hold DNA and histones together, and the interactions between nucleosomes. Cui *et al.* used OT to stretch the first (native) chromatin fiber in 2000. They described the mechanical properties of chromatin fibers that were recovered from chicken erythrocytes [107], but were unable to identify the unfolding of individual nucleosomes in the force-extension curves. Bennink *et al.* stretched chromatin fibers, reconstituted *in situ* on λ-DNA, with OT in 2001 [108]. They noticed discrete transitions that were attributed to unfolding of full nucleosomes. However, the step sizes were unequal in length. Brower-Toland *et al.* examined the unfolding of regular, *in vitro* reconstituted, nucleosomal arrays with a feedback-enhanced OT in 2002 and were the first to resolve two transitions in the unfolding of individual nucleosomes. The outer turn (76 base pairs) unfolded at low forces and the inner turn (80 base pairs) required higher force [109].

Kruithof *et al.* used MT to analyze the low-force transition in chromatin fibers and was able to resolve a highly compliant helical folding for the 30-nm chromatin fiber [80]. Their research laid the foundation for the development of the statistical mechanics model by Meng *et al.* [81], which identified a previously undetected third nucleosomal conformation, the extended conformation, between the singly-wrapped and the completely unwrapped conformation. In this publication the relationship between linker DNA length and fiber structure was first explored with MT. In 2016, Li *et al.* were able to separate the unstacking transition from the DNA unwrapping of the outer turn with MT, by stretching the chromatin fibers in zero-salt conditions [110].

In our group, Hermans *et al.* used multiplexed MT to study native chromatin [111]. These fibers, composed from 18S ribosomal DNA, were extracted from the nucleus of the yeast *Saccharomyces cerevisiae* using Locked Nucleic Acid (LNA) probes. Since native chromatin is assembled *in vivo*, it is highly heterogeneous due to PTMs of the histones and the absence of nucleosome positioning sequences Therefore, analysis requires higher statistics to recover common features. Hermans *et al.* measured a 24.8 nm width of the stepwise unwrapping transition, which was comparable to *in vitro* reconstituted chromatin.

In this thesis, we used multiplexed MT to thoroughly examine the effect of linker DNA length on the mechanical parameters and higher-order structure of eukaryotic chromatin (*see* Chapter 5). Furthermore, we stretched archaeal chromatin, composed of HMfA or HMfB proteins, to infer their higher-order structures (*see* Chapter 6). Figure 1.9 depicts our experimental setup as described by Kaczmarczyk *et al.* [112]. During a typical experiment, a chromatin fiber, flanked by bare DNA handles, is tethered to a glass cover slip by means of a digoxigenin – anti-digoxigenin interaction. The other end of the molecule is tethered to a paramagnetic bead by means of a streptavidin – biotin interaction. Mounted above the sample is a set of horizontally oriented magnets that yield a uniform magnetic field over the entire field-of-view [113]. The force on the beads scales exponentially with the height of the magnets and causes the tethered DNA-chromatin complex to stretch and unfold. In addition, the magnetic field can be rotated to exert torque, which can build up in the tether if it is torsionally constrained. Torque has a distinct effect on the structure of DNA and chromatin [81, 114–116]. In this thesis, we used torque to stabilize the archaeal HMfB fiber (*see* Chapter 6).

## 1.6 Particle tracking

The diameter of DNA (2 nm) or chromatin (30 nm) is too small to resolve with a light microscope. Hence, to detect structural changes in chromatin fibers we measure their extension by tracking the position of the paramagnetic bead attached to the tether. In our experimental setup, we use a collimated LED to illuminate our beads, resulting in a characteristic diffraction pattern that is recorded by a camera. This circular hologram arises from interference between the incident beam and the beam scattered off the bead. The size of the diffraction pattern scales with bead height. Therefore, the three-dimensional position of the magnetic beads can be deduced from the two-dimensional hologram.

There are a plethora of bead tracking techniques available for MT. When only the in-plane position of the bead is desired, the $xy$-position of the bead could, for instance, be resolved by a center-of-mass calculation [117–119], or

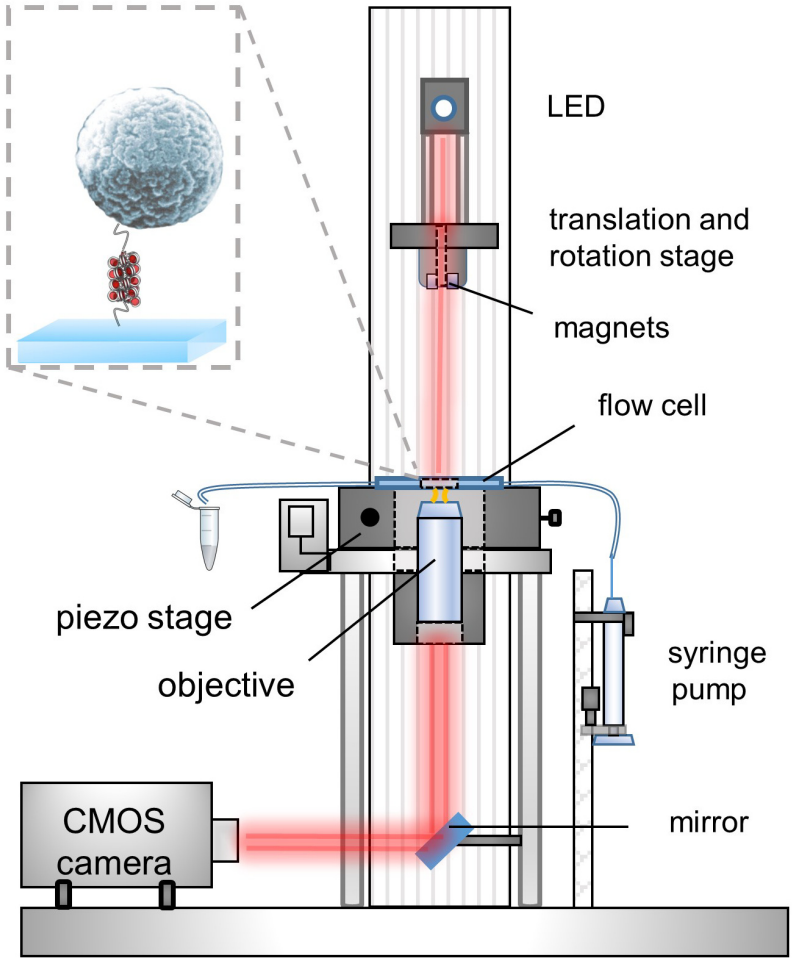



**Figure 1.9**
**Chromatin fibers are stretched by a multiplexed magnetic tweezers apparatus.** The setup was able to measure hundreds of chromatin molecules in parallel. The inset shows a chromatin fiber, flanked by long handles of bare DNA, tethered to a paramagnetic bead (not to scale). Image reprinted from [112].

by fitting the diffraction pattern with a Gaussian profile [120–122]. These algorithms are simple, fast, and yield nm precision in two dimensions [123].

Further analysis of the hologram is required to extract the three-dimensional position of the bead. Lee *et al.* fitted the diffraction pattern with Lorentz–Mie scattering theory, which resolved the refractive index as well as the three-dimensional position of the bead with nm precision, at the cost of being rather computationally intensive [124].

A more frequently implemented algorithm in MT is based on the auto-correlation method introduced by Gelles *et al.* [115, 123, 125–128]. The diffraction pattern of a bead is cross-correlated with its mirror image, a predefined kernel, or a previous image. The cross-correlation yields a peak, corresponding to the $xy$-position of the bead. At the peak, a radial profile is reconstructed. This profile can be compared to a pre-measured look-up table (LUT) to extract the $z$-position of the bead. The cross-correlation method yields nm accuracy, speed, and simplicity. Nevertheless, without further optimization it can only process a couple of beads in real-time and there is cross-talk between the obtained coordinates in three dimensions. In Chapter 3 we describe a new tracking method, which is lightweight, robust, and therefore optimized for multiplexed experiments.

### 1.6.1 Stretching DNA

To describe the mechanical properties of chromatin we must first have a clear understanding of the mechanical properties of DNA. As physicists, we can simplify the underlying molecular complexity of DNA, and view it as a polymer that is characterized by its bending stiffness. This stiffness is often expressed in terms of the persistence length $P$. Formally, the persistence length of a polymer is the length after which statistical correlations in the direction of the polymer are lost. This implies that a polymer that is much longer than the persistence length will be in a random coil conformation, whereas a polymer shorter than the persistence length can be described by a straight rod. Another way of defining persistence length is the length over which a bend over 1 radian can be made with energy cost $k_{\mathrm{B}}T$ [129]. Since most other biopolymers have a persistence length smaller than 1 nm, double-stranded DNA is considered a semi-flexible polymer with $P = 50$ nm [130]. When a polymer is subjected to a sufficiently high force, its backbone starts to extend beyond its contour length. The extensibility of polymers is expressed in the stretch modulus $S$. This parameter describes the force that would stretch the polymer to twice its contour length, and equals approximately 1000 picoNewton (pN) for DNA in physiological conditions [81, 131, 132]. These two parameters, $P$ and $S$, are sufficient to describe the mechanical response of DNA to force $F$.

Two mathematical models are frequently used to describe the mechanical response of biopolymers to increasing force. To characterize DNA at forces

up to several pN, the freely-jointed chain model (FJC) can be used. The FJC describes a polymer as a chain of $N$ freely jointed monomers of fixed length $b$ (the Kuhn length) which follow random-walk statistics. The contour length of the polymer is $L = bN$, and the average end-to-end distance of the chain is $b\sqrt{N}$. The extension $z_{\mathrm{FJC}}$ is described by

$$z_{\mathrm{FJC}}(F) = L\left(\coth\left(\frac{Fb}{k_{\mathrm{B}}T}\right) - \frac{k_{\mathrm{B}}T}{Fb}\right), \tag{1.1}$$

where the Kuhn length $b$ can be expressed as $P/2$ [133]. For low forces, $Fb \ll k_{\mathrm{B}}T$, the effective spring constant $k$ follows [134]

$$k = \frac{3k_{\mathrm{B}}T}{bL}. \tag{1.2}$$

The FJC, however, does not accurately capture the stretching of DNA when the force is higher [129]. For a more accurate description of a larger force regime, DNA can be modeled with the worm-like chain model (WLC). The WLC, also known as the elastic rod model or the continuous Kratky-Porod model [132], introduces cooperativity between the neighboring monomers and describes the molecule as a flexible rod that curves smoothly as a result of thermal fluctuations. The extension $z_{\mathrm{WLC}}$ is described by

$$z_{\mathrm{WLC}}(F) = L\left(1 - \frac{1}{2}\sqrt{\frac{k_{\mathrm{B}}T}{FP}} + \frac{F}{S}\right). \tag{1.3}$$

The last term is an extension of the WLC and included stretch modulus $S$. The WLC, depicted in Figure 1.10 by the dark dashed line, describes the extension of DNA under forces below 65 pN to a good approximation and will be mostly used throughout this thesis.

When DNA is stretched, it describes a force-extension curve that spans several regimes. The first regime, the linear entropic elasticity regime, takes place below the characteristic force $F = k_{\mathrm{B}}T/P = 0.08$ pN. Here, the molecule behaves as a Hookean spring [129]. The second regime, the non-linear entropic elasticity regime, is observed when this characteristic force is exceeded and the reduction of the conformational entropy defines the extension. The force-extension graph describes a characteristic arc, where several pN is sufficient to induce significant distance between the ends. The third regime, the enthalpic stretching regime, becomes apparent for forces larger than 10 pN, where the extension approaches the contour length. The DNA deforms, causing the extension to exceed the contour length [135, 136].

### 1.6.2 Stretching chromatin

Stretching chromatin is far more complicated than stretching DNA. To illustrate this, imagine stretching a chromatin fiber with 20 nucleosomes. We then need

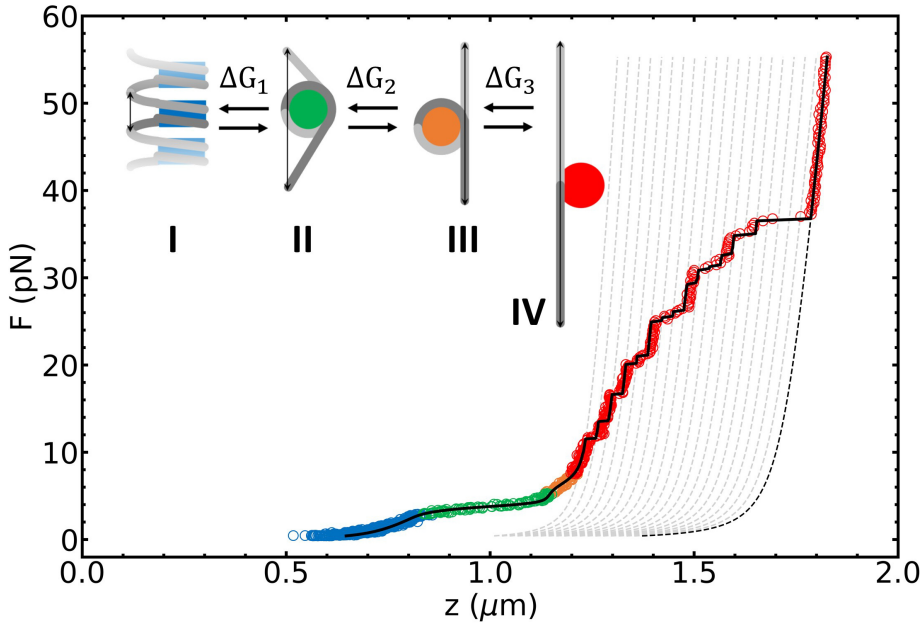



**Figure 1.10**
**Force spectroscopy experiments resolve the higher-order structure of chromatin.** A typical force-extension curve was measured with the multiplexed MT setup, depicting the characteristic unfolding of a chromatin fiber from which the fiber stiffness $k$, rupture energy $\Delta G_1$, and the partial unwrapping energy $\Delta G_2$ could be extracted. Each nucleosome conformation and its corresponding extension is depicted in the inset. The colors of the scatter plot roughly correspond to the different conformations.

to describe the behavior or 162 molecules ($20 \times 8$ histones + 2 DNA strands). Moreover, all histones can detach from the fiber at any given moment. In our experimental approach, the chromatin fiber is furthermore flanked by bare DNA handles. Hence, the total measured extension is the sum of the extension of the handles, dictated by the WLC, and the extension of the fiber, which is composed of the nucleosomes that can be various conformations.

When chromatin fibers unfold by force, their individual nucleosomes progress through a sequence of transitions and conformations, as illustrated in the inset of Figure 1.10. In the case of eukaryotic chromatin, the nucleosome conformations are: (I) the **fully wrapped fiber** conformation, (II) the **partially unwrapped nucleosome** conformation, (III) the **singly wrapped nucleosome** conformation, and (IV) the **fully unwrapped nucleosome** conformation. To complete these transitions, the work done by the stretching force must successively balance the rupture energy $\Delta G_1$, the interaction energy $\Delta G_2$, or the wrapping energy $\Delta G_3$. The first two transitions are in thermodynamic equi-

librium, but the third is not, as illustrated by the stochastic stepwise behavior of the transition.

The force-extension curve of unfolding a complete chromatin fiber with 20 nucleosomes is shown in Figure 1.10. At low forces, the 20 nucleosomes that comprise this fiber are stacked and fully wrapped. In this regime, the fiber stretches linearly with force. Increasing the force causes the nucleosomes to unstack and ultimately unwrap all the nucleosomal DNA until the curve follows the WLC of the DNA tether.

Because there are so many factors that play a role in unfolding chromatin, it becomes increasingly complex to model. This is, however, where statistical physics provides a solution. The total set of nucleosomes, with each nucleosome in a specific conformation, composes a state, featuring its typical extension and free energy. The total extension $z_{\mathrm{tot}}(F)$ and total free energy $g_{\mathrm{tot}}(F)$ of this state are subsequently calculated by adding the extensions and free energies of all nucleosomes that exist in this state. The force-extension curve can then be fitted with the Boltzmann weighted sum of all states:

$$\langle z_{\mathrm{tot}}(F)\rangle = \frac{1}{Z}\sum\nolimits_i z_i(F)\exp(-\frac{g_i(F)+W}{k_{\mathrm{B}}T}), \tag{1.4}$$

where $z_i(F)$ and $g_i(F)$ are the extension and the free energy of nucleosome $i$, and $Z$ describes the partition function $Z = \sum_i \exp\left(-(g_i(F)+W)/k_{\mathrm{B}}T\right)$. As $z_i(F)$ and $g_i(F)$ depend on the mechanical properties of the nucleosomes in the fiber, fitting the equation yields the fiber stiffness $k$, rupture energy $\Delta G_1$, and the interaction energy $\Delta G_2$. In Chapter 5 we have characterized a variety of chromatin fibers, with different linker length, linker sequence, histone variants, and the number of repeats.

The same approach can be followed to describe the extension of other protein-DNA complexes under force. In Chapter 6 we have identified the various conformations of the HMf dimer that compose a hypernucleosome, and developed an analogous statistical mechanics model that revealed stacking energy $g_{\mathrm{stack}}$, and wrapping energy $g_{\mathrm{wrap}}$ of HMfA- and HMfB-DNA complexes.

## 1.7 Rigid base pair Monte Carlo simulations

To complement our force spectroscopy experiments we performed Monte Carlo (MC) simulations of single chromatin fibers. MC simulations are a great way to study complicated stochastic physical systems. By coarse-graining them, we can simulate these systems in experimentally accessible size and time scales, while retaining their important physical properties. These simulations complement the experimental data by providing structural insights that cannot be directly obtained from force spectroscopy experiments alone.

Chromatin fibers have been simulated before in varying degrees of complexity. For instance, De Bruin *et al.* [137], Dobrovolskaia *et al.* [138], and Lequieu

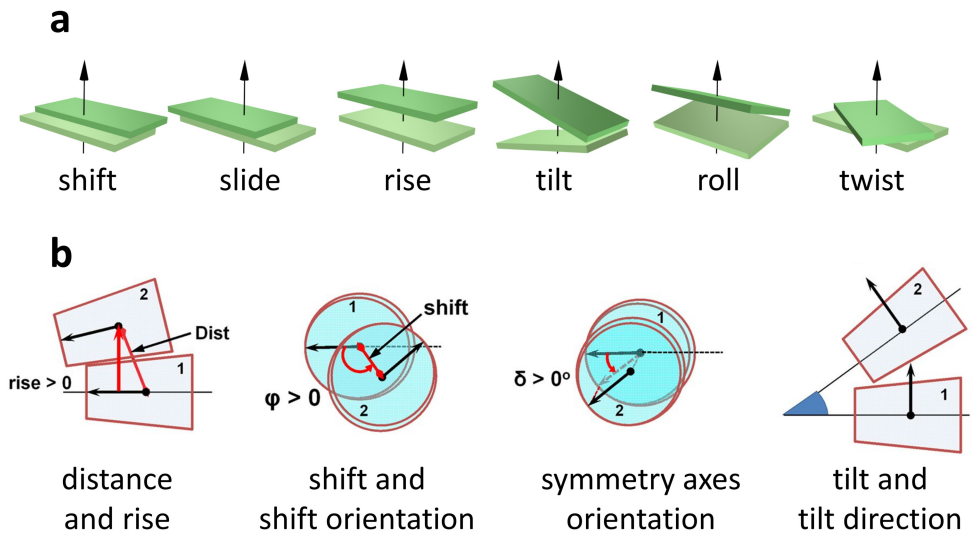



**Figure 1.11**
**The structure of both DNA and chromatin can be coarse-grained into rigid body models.** a) The structure of DNA can be simulated as a chain of rigid base pairs. Each base pair has six degrees-of-freedom with respect to its neighbor: three translational ones (shift, slide, rise), and three rotational ones (tilt, roll, twist). Figure reprinted from [146]. b) The orientation of stacked nucleosomes can also be described by step parameters as described by Korolev *et al.* Figure adapted from [147].

*et al.* [139] detailed the unwrapping of single nucleosomes. Kepper *et al.* included interacting nucleosomes as single units with straight linker DNA [140]. Collepardo-Guevara and Schlick modeled flexible DNA linker between nucleosomes, but, did not allow nucleosomal DNA unwrapping [141, 142]. The recent work of Norouzi and Zhurkin included flexible linker DNA, nucleosomal DNA unwrapping, and interacting nucleosomes, however, their simulations only focused on the zig-zag structure [143–145].

In Chapter 4 of this thesis, we introduce MC simulations of chromatin fibers that include flexible linker DNA, geometrically unconstrained nucleosomes, nucleosome stacking, nucleosomal DNA unwrapping, DNA handles, and various higher-order structures. In this way, we could faithfully simulate force spectroscopy on single chromatin fibers *in silico*.

We based our approach on the rigid base pair model [148]. Rigid base pair Monte Carlo (rbMC) simulations compute DNA trajectories, by treating each base pair as a rigid body that can move with respect to its neighbor. The relation between one base pair and its neighbor is determined by six degrees-of-freedom, which are depicted in Figure 1.11a. The degrees-of-freedom are divided into translational step parameters (shift, slide, rise) and rotational step

parameters (tilt, roll, twist). Six parameters per base pair were sufficient to completely describe the 3D trajectory of a DNA molecule. Specific instances of these step parameters are drawn at random from a distribution that is based on the molecular structure of a B-DNA molecule [149]. We can then probe the conformational space of a DNA molecule by a standard Metropolis-Hastings Monte Carlo algorithm [150].

To simulate chromatin fibers, we needed to include nucleosomes in the rbMC simulations. The orientation of pairs of stacked nucleosomes can be simulated with analogous step parameters used for base pairs. Using step parameters to describe the relative orientation of nucleosomes has been done before, for instance by Korolev *et al.* who used this approach to analyze nucleosome stacking in crystals, which is depicted in Figure 1.11b [147]. In our work, the mean value of each nucleosome step parameters were derived from the global structure of a chromatin fiber, *i.e.* a left-handed fiber with a diameter of 33 nm and a nucleosome line density of 2 nm [79]. The nucleosomal step parameters were subsequently superimposed onto the rbMC simulations. To complete our simulations, energy penalties were added for nucleosomal DNA unwrapping, nucleosome unstacking, surface/bead interaction, and steric hindrance. We used these rbMC simulations to study the effect of linker DNA length on chromatin structure.

## 1.8 Outline of this thesis

In this thesis, the mechanical properties of chromatin fibers are characterized with multiplexed MT. **Chapter 2** details on how MT-based experiments on single chromatin fibers are performed. The chapter includes detailed protocols on experiments and sample preparation, such as DNA cloning and tether assembly in the flow cell. The force spectroscopy workflow and data processing are described, as well as force-calibration of the multiplexed MT apparatus. **Chapter 3** focusses on real-time image analysis. A fast and robust bead tracking algorithm is introduced which is optimized for multiplexed experiments. The algorithm performs with nanometric localization accuracy in a $z$-range of over 10 µm. Its performance under typical image aberrations is evaluated. Its easy implementation, efficiency, and robustness can improve any biophysical technique which relies on bead tracking, especially where high throughput is required. **Chapter 4** introduces the rbMC chromatin model. We extended a rigid base pair DNA model to include (un-)wrapping of nucleosomal DNA and (un-)stacking of nucleosomes in differently folded chromatin fibers. We include flexible linker DNA and DNA handles, and evaluate their effect on the higher-order structure. This model is used to simulate complete force spectroscopy experiments on eukaryotic chromatin fibers *in silico*. **Chapter 5** describes the experiments and simulations of the unfolding pathway of reconstituted

chromatin fibers as a function of single base pair increments in linker DNA length. We analyzed nucleosomes stacking in both 1- and 2-start topologies and show that stacking preferences are indeed determined by linker DNA length. Both computational and experimental results reveal a periodic variation of the folding energies due to the limited flexibility of the linker DNA. We present evidence that the sequence of the linker DNA also affects the mechanical properties of the chromatin fiber. Lastly, we evaluated the significance of H4 tail-mediated stacking interactions in two types of chromatin fibers. In **Chapter 6** we studied the mechanism of DNA compaction in Archaea. We used multiplexed MT to stretch complexes of the *M. fervidus* histones HMfA and HMfB on single DNA tethers. Our experiments demonstrate hypernucleosome assembly which results from the cooperative binding of histones to DNA, facilitated by stacking interactions between neighboring histone dimers. Furthermore, rotational force spectroscopy was used to show that the HMfB-DNA complex has a left-handed chirality, and interestingly, that torque can drive it in a right-handed conformation. We show that the structure of the hypernucleosome thus depends on stacking interactions, torque, and force.

# Bibliography I


[1] Viswanathan Chinnusamy and Jian-Kang Zhu. "Epigenetic regulation of stress responses in plants". In: *Current opinion in plant biology* 12.2 (2009), pp. 133–139.

[2] Silvia Juliana Serrano-Gomez, Mazvita Maziveyi, and Suresh K Alahari. "Regulation of epithelial-mesenchymal transition through epigenetic and post-translational modifications". In: *Molecular cancer* 15.1 (2016), p. 18.

[3] Simone Sidoli, Lei Cheng, and Ole N Jensen. "Proteomics in chromatin biology and epigenetics: Elucidation of post-translational modifications of histone proteins by mass spectrometry". In: *Journal of proteomics* 75.12 (2012), pp. 3419–3433.

[4] Alfred J Robison and Eric J Nestler. "Transcriptional and epigenetic mechanisms of addiction". In: *Nature reviews neuroscience* 12.11 (2011), pp. 623–637.

[5] Thu-Hang Pham et al. "Dynamic epigenetic enhancer signatures reveal key transcription factors associated with monocytic differentiation states". In: *Blood, The Journal of the American Society of Hematology* 119.24 (2012), e161–e171.

[6] Lai Wei et al. "Discrete roles of STAT4 and STAT6 transcription factors in tuning epigenetic modifications and transcription during T helper cell differentiation". In: *Immunity* 32.6 (2010), pp. 840–851.

[7] ER Gibney and CM Nolan. "Epigenetics and gene expression". In: *Heredity* 105.1 (2010), pp. 4–13.

[8] Diane E Handy, Rita Castro, and Joseph Loscalzo. "Epigenetic modifications: basic mechanisms and role in cardiovascular disease". In: *Circulation* 123.19 (2011), pp. 2145–2156.

[9] Charlotte Ling and Leif Groop. "Epigenetics: a molecular link between environmental factors and type 2 diabetes". In: *Diabetes* 58.12 (2009), pp. 2718–2725.

[10] Michael K Skinner, Mohan Manikkam, and Carlos Guerrero-Bosagna. "Epigenetic transgenerational actions of environmental factors in disease etiology". In: *Trends in Endocrinology & Metabolism* 21.4 (2010), pp. 214–222.

[11] John C Mathers, Gordon Strathdee, and Caroline L Relton. "Induction of epigenetic alterations by dietary and other environmental factors". In: *Advances in genetics*. Vol. 71. Elsevier, 2010, pp. 3–39.

[12] Richard P. (Richard Phillips) Feynman. *The pleasure of finding things out: the best short works of Richard P. Feynman*. Perseus Books, 1999.

[13] Charles Darwin. *On the origin of species*. 1859.

[14] Gregor Mendel. "Versuche über Plflanzenhybriden". In: *Verhandlungen des naturforschenden Vereines in Brünn* Bd. IV (1866), pp. 3–47.

[15] Francesca Mattiroli et al. "Structure of histone-based chromatin in Archaea". In: *Science* 357.6351 (2017), pp. 609–612. ISSN: 0036-8075. DOI: 10.1126/science.aaj1849.

[16] B Alberts et al. *Molecular biology of the cell, 6th ed.* NY: WW Norton & Company, 2014.

[17] William E Balch et al. "An ancient divergence among the bacteria". In: *Journal of Molecular Evolution* 9.4 (1977), pp. 305–311.

[18] George E Fox et al. "Classification of methanogenic bacteria by 16S ribosomal RNA characterization". In: *Proceedings of the National Academy of Sciences of the United States of America* 74.10 (1977), p. 4537.

[19] Carl R Woese and George E Fox. "Phylogenetic structure of the prokaryotic domain: the primary kingdoms". In: *Proceedings of the National Academy of Sciences* 74.11 (1977), pp. 5088–5090.

[20] Leslie A. Pray. "Discovery of DNA Structure and Function: Watson and Crick". In: *Nature Education* 1(1):100 (2008).

[21] PA Levene and WA Jacobs. "Über Inosinsäure". In: *Berichte der Deutschen Chemischen Gesellschaft* 42.1 (1909), pp. 1198–1203.

[22] D Elson and E Chargaff. "On the desoxyribonucleic acid content of sea urchin gametes". In: *Experientia* 8.4 (1952), pp. 143–145.

[23] Erwin Chargaff, Rakoma Lipshitz, and Charlotte Green. "Composition of the desoxypentose nucleic acids of four genera of sea-urchin". In: *J Biol Chem* 195.1 (1952), pp. 155–160.

[24] James D Watson and Francis HC Crick. "The structure of DNA". In: *Cold Spring Harbor symposia on quantitative biology*. Vol. 18. Cold Spring Harbor Laboratory Press. 1953, pp. 123–131.

[25] Rosalind E Franklin and Raymond George Gosling. "The structure of sodium thymonucleate fibres. I. The influence of water content". In: *Acta Crystallographica* 6.8-9 (1953), pp. 673–677.

[26] Rosalind E Franklin and RG Gosling. "The structure of sodium thymonucleate fibres. II. The cylindrically symmetrical Patterson function". In: *Acta Crystallographica* 6.8-9 (1953), pp. 678–685.

[27] Maurice Hugh Frederick Wilkins, Alex R Stokes, and Herbert R Wilson. "Molecular structure of nucleic acids: molecular structure of deoxypentose nucleic acids". In: *Nature* 171.4356 (1953), pp. 738–740.

[28] MH Wilkins and JT Randall. "Crystallinity in sperm heads: molecular structure of nucleoprotein in vivo." In: *Biochimica et biophysica acta* 10.1 (1953), p. 192.

[29] Linus Pauling and Robert B Corey. "The pleated sheet, a new layer configuration of polypeptide chains". In: *Proceedings of the National Academy of Sciences of the United States of America* 37.5 (1951), p. 251.

[30] Linus Pauling, Robert B Corey, and Herman R Branson. "The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain". In: *Proceedings of the National Academy of Sciences* 37.4 (1951), pp. 205–211.

[31] Linus Pauling and Robert B Corey. "Configurations of polypeptide chains with favored orientations around single bonds: two new pleated sheets". In: *Proceedings of the National Academy of Sciences of the United States of America* 37.11 (1951), p. 729.

[32] Philip Nelson. *Biological physics*. WH Freeman New York, 2004.

[33] Ian R Gould and Peter A Kollman. "Theoretical investigation of the hydrogen bond strengths in guanine-cytosine and adenine-thymine base pairs". In: *Journal of the American Chemical Society* 116.6 (1994), pp. 2493–2499.

[34] Francis Crick. "Central dogma of molecular biology". In: *Nature* 227.5258 (1970), p. 561.

[35] Howard M Temin, S Mizutami, et al. "RNA-dependent DNA polymerase in virions of Rous sarcoma virus." In: *Nature* 226 (1970), pp. 1211–1213.

[36] Minoru Takata et al. "Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells". In: *The EMBO journal* 17.18 (1998), pp. 5497–5508.

[37] Chance Meers, Havva Keskin, and Francesca Storici. "DNA repair by RNA: Templated, or not templated, that is the question". In: *DNA repair* 44 (2016), pp. 17–21.

[38] David Baltimore. "Retroviruses and retrotransposons: the role of reverse transcription in shaping the eukaryotic genome". In: *Cell* 40 (1985), pp. 481–482.

[39] Matthew Cobb. "60 years ago, Francis Crick changed the logic of biology". In: *PLoS biology* 15.9 (2017), e2003243.

[40] Patrick D Hsu, Eric S Lander, and Feng Zhang. "Development and applications of CRISPR-Cas9 for genome engineering". In: *Cell* 157.6 (2014), pp. 1262–1278.

[41] Le Cong et al. "Multiplex genome engineering using CRISPR/Cas systems". In: *Science* 339.6121 (2013), pp. 819–823.

[42] Prashant Mali et al. "RNA-guided human genome engineering via Cas9". In: *Science* 339.6121 (2013), pp. 823–826.

[43] Gregory J Hannon. "RNA interference". In: *nature* 418.6894 (2002), p. 244.

[44] Adrian Bird. "Perceptions of epigenetics". In: *Nature* 447.7143 (2007), p. 396.

[45] Luis C Junqueira and Anthony L Mescher. *Junqueira's basic histology*. New York: McGraw-Hill Medical, 2013.

[46] Nick Gilbert et al. "Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers". In: *Cell* 118.5 (2004), pp. 555–566.

[47] John L Rinn et al. "Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs". In: *cell* 129.7 (2007), pp. 1311–1323.

[48] MA Kelly et al. “Molecular aspects of type 1 diabetes”. In: *Molecular Pathology* 56.1 (2003), p. 1.

[49] Samuel T Jerram, Mary N Dang, and R David Leslie. “The role of epigenetics in type 1 diabetes”. In: *Current diabetes reports* 17.10 (2017), p. 89.

[50] Anders H Olsson et al. “Genome-wide associations between genetic and epigenetic variation influence mRNA expression and insulin secretion in human pancreatic islets”. In: *PLoS genetics* 10.11 (2014), e1004735.

[51] Tamara L Caterino and Jeffrey J Hayes. “Chromatin structure depends on what’s in the nucleosome’s pocket”. In: *Nature structural & molecular biology* 14.11 (2007), p. 1056.

[52] Sergei A Grigoryev. “Chromatin higher-order folding: a perspective with linker DNA angles”. In: *Biophysical journal* 114.10 (2018), pp. 2290–2297.

[53] Alison E White, Aaron R Hieb, and Karolin Luger. “A quantitative investigation of linker histone interactions with nucleosomes and chromatin.” In: *Scientific reports* 6 (2016), p. 19122. ISSN: 2045-2322. DOI: 10.1038/srep19122.

[54] Vladislav Levchenko, Beverly Jackson, and Vaughn Jackson. “Histone release during transcription: displacement of the two H2A- H2B dimers in the nucleosome is dependent on different levels of transcription-induced positive stress”. In: *Biochemistry* 44.14 (2005), pp. 5357–5372.

[55] Orkide Ordu, Alexandra Lusser, and Nynke H Dekker. “Recent insights from in vitro single-molecule studies into nucleosome structure and dynamics”. In: *Biophysical reviews* 8.1 (2016), pp. 33–49.

[56] Daniela Rhodes. “Chromatin structure: The nucleosome core all wrapped up”. In: *Nature* 389.6648 (1997), p. 231.

[57] 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.

[58] A L Olins and Donald E Olins. “Spheroid chromatin units (v bodies).” In: *Science (New York, N.Y.)* 183.4122 (1974), pp. 330–2. ISSN: 0036-8075. DOI: 10.1126/science.183.4122.330.

[59] Roger D Kornberg. “Chromatin structure: a repeating unit of histones and DNA”. In: *Science* 184.4139 (1974), pp. 868–871.

[60] Timothy J Richmond and Curt A Davey. “The structure of DNA in the nucleosome core”. In: *Nature* 423.6936 (2003), p. 145.

[61] Hiroaki Tachiwana et al. “Crystal structure of the human centromeric nucleosome containing CENP-A”. In: *Nature* 476.7359 (2011), p. 232.

[62] Tomoya Kujirai et al. “Structure and function of human histone H3.Y nucleosome”. In: *Nucleic Acids Research* 44.13 (2016), pp. 6127–6141. ISSN: 13624962. DOI: 10.1093/nar/gkw202.

[63] Dorine Rossetto, Nikita Avvakumov, and Jacques Côté. “Histone phosphorylation: a chromatin modification involved in diverse nuclear events”. In: *Epigenetics* 7.10 (2012), pp. 1098–1108.

[64] Marissa Vignali et al. “ATP-dependent chromatin-remodeling complexes”. In: *Molecular and cellular biology* 20.6 (2000), pp. 1899–1910.

[65] Leanne De Koning et al. “Histone chaperones: an escort network regulating histone traffic”. In: *Nature structural & molecular biology* 14.11 (2007), p. 997.

[66] Ali Shilatifard. “Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression”. In: *Annu. Rev. Biochem.* 75 (2006), pp. 243–269.

[67] Zhenhai Li and Hidetoshi Kono. “Distinct roles of histone H3 and H2A tails in nucleosome stability”. In: *Scientific reports* 6 (2016), p. 31437.

[68] Gaurav Arya and Tamar Schlick. “Role of histone tails in chromatin folding revealed by a mesoscopic oligonucleosome model”. In: *Proceedings of the National Academy of Sciences* 103.44 (2006), pp. 16236–16241.

[69] 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.

[70] 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.

[71] 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.

[72] Benedetta Dorigo et al. “Nucleosome arrays reveal the two-start organization of the chromatin fiber.” In: *Science (New York, N.Y.)* 306.5701 (Nov. 2004), pp. 1571–3. ISSN: 1095-9203. DOI: 10.1126/science.1103124.

[73] Andrew Bassett et al. “The folding and unfolding of eukaryotic chromatin”. In: *Current opinion in genetics & development* 19.2 (2009), pp. 159–165.

[74] Gungor Ozer, Antoni Luque, and Tamar Schlick. “The chromatin fiber: Multiscale problems and approaches”. In: *Current Opinion in Structural Biology* 31 (2015), pp. 124–139. ISSN: 1879033X. DOI: 10.1016/j.sbi.2015.04.002.

[75] F Thoma, Th Koller, and A Klug. “Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin.” In: *The Journal of cell biology* 83.2 (1979), pp. 403–427.

[76] C. L F Woodcock, L. L Y Frado, and J. B. Rattner. “The higher-order structure of chromatin: Evidence for a helical ribbon arrangement”. In: *Journal of Cell Biology* 99.1 I (1984), pp. 42–52. ISSN: 00219525. DOI: 10.1083/jcb.99.1.42.

[77] 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.

[78] Philip JJ Robinson and Daniela Rhodes. “Structure of the ’30 nm’ chromatin fibre: A key role for the linker histone”. In: *Current Opinion in Structural Biology* 16.3 (2006), pp. 336–343. ISSN: 0959440X. DOI: 10.1016/j.sbi.2006.05.007.

[79] 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.

[80] 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.

[81] 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.

[82] 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.

[83] Jacques Dubochet and Markus Noll. “Nucleosome arcs and helices”. In: *Science* 202.4365 (1978), pp. 280–286.

[84] P T T Lowary and J Widom. “New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning.” In: *Journal of molecular biology* 276.1 (1998), pp. 19–42. ISSN: 0022-2836. DOI: 10.1006/jmbi.1997.1494. eprint: NIHMS150003.

[85] Sarah J Correll, Michaela H Schubert, and Sergei A Grigoryev. “Short nucleosome repeats impose rotational modulations on chromatin fibre folding”. In: *The EMBO Journal* 31.10 (2012), pp. 2416–2426. ISSN: 1460-2075. DOI: 10.1038/emboj.2012.80.

[86] Sergei A Grigoryev et al. “Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions”. In: *Proceedings of the National Academy of Sciences* 106.32 (2009), pp. 13317–13322.

[87] CL Woodcock et al. “A chromatin folding model that incorporates linker variability generates fibers resembling the native structures”. In: *Proceedings of the National Academy of Sciences* 90.19 (1993), pp. 9021–9025.

[88] J Widom. “A relationship between the helical twist of DNA and the ordered positioning of nucleosomes in all eukaryotic cells.” In: *Proceedings of the National Academy of Sciences of the United States of America* 89.February (1992), pp. 1095–1099. ISSN: 0027-8424. DOI: 10.1073/pnas.89.3.1095.

[89] Anton Valouev et al. “Determinants of nucleosome organization in primary human cells”. In: *Nature* 474.7352 (2011), p. 516.

[90] Kristin Brogaard et al. “A map of nucleosome positions in yeast at base-pair resolution”. In: *Nature* 486.7404 (2012), p. 496.

[91] Ji-Ping Wang et al. “Preferentially quantized linker DNA lengths in Saccharomyces cerevisiae”. In: *PLoS computational biology* 4.9 (2008), e1000175.

[92] Tsung-Han S Hsieh et al. “Mapping nucleosome resolution chromosome folding in yeast by micro-C”. In: *Cell* 162.1 (2015), pp. 108–119.

[93] Nikolay Korolev, Lars Nordenskiöld, and Alexander P. Lyubartsev. "Multiscale coarse-grained modelling of chromatin components: DNA and the nucleosome". In: *Advances in Colloid and Interface Science* (2016), pp. 1–13. ISSN: 00018686. DOI: 10.1016/j.cis.2016.02.002.

[94] Kazuhiro Maeshima, Saera Hihara, and Mikhail Eltsov. "Chromatin structure: does the 30-nm fibre exist in vivo?" In: *Current opinion in cell biology* 22.3 (2010), pp. 291–297.

[95] Joan-Ramon Daban. "Electron microscopy and atomic force microscopy studies of chromatin and metaphase chromosome structure". In: *Micron* 42.8 (2011), pp. 733–750.

[96] 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.

[97] Viviana I Risca et al. "Variable chromatin structure revealed by in situ spatially correlated DNA cleavage mapping". In: *Nature* 541.7636 (2017), p. 237.

[98] Beat Fierz. "Dynamic chromatin regulation from a single molecule perspective". In: *ACS chemical biology* 11.3 (2015), pp. 609–620.

[99] Beat Fierz and Michael G Poirier. "Biophysics of Chromatin Dynamics". In: *Annual review of biophysics* 48 (2019), pp. 321–345.

[100] Suhas SP Rao et al. "A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping". In: *Cell* 159.7 (2014), pp. 1665–1680.

[101] Jesse R Dixon et al. "Topological domains in mammalian genomes identified by analysis of chromatin interactions". In: *Nature* 485.7398 (2012), p. 376.

[102] Ana Pombo and Niall Dillon. "Three-dimensional genome architecture: players and mechanisms". In: *Nature reviews Molecular cell biology* 16.4 (2015), p. 245.

[103] Matthias Merkenschlager and Elphege P Nora. "CTCF and cohesin in genome folding and transcriptional gene regulation". In: *Annual review of genomics and human genetics* 17 (2016), pp. 17–43.

[104] Johannes Nuebler et al. "Chromatin organization by an interplay of loop extrusion and compartmental segregation". In: *Proceedings of the National Academy of Sciences* 115.29 (2018), E6697–E6706.

[105] A Nagy K. Neuman, Keir C Neuman, and Attila Nagy. "Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy". In: *Nature methods* 5.6 (2008), pp. 491–505. ISSN: 1548-7091. DOI: 10.1038/NMETH.1218. eprint: NIHMS150003.

[106] Gerrit Sitters et al. "Acoustic force spectroscopy." In: *Nature methods* 12.1 (2015), pp. 47–50. ISSN: 1548-7105. DOI: 10.1038/nmeth.3183.

[107] Yujia Cui and Carlos Bustamante. "Pulling a single chromatin fiber reveals the forces that maintain its higher-order structure". In: *Proceedings of the National Academy of Sciences* 97.1 (2000), pp. 127–132.

[108] Martin L Bennink et al. "Unfolding individual nucleosomes by stretching single chromatin fibers with optical tweezers". In: *Nature Structural & Molecular Biology* 8.7 (2001), p. 606.

[109] Brent D Brower-Toland et al. “Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA.” In: *Proceedings of the National Academy of Sciences of the United States of America* 99.4 (2002), pp. 1960–5. ISSN: 0027-8424. DOI: 10.1073/pnas.022638399.

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

[111] Nicolaas Hermans et al. “Toehold-enhanced LNA probes for selective pull down and single-molecule analysis of native chromatin”. In: *Scientific reports* 7.1 (2017), p. 16721.

[112] Artur Kaczmarczyk et al. “Probing Chromatin Structure with Magnetic Tweezers”. In: *Nanoscale Imaging*. Springer, 2018, pp. 297–323.

[113] ID Vilfan et al. “Magnetic tweezers for single-molecule experiments”. In: *Handbook of Single-Molecule Biophysics*. Springer, 2009, pp. 371–395.

[114] He Meng et al. “Coexistence of Twisted, Plectonemic, and Melted DNA in Small Topological Domains”. In: *Biophysical Journal* 106.5 (2014), pp. 1174–1181. ISSN: 00063495. DOI: 10.1016/j.bpj.2014.01.017.

[115] T. R. Strick et al. “The elasticity of a single supercoiled DNA molecule.” In: *Science (New York, N.Y.)* 271.5257 (1996), pp. 1835–1837. ISSN: 0036-8075. DOI: 10.1126/science.271.5257.1835.

[116] John F. Marko. “Torque and dynamics of linking number relaxation in stretched supercoiled DNA”. In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 76.2 (2007), pp. 1–13. ISSN: 15393755. DOI: 10.1103/PhysRevE.76.021926.

[117] Richik N Ghosh and Watt W Webb. “Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules”. In: *Biophysical journal* 66.5 (1994), pp. 1301–1318.

[118] G M Lee, A Ishihara, and K Jacobson. “Direct observation of brownian motion of lipids in a membrane.” In: *Proceedings of the National Academy of Sciences of the United States of America* 88.14 (1991), pp. 6274–8. ISSN: 0027-8424. DOI: 10.1073/pnas.88.14.6274.

[119] Lisa Gottesfeld Brown. “A survey of image registration techniques”. In: *ACM computing surveys (CSUR)* 24.4 (1992), pp. 325–376.

[120] Catherine M Anderson et al. “Tracking of cell surface receptors by fluorescence digital imaging microscopy using a charge-coupled device camera. Low-density lipoprotein and influenza virus receptor mobility at 4 degrees C.” In: *Journal of cell science* 101 ( Pt 2 (1992), pp. 415–425. ISSN: 0021-9533.

[121] Gerhard J Schütz, Hansgeorg Schindler, and Thomas Schmidt. “Single-molecule microscopy on model membranes reveals anomalous diffusion”. In: *Biophysical journal* 73.2 (1997), pp. 1073–1080.

[122] Andrew J Berglund et al. “Fast, bias-free algorithm for tracking single particles with variable size and shape”. In: *Optics express* 16.18 (2008), pp. 14064–14075.

[123] M K Cheezum, W F Walker, and W H Guilford. "Quantitative comparison of algorithms for tracking single fluorescent particles." In: *Biophysical journal* 81.4 (2001), pp. 2378–2388. ISSN: 00063495. DOI: 10.1016/S0006-3495(01)75884-5.

[124] Sang-Hyuk Lee et al. "Characterizing and tracking single colloidal particles with video holographic microscopy." In: *Optics express* 15.26 (2007), pp. 18275–18282. ISSN: 1094-4087. DOI: 10.1364/OE.15.018275. arXiv: 0712.1738.

[125] Jeff Gelles, Bruce J. Schnapp, and Michael P. Sheetz. "Tracking kinesin-driven movements with nanometre-scale precision." In: *Nature* 331.6155 (1988), pp. 450–453. ISSN: 0028-0836. DOI: 10.1038/331450a0. arXiv: arXiv:1011.1669v3.

[126] Akihiro Kusumi, Yasushi Sako, and Mutsuya Yamamoto. "Confined lateral diffusion of membrane receptors as studied by single particle tracking (nanovid microscopy). Effects of calcium-induced differentiation in cultured epithelial cells". In: *Biophysical journal* 65.5 (1993), pp. 2021–2040.

[127] Charlie Gosse and Vincent Croquette. "Magnetic tweezers: micromanipulation and force measurement at the molecular level." In: *Biophysical Journal* 82.6 (2002), pp. 3314–3329. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(02)75672-5.

[128] Astrid van der Horst and Nancy R Forde. "Calibration of dynamic holographic optical tweezers for force measurements on biomaterials." In: *Optics express* 16.25 (Dec. 2008), pp. 20987–1003. ISSN: 1094-4087.

[129] John F. Marko and Eric D. Siggia. "Stretching DNA". In: *Macromolecules* 28.26 (Dec. 1995), pp. 8759–8770. ISSN: 15205835. DOI: 10.1021/ma00130a008.

[130] C Bustamante et al. "Single-molecule studies of DNA mechanics". In: *Current opinion in structural biology* 10.3 (June 2000), pp. 279–85. ISSN: 0959-440X. DOI: 10.1016/S0959-440X(00)00085-3.

[131] Michelle D Wang et al. "Stretching DNA with Optical Tweezers". In: *Biophysical Journal* 72.March (1997), pp. 1335–1346.

[132] Carlos Bustamante, Zev Bryant, and Steven B. Smith. "Ten years of tension: single-molecule DNA mechanics". In: *Nature* 421.6921 (2003), pp. 423–427. ISSN: 0028-0836. DOI: 10.1038/nature01405.

[133] Yih-Fan Chen et al. "Entropic boundary effects on the elasticity of short DNA molecules". In: *Physical Review E* 80.2 (2009), p. 020903.

[134] TR Strick et al. "Physical approaches to the study of DNA". In: *Journal of statistical physics* 93.3-4 (1998), pp. 647–672.

[135] John F Marko and Simona Cocco. "The micromechanics of DNA". In: *Physics World* 16.3 (2003), p. 37.

[136] Theo Odijk. "Stiff chains and filaments under tension". In: *Macromolecules* 28.20 (1995), pp. 7016–7018.

[137] Lennart De Bruin et al. "Why do nucleosomes unwrap asymmetrically?" In: *Journal of Physical Chemistry B* 120.26 (2016), pp. 5855–5863. ISSN: 15205207. DOI: 10.1021/acs.jpcb.6b00391.

[138] Irina V Dobrovolskaia and Gaurav Arya. "Dynamics of forced nucleosome unraveling and role of nonuniform histone-DNA interactions". In: *Biophysical journal* 103.5 (2012), pp. 989–998.

[139] Joshua Lequieu et al. "Tension-dependent free energies of nucleosome unwrapping". In: *ACS central science* 2.9 (2016), pp. 660–666.

[140] 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.

[141] Rosana Collepardo-Guevara and Tamar Schlick. "Crucial role of dynamic linker histone binding and divalent ions for DNA accessibility and gene regulation revealed by mesoscale modeling of oligonucleosomes". In: *Nucleic acids research* 40.18 (2012), pp. 8803–8817.

[142] 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.

[143] Mikhail V Bass et al. "Nucleosome spacing periodically modulates nucleosome chain folding and DNA topology in circular nucleosome arrays". In: *Journal of Biological Chemistry* 294.11 (2019), pp. 4233–4246.

[144] Davood Norouzi and Victor B Zhurkin. "Dynamics of chromatin fibers: Comparison of Monte Carlo simulations with force spectroscopy". In: *Biophysical journal* 115.9 (2018), pp. 1644–1655.

[145] Davood Norouzi and Victor B. Zhurkin. "Topological polymorphism of the two-start chromatin fiber". In: *Biophysical Journal* 108.10 (2015), pp. 2591–2600. ISSN: 15420086. DOI: 10.1016/j.bpj.2015.04.015.

[146] Marco Tompitak. "The mechanical genome: inquiries into the mechanical function of genetic information". PhD thesis. Leiden University, 2017.

[147] 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.

[148] Wilma K Olson et al. "A standard reference frame for the description of nucleic acid base-pair geometry". In: *Journal of molecular biology* 313.1 (2001), pp. 229–237.

[149] 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.

[150] W Keith Hastings. "Monte Carlo sampling methods using Markov chains and their applications". In: (1970).

# Chapter 2

# Methods for Unravelling DNA organization with single-molecule Force Spectroscopy using Magnetic Tweezers

Genomes carry the genetic blueprint of all living organisms. Their organization requires strong condensation as well as carefully regulated accessibility to specific genes for proper functioning of their hosts. The study of the structure and dynamics of the proteins that organize the genome has benefited tremendously from the development of single-molecule force spectroscopy techniques that allow for real-time, nanometer accuracy measurements of the compaction of DNA and manipulation with pico-Newton scale forces. Magnetic tweezers in particular have the unique ability to complement such force spectroscopy with the control over the linking number of the DNA molecule, which plays an important role when DNA organizing proteins form or release wraps, loops and bends in DNA. Here, we describe all the necessary steps to prepare DNA substrates for magnetic tweezers experiments, assemble flow cells, tether DNA to a magnetic bead inside flow cell and manipulate and record the extension of such DNA tethers. Furthermore, we explain how mechanical parameters of nucleo-protein filaments can be extracted from the data.

This chapter is based on: Brouwer T. B., Kaczmarczyk A., Pham C., van Noort S. J. T.: Unravelling DNA organization with single-molecule Force Spectroscopy using Magnetic Tweezers, **Bacterial Chromatin**, Humana Press, New York, NY, 2018: 317-349.

# 2.1 Introduction

## 2.1.1 Chromatin organization by DNA-binding proteins

Genomes carry the genetic blueprint of all living organisms. Their organization requires strong condensation in a structure called chromatin, as well as carefully regulated accessibility to genes in chromatin for proper functioning of their hosts. The organization of chromatin is carried out by a large group of proteins that vary in abundance, structure and function. Different kingdoms of life have evolved to implement different solutions for a compact, but dynamic organization of chromatin. Nonetheless, many of the involved proteins share common mechanisms to deform the DNA by inducing kinks, bends, wraps and loops in the DNA trajectory. Examples of common DNA organizing proteins are histone proteins found in eukaryotes and in Archaea and smaller and more diverse proteins like HU, integration host factor IHF and histone-like nucleoid-structuring protein H-NS in prokaryotes [1].

For a mechanistic understanding of the function of these proteins, required for a robust comprehension of chromatin-associated processes such as transcription, a detailed knowledge of the kinetics of DNA binding, as well as the mechanical consequences of their binding to DNA is needed. In fact, there is a tight relationship between DNA conformation and its interaction with proteins: wrapping of DNA around proteins results in condensation of the DNA, both by bringing together the ends of the DNA and by inducing supercoiling in the surrounding DNA, which results in additional condensation by plectoneme formation. DNA supercoiling, induced by externally applied torque, will facilitate the binding of proteins that wrap or bend the DNA with similar chirality. Force and extension play similar related roles: some proteins induce stretching of DNA upon binding. Pre-stretching of DNA can facilitate binding or function of these proteins, making force an important parameter in the study of nucleo-protein filaments [2].

DNA forms the common thread in chromatin organization, and it is informative to gather and compare mechano-chemical experimental data on chromatin organization by different proteins and from different species, to gather a comprehensive picture of DNA organization. In terms of DNA mechanobiology, as well as from an evolutionary perspective, it is highly desirable to establish a physical understanding of chromatin organization throughout the kingdoms of life. In this chapter, we present protocols to set up single-molecule force spectroscopy experiments on DNA tethers interacting with a variety of chromatin proteins. Although some examples stem from studies of eukaryotic proteins, it is our experience that the methods for sample preparation, execution of the experiments and data analysis are similar, and we provide protocols that can be used to study bacterial chromatin, as well as other nucleoprotein complexes.

### 2.1.2 Magnetic tweezers

Magnetic Tweezers (MT) can probe the mechanical response of individual biomolecules to applied force and twist. The first MT experiments have been instrumental in understanding the mechanical behavior of bare DNA [3–5]. Since then, mechanical and structural properties of more complex biomolecules and DNA-protein complexes have been studied [6–11]. MT are especially suitable to this end since they require limited hardware investments, uniquely control both force and linking number, and have a relatively high throughput compared to optical tweezers, the most common alternative single-molecule force spectroscopy technique. Recent developments are an increase in maximum force to beyond the overstretching force for DNA [8, 12, 13], multiplexing capabilities, such that hundreds of molecules can be measured in parallel [14, 15] and freely orbiting magnetic tweezers (FOMT) [16] or magnetic torque tweezers (MTT) [17], that allow for (partial) relief of twist or direct torque measurement. Another advance is the replacement of permanent magnets by electromagnets in electromagnetic tweezers [5, 18]. Since there are no moving parts in electromagnetic tweezers, the only factor limiting the speed of the experiments is the viscous drag of the magnetic beads themselves. Finally, combinations with (single-molecule) fluorescence have been reported [19–22]. These modern variations of the technique require similar sample preparation protocols and can be seen as extensions of typical MT experiments.

A typical MT experiment involves a single DNA molecule tethered between a paramagnetic bead and a functionalized glass slide. The three-dimensional position of the bead is recorded in time by video microscopy and extracted by image processing. The bead can be manipulated by translation or rotation of a pair of magnets that are held above the tether. The force and torque that are exerted on the paramagnetic bead result in a restoring force and torque in the tethered molecule, which provides insight into the structure of the tether.

### 2.1.3 Measurement schemes

MT can be operated in several measurement schemes. Traditionally, the magnet is fixed in a preset position during a measurement, resulting in a constant force during the experiment. For bare DNA, such a constant force measurement not only accurately reveals the extension of the tether, one can also directly calculate the force by quantification of the lateral movements of the bead [4]. Application of the equipartition theorem yields that the force must equal the product of thermal energy and extension, divided by the variance of the lateral fluctuations (*see* section 3.7). Repeated measurements at different magnet positions yield a calibration of the force as a function of the magnet position [4, 23–26], which typically follows a mono- or bi-exponential decay [25]. The same measurement also provides an accurate map of the force-extension relation

of DNA, which can be described by a Worm-like Chain (WLC) model. For transient protein-DNA interactions, a constant force experiment can directly reveal changes in extension of the tether upon flushing the protein solution in or out [27–29].

Having established a force calibration relation, one can revert to dynamic force spectroscopy, in which the extension of the tether is monitored during the repositioning of the magnet [30]. Such a measurement scheme aids in revealing the force dependence of structures that may rupture or dissociate when increasing forces are applied. For instance, we have noticed that histone proteins readily dissociate when chromatin fibers are exposed to forces exceeding several pN [11].

Next to shifting the position of the magnets with respect to the tethered bead, the magnets can be rotated around the axis of the tether. Since the torque on the bead is proportional to the strength of the magnetic field (*see* **Note 1**), and is typically much larger than the opposing torque provided by the DNA tether, the bead will immediately follow the rotation of the magnets. This can lead to accumulation of twist in the tether, if the twist is not relieved by swiveling around the bond between the tether and bead or surface, or by swiveling around a nick in the DNA. Constraining the rotation of the DNA ends requires more than one chemical bond between the DNA strand and the bead or surface, which can be achieved by introducing multiple affinity tags in the DNA ends. Such a DNA substrate allows for direct control of the linking number of the topological domain formed by the bead and the surface. Here, we describe the preparation of both torsionally constrained and torsionally free DNA substrates, allowing for a broad spectrum of applications.

Controlling the linking number of the DNA tether opens up a range of measurement schemes. At constant twist, force-extension measurements can be done in a similar fashion as with torsionally unconstrained DNA. Any twist-stretch coupling will be evident in the force-extension curve. Protein-induced (un-)wrapping will accumulate, which may affect the binding of subsequent proteins. A torsionally constrained tether can be pre-twisted before force spectroscopy experiments, to probe the chiral properties of the tether over a wide range of linking numbers [31]. The mechanical response of the tether to torque can also be measured directly by twisting the tether under constant force. In conclusion, a wide variety of single-molecule manipulation schemes is available with MT force spectroscopy.

### 2.1.4 Data analysis

The mechanical properties of DNA provide a reference for data analysis of nucleo-protein filaments. The extension of DNA as a function of force can be described accurately by the WLC model [32]. At forces larger than 10 pN, the extensible WLC model provides a better fit [32, 33]. DNA extension

as a function of twist requires a more elaborate analysis, using numerical methods, due to the coexistence of three conformations, *i.e.* twisted, melted and plectonemic DNA [31]. However, for force and twist regimes where only two of these states prevail, analytical solutions exist [34]. In general, it is important to make sure that the force-extension curves of bare DNA, under conditions relevant for studying chromatin, fit well to these models, prior to the study of more advanced structures, as protein-induced changes in extension can be difficult to interpret without proper reference.

Wrapping, bending, twisting or deforming DNA otherwise, will result in a change in extension of the tether that may depend on force and twist. In a force-extension curve, this may lead to a change in persistence length, contour length, stretch modulus and/or twist persistence length. In the data analysis part of the methods section, we review some models that can be fitted to extract these parameters, as well as more dynamic methods to extract the kinetics of protein binding. Below, we describe in detail how to make DNA substrates suitable for MT. Note that many variations of DNA substrates exist that the reader can adapt to specific needs. We report a detailed protocol for the assembly of DNA tethered beads in a flow cell, ready for measurements, and share some typical procedures for data analysis. These should provide a solid foundation to study chromatin at the single-molecule level with MT.

## 2.2 Materials

### 2.2.1 Stock solutions

All aqueous buffers are prepared with Milli-Q water and stored at 4°C.

- *Measurement buffer*: 10 mM HEPES pH 7.5, 100 mM NaCl, 10 mM $NaN_3$ (sodium azide), 0.2% Bovine Serum Albumin (BSA) (heat shock fraction, pH 7, $\geq$ 98%), 0.1% TWEEN 20 (*see* **Note 2**).
- *Passivation buffer*: 10 mM sodium azide, 3.6% BSA and 0.1% TWEEN 20 (*see* **Note 2**).

### 2.2.2 Isolation of DNA plasmids

- Desired DNA plasmid (*see* **Note 3**)
- XL1-Blue competent cells
- Heat block, such as the Eppendorf ThermoMixer
- Orbital shaker
- Microcentrifuge, such as the Eppendorf 5424R
- Lysogeny broth (LB) medium (*see* **Note 4**)
- LB agar plates (*see* **Note 4**)
- 500 mL Erlenmeyer flask
- DNA plasmid isolation kit, such as the NucleoBondXtra Midi kit

### 2.2.3 Digestion and labeling of DNA

- Isolated DNA plasmid (*see* **Note 3**)
- Restriction enzymes (*see* **Note 3**)
- Reaction buffer for digestion
- DNA purification kit, such as the Wizard SV Gel & PCR cleanup kit
- Spectrophotometer for micro volume quantitation of nucleic acids, for instance the BioDrop µLITE
- Agarose for DNA/RNA-electrophoresis
- Horizontal electrophoresis system

- dNTPs (100 mM)
- DNA Ladder
- Ethidium Bromide (10 mg/mL)
- Klenow fragment
- Reaction buffer for Klenow Fragment
- Digoxigenin–labeled ddUTP (1 mM)
- Biotin–labeled ddUTP (1 mM)
- DNA plasmid as a template for DNA handles (*see* **Note 3**)
- Forward primer to construct DNA handles (*see* **Note 3**)
- Backward primer to construct DNA handles (*see* **Note 3**)
- T4 DNA Ligase
- Reaction buffer for T4 DNA Ligase
- DNA PCR kit, such as the FastStart Taq DNA Polymerase kit
- Digoxigenin–labeled dUTP (1 mM)
- Biotin–labeled dUTP (1 mM)
- Sodium dodecyl sulfate (SDS) $\geq$ 98.5% (GC)
- PCR Thermal Cycler, such as the Bio-Rad T100
- UV Transilluminator, such as the Bio-Rad UVT 2000
- Imaging system for visualization of electrophoresis gels, such as the ChemiDoc Imaging Systems

### 2.2.4 Flow cell assembly

- High-precision crossover tweezers
- Carbon steel scalpel
- $N_2$ spray gun
- 2-Propanol
- Custom-built aluminum flow cell body

- Custom-built Perspex mold
- Embossing Tape (9 mm)
- Cover slip 24 × 40 mm (#1.5 thickness)
- Cover slip 24 × 60 mm (#1.5 thickness)
- PDMS kit, for instance the Sylgard 184 Silicone Elastomer Kit
- FEP Tubing (1/16" outer diameter × 0.020" inner diameter)
- Flexible electrical wire (of approximately 0.02" outer diameter)
- Fitting nut for 1/16" outer diameter tubing
- Ferrule for 1/16" outer diameter tubing
- Vacuum controller
- 5 mL syringe
- M4 bolts
- Ultrasonic cleaning bath
- Glass staining trough
- 100 mL glass beaker
- Microscope slide (75 × 26 × 1 mm)

### 2.2.5 Bead-tether assembly

- DNA substrate (torsionally constrained or torsionally free)
- Magnetic rack for bead separation
- Streptavidin – coated paramagnetic beads (*see* **Note 5**)
- Nitrocellulose (0.1% in amyl acetate)
- Anti-digoxigenin (AD)
- 1 mL syringe
- 18G single-use needle
- Flexible tubing (0.8 mm inner diameter / 0.8 mm wall)
- Peristaltic pump (which can exert flow rates of approximately 50 µL / minute)

### 2.2.6 Microscope

- Microscope objective (flat-field, NA = 1.3, 40×, oil immersion, field number 25)
- CMOS camera (25 Mpix, 8 bit, 30 fps, Camera Link – Full interface)
- Infinity-corrected tube lens (f = 200 mm)
- Kinematic pitch/yaw adapter
- LED-collimator (645 nm, 100 µW, 20mA)
- Frame grabber for Camera Link – Full interface
- Multi-core PC (*e.g.* 10-cores) with 32GB DDR3 memory
- *xyz* piezo stage (20 µm travel in *z*, 0.1 nm resolution)
- Cube magnets (N50 magnetized, 5 mm)
- Stepper motor controller (6-axis, 1-256 micro-steps)
- Hollow-shaft stepper motor (2-phase, NEMA 8)
- Two translation stages to mount objective and magnets (20 mm travel, 0.1 µm resolution)
- *xy* manual positioning stage (20 mm travel in *xy*, 65 mm aperture)
- Ø1" and Ø2" broadband dielectric mirrors (λ = 400 - 750 nm)

## 2.3 Methods

### 2.3.1 Isolation of DNA plasmids

1. Take 100 µL of XL1-Blue Competent Cells (at a concentration of approximately $3 \times 10^8$ cells / mL) from the -80°C freezer and let thaw on ice (*see* **Note 6**).

2. Add approximately 1.0 µL of the desired DNA plasmid (*see* **Note 7**) to the Eppendorf tube, gently mix and incubate on ice for a minimum of 30 minutes.

3. Heat-shock the sample for 90-120 seconds at 42°C by means of the thermomixer. Do not shake the sample or exceed the indicated time.

4. Place the tube back on ice for 60 seconds and add 900 µL of LB medium. Gently shake (240 rpm) and incubate the tube for 30-60 minutes at 37°C.

   item Centrifuge the tube for 60 seconds at 20,000 g, remove approximately 90% of the supernatant and re-suspend the sample.

5. Spread the sample on an LB agar plate containing the appropriate antibiotic (here: ampicillin) and incubate (upside down) at 37°C overnight.

6. Pick up one colony and put it into a 500 mL Erlenmeyer flask containing 250 mL LB medium and the appropriate antibiotic (here: ampicillin). Incubate on the shaker (320 rpm) at 37°C, overnight.

7. Isolate and purify the DNA plasmid DNA the NucleoBond® Xtra Midi kit (Macherey-Nagel) following the manufacturers' protocol.

### 2.3.2 Digestion and labeling of torsionally free DNA constructs

1. Digest the DNA plasmid by mixing the reagents listed in Table 2.1 in an Eppendorf tube and incubate at 37°C, overnight (*see* **Notes 8 and 9**).

2. Inactivate BsaI by incubating for 20 minutes at 65°C (*see* **Note 10**).

3. Purify DNA with Promega Wizard SV Gel & PCR cleanup kit following the manufacturers' protocol (*see* **Notes 11 and 12**). Determine the DNA concentration after purification by measuring A260 nm peak absorbance with the spectrophotometer and store a microliter of the reaction mixture for agarose gel electrophoresis.

4. Label one end of the digested DNA by mixing the reagents listed in Table 2.2 in an Eppendorf tube and incubate at 37°C, for 2 hours (*see* **Note 13**). The end concentration of (labeled) dNTPs should be 20 µM (*see* **Note 14**). Note that in this step the DNA is labeled with ddUTP rather than dUTP (*see* **Note 15**).

5. Purify DNA and store a microliter of reaction mixture for agarose gel electrophoresis.

6. Digest the DNA labeled in the previous steps with a single digoxigenin by mixing the reagents listed in Table 2.3 in an Eppendorf tube and incubate at 37°C, overnight.

7. Inactivate BseYI by incubating for 20 minutes at 80°C.

8. Purify DNA and store a microliter of reaction mixture for agarose gel electrophoresis.

| | | |
|---|---|---|
| pUC18 DNA plasmid (with 601-array) (2000 ng/µL) | 20 µL | 40 µg |
| NEBuffer (10× concentrated) | 80 µL | – |
| Milli-Q water | 696 µL | – |
| BsaI (10 U/µL) | 4 µL | 40 units |
| **Total volume** | **800 µL** | |

**Table 2.1**
Composition of digestion reaction mixture I, used for preparation of the torsionally free construct.

| | | |
|---|---|---|
| BsaI-digested pUC18 DNA (with 601-array) (400 ng/µL) | 50 µL | 20 µg |
| Reaction buffer for Klenow Fragment (10× concentrated) | 10 µL | – |
| Digoxigenin-11-ddUTP (1 mM) | 2 µL | 20 µM |
| dGTP (1 mM) | 2 µL | 20 µM |
| dCTP (1 mM) | 2 µL | 20 µM |
| Milli-Q water | 24 µL | – |
| Klenow Fragment, LC (2 U/µL) | 10 µL | 20 units |
| **Total volume** | **100 µL** | |

**Table 2.2**
Composition of labeling reaction mixture I, used for preparation of the torsionally free construct.

9. Label the second extremity of the DNA by mixing the reagents listed in Table 2.4 in an Eppendorf tube and incubate at 37°C, for 2 hours.

10. Purify DNA and store a microliter of reaction mixture for agarose gel electrophoresis.

11. Check the digestion reaction products by electrophoresis using a 1% agarose gel. Pre-stain the gel with ethidium bromide by diluting the stock 10,000× in the agarose solution before it cures (*see* **Note 16**). Load the gel with the products obtained at different steps of the labeling procedure, flanked by two DNA ladders. Load approximately 50-100 ng of DNA per lane. Run the gel in TBE buffer in a horizontal electrophoresis system for 1-2 hours at 90 mV or until bands are adequately separated. Visualize the electrophoresis gel in an imaging system adjusted for ethidium bromide staining.

### 2.3.3 Digestion and labeling of torsionally constrained DNA constructs

1. Digest the DNA plasmid by mixing the reagents listed in Table 2.5 in an Eppendorf tube and incubate at 37°C, overnight (*see* **Notes 8 and 9**).

| | | |
|---|---|---|
| Single-labeled pUC18 DNA (with 601-array) (200 ng/µL) | 50 µL | 10 µg |
| NEBuffer 3.1 (10× concentrated) | 20 µL | – |
| Milli-Q water | 128 µL | – |
| BseYI (5 U/µL) | 2 µL | 10 units |
| **Total volume** | **200 µL** | |

**Table 2.3**
Composition of DNA plasmid digestion reaction mixture II, used for preparation of the torsionally free construct.

| | | |
|---|---|---|
| BseYI-digested single-labeled pUC18 DNA (with 601-array) (150 ng/µL) | 50 µL | 7.5 µg |
| Reaction buffer for Klenow Fragment (10× concentrated) | 10 µL | – |
| Biotin-16-ddUTP (1 mM) | 2 µL | 20 µM |
| dCTP (1 mM) | 2 µL | 20 µM |
| Milli-Q water | 32.25 µL | – |
| Klenow Fragment, LC (2 U/µL) | 3.75 µL | 7.5 units |
| Total volume | 100 µL | |

**Table 2.4**
Composition of labeling reaction mixture II, used for preparation of the torsionally free construct.

The two enzymes are used at the same time.

2. Inactivate restriction enzymes by incubating for 20 minutes at 80°C.

3. Separate the digestion products by electrophoresis using a 1% agarose gel. Pre-stain the gel with ethidium bromide by diluting the stock 10,000× in the agarose solution before it cures (*see* **Note 16**). Load the digested DNA in a large slot, obtained by using 1.0 mm comb and merging several lanes (with tape), flanked by two DNA ladders. Before loading, add SDS to an end concentration of 0.5% (*see* **Note 17**).

4. Image the gel on the UV transilluminator to locate the bands of interest. Minimize the time on the UV transilluminator (*see* **Note 18**). Use a scalpel to cut the digested DNA from the gel, to separate the construct of interest from the DNA backbone.

5. Purify the digested DNA from gel with the Promega Wizard SV Gel & PCR cleanup kit following the manufacturers' protocol (*see* **Note 19**).

6. Construct two sets of handles with multiple biotin or digoxigenin affinity tags by polymerase chain reaction (PCR) (*see* **Notes 20 and 21**). Mix the reagents listed in Table 2.6 in two separate Eppendorf tubes and place them in the thermal cycler. Use the PCR program depicted in Table 2.7

| | | |
|---|---|---|
| pUC18 DNA plasmid (with 601-array) (2000 ng/µL) | 50 µL | 100 µg |
| NEBuffer 3.1 (10× concentrated) | 200 µL | – |
| Milli-Q water | 1720 µL | – |
| BsaI (10 U/µL) | 10 µL | 100 units |
| BseYI (5 U/µL) | 20 µL | 100 units |
| **Total volume** | **2000 µL** | |

**Table 2.5**
Composition of the reaction mixture for digestion of the DNA plasmid for the torsionally constrained construct.

and store the tubes at 4°C. Here, the DNA is labeled with dUTP rather than ddUTP (*see* **Notes 15 and 22**).

7. Digest the biotin PCR product with BsaI and the digoxigenin PCR product with BseYI by mixing the reagents listed in Table 2.8 in two different Eppendorf tubes and incubate at 37°C, overnight.

8. Inactivate restriction enzymes by incubating for 20 minutes at 80°C.

9. Purify digested PCR products with the Promega Wizard SV Gel & PCR cleanup kit following the manufacturers' protocol (*see* **Notes 12 and 23**).

10. Ligate the digested DNA plasmid with the digested PCR products containing the multiple biotin or digoxigenin affinity tags. The digested DNA and PCR products are mixed in equimolar ratio. Mix the reagents listed in Table 2.9 in an Eppendorf tube and incubate at 4°C, overnight. Before ligase is added, store a microliter of reaction mixture for agarose gel electrophoresis (*see* **Notes 24 and 25**).

11. Inactivate T4 ligase by incubating for 20 minutes at 65°C.

12. Purify ligated DNA, elute at least two times (*see* **Note 26**). Store a microliter of reaction mixture for agarose gel electrophoresis.

13. Check the ligation reaction by electrophoresis using a 1% agarose gel. Pre-stain the gel with ethidium bromide by diluting the stock 10,000× in the agarose solution before it cures (*see* **Note 16**). Load the gel with unligated and ligated reaction products, flanked by two DNA ladders. Load approximately 50-100 ng of DNA per lane. Run the gel in TBE buffer in a horizontal electrophoresis system for 1-2 hours at 90 mV or until bands are adequately separated. Visualize the electrophoresis gel in an imaging system adjusted for ethidium bromide staining.

| | | |
|---|---|---|
| Template pUC18 DNA plasmid (100 ng/µL) | 1 µL | 100 ng |
| FastStart Taq DNA Polymerase (5 U/µL) | 0.5 µL | 2.5 units |
| Forward primer to construct DNA handles (10 µM) | 1 µL | 100 nM |
| Backward primer to construct DNA handles (10 µM) | 1 µL | 100 nM |
| PCR Reaction Buffer + $MgCl_2$ (10× concentrated) | 10 µL | – |
| dNTP (10 mM) | 1 µL | 100 µM |
| Biotin-16-dUTP (1 mM) / Digoxigenin-11-dUTP (1 mM) | 1 µL | 10 µM |
| Milli-Q water | 84.5 µL | – |
| **Total volume** | **100 µL** | |

**Table 2.6**
Composition of the PCR mixture for production of DNA handles for the torsionally constrained construct.

| **Step** | **Temperature (°C)** | **Time (s)** | **Function** |
|---|---|---|---|
| 1 | 95 | 240 | Melt the DNA |
| 2 | 95 | 30 | Melt the DNA |
| 3 | 50 | 45 | Anneal the primer to the ssDNA |
| 4 | 72 | 60 | Elongation by DNA polymerase |
| 5 | 72 | 300 | Fill in any protruding ends |

**Table 2.7**
The PCR program to produce the DNA handles. Steps 2–4 are cycled 30 times.

| | | |
|---|---|---|
| Multi-bio. PCR product / Multi-dig. PCR product (200 ng/µL) | 100 µL | 20 µg |
| NEBuffer 3.1 (10× concentrated) | 100 µL | – |
| Milli-Q water | 798 µL / 796 µL | – |
| BsaI (10 U/µL) / BseYI (5 U/µL) | 2 / 4 µL | 20 units |
| **Total volume** | **1000 µL** | |

**Table 2.8**
Composition of the digestion reaction mixture for generation of the DNA handles of the torsionally constrained construct.

| | | |
|---|---|---|
| Digested pUC18 DNA (with 601-array) (100 ng/µL) | 300 µL | 30 µg |
| Digested multi-biotin PCR product (100 ng/µL) | 50 µL | 5 µg |
| Digested multi-digoxigenin PCR product (100 ng/µL) | 50 µL | 5 µg |
| T4 DNA Ligase Reaction Buffer (10× concentrated) | 80 µL | – |
| Total volume | 315 µL | – |
| T4 DNA ligase (400 U/µL) | 5 µL | 2000 units |
| **Total volume** | **800 µL** | |

**Table 2.9**
Composition of the ligation reaction mixture to produce the torsionally constrained construct.

### 2.3.4 Flow cell assembly

1. Mix the two components of the PDMS in a ratio 10 : 1 (viscous monomer : non-viscous elastomer), approximately 2 mL per flow cell. Mix carefully and thoroughly (*see* **Note 27**). Degas the PDMS by keeping the mixture at 300 mBar for 2 hours using the vacuum controller.

2. Clean the Plexiglas molds and aluminum flow cell body with 2-propanol and Milli-Q water, and dry in a stream of $N_2$. On the mold, place a strip of embossing tape which shapes the channel. Using a needle, puncture two holes through the tape, which are used to pass through two electrical wires. Place the $24 \times 40$ mm cover slip into the aluminum body. For the in- and outlet of the flow cell, insert two pieces of 3 cm flexible electrical wire into approximately 10 cm of FEP tubing. Assemble the fitting nut and ferrule onto the tubing. Mount the tubing onto the side of the flow cell. Guide the tubing through the side of the flow cell and secure the fitting nut tightly. Guide the electrical wire through the Plexiglas molds into the tubing. Fix the flow cell body and mold with the M4 bolts. The physical assembly of the flow cells is illustrated in Figure 2.1.

3. Using a 5 mL syringe, slowly flow the degassed PDMS into the flow cell. Prevent bubbles as much as possible. Cure the flow cells at 65°C for at least 5 hours (up to overnight).

4. Remove electrical wires, creating connection channels in the PDMS to the in- and outlet of flow cell. Remove the Perspex mold. Use a scalpel to widen the channel in the flow cell to approximately 10 mm. Make sure the channel is connected to the in- and outlet of the flow cell.

5. Put multiple $24 \times 60$ mm cover slips in a glass staining trough and sonicate in ultrasonic cleaning bath for 15 minutes in 2-propanol. Dry the cover slips in a stream of $N_2$ or allow the cover slips to dry in air.

6. Put 5 mL of 0.1 % nitrocellulose in amyl acetate in a 100 mL glass beaker. Place a clean microscope object slide vertically in the solution. Using reversed tweezers, place one of the sonicated cover slips into the beaker directly next to the microscope slide. Capillary forces will fill the gap between the slides with nitrocellulose solution and will quickly coat the cover slip (*see* **Note 28**).

7. Once the gap between the cover slip and the object slide is filled, use reversed tweezers to carefully slide the cover slip from the object slide. Dry the cover slip in a stream of $N_2$. Place the cover slip with the nitrocellulose-coated side at the bottom of the flow cell, sealing the flow channel or store the coated cover slip for later use.

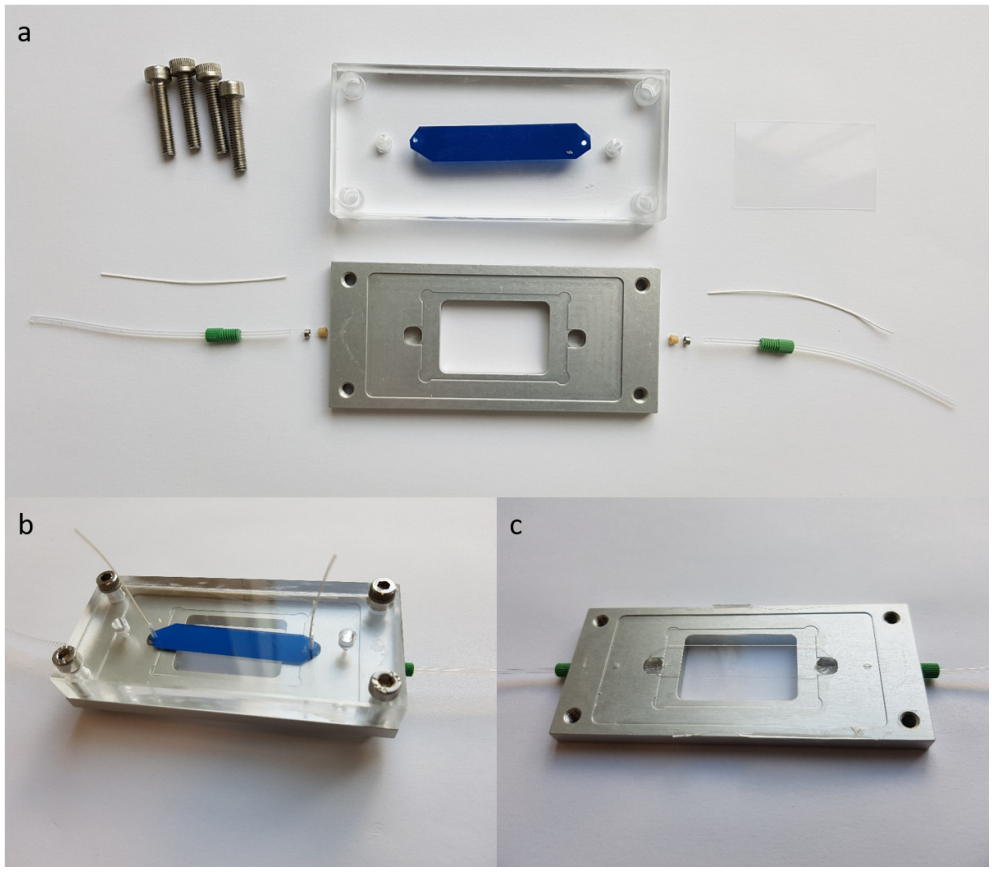



**Figure 2.1**
**The assembly of the custom-built flow cells is a three-step process.** a) Tubes and slides are assembled. The 24 × 40 mm glass slide is placed in the designated cavity of the aluminum body. Electric wire is inserted into the FEP tubing and guided upward through the mold. b) After mounting the Perspex mold, the assembled flow cell is ready for casting PDMS. c) Once the PDMS is cured, the electric wire and mold can be removed and the flow cell is covered by a 24 × 60 mm biochemically functionalized glass slide.

### 2.3.5 Bead-tether assembly

1. Connect the tube on one side of the flow cell to an empty syringe using a short length of flexible silicone tubing and a needle. Carefully flush the flow cell with 1 mL of Milli-Q water, incubate for approximately 30 seconds.

2. Replace the Milli-Q water with 300 µL of 10 ng/µL anti-digoxigenin solution. Incubate at 4°C for 2 hours.

3. Passivate the flow cell by flowing in 1 mL of passivation buffer. Incubate

at 4°C overnight (*see* **Note 29**).

4. Flush the flow cell with 1 mL of measurement buffer.

5. Dilute 1 ng of DNA in 500 µL measurement buffer. Flush the flow cell with DNA and incubate at room temperature for 10 minutes (*see* **Note 30**).

6. Vortex and wash the 1 µg/µL paramagnetic beads in Milli-Q water by means of the magnetic rack and dilute 1 µL of paramagnetic beads in 500 µL measurement buffer. Flush the flow cell with beads and incubate at room temperature for 10 minutes (*see* **Note 31**).

7. Slowly flush out the excess beads from the flow cell with 500 µL of measurement buffer using a peristaltic pump at 500 µL/minute (*see* **Note 32**).

8. The flow cell is now ready for measurements on bare DNA or incubation with DNA binding proteins.

### 2.3.6 Initial bead selection and height calibration

1. Position the magnets at such a distance that force is negligible. In our setup, depicted in Figure 2.2 (*see* **Note 33**), we position the magnets 10 mm above the sample, resulting in a force lower than 0.05 pN.

2. Place the flow cell containing the (pre-incubated) sample onto the microscope.

3. Put the beads in focus using the objective mounted onto the stepper motor. With the piezo stage, move the objective 10 µm above focus so distinct diffraction rings appear around the beads (*see* **Note 34**).

4. Select regions-of-interest (ROIs) of 150 × 150 pixels containing single, isolated beads (*see* **Note 35**).

5. Identify an immobile bead as reference bead, and use this bead for phase calibration. Move the bead 20 µm in $z$ while cross-correlating the bead images with a set of computer-generated reference images. Calibrate the phase of the cross-correlation peak with bead height. When there are no immobile beads in the field of view, phase calibration can also be done at an intermediate force, *i.e.* 1 pN, to reduce the thermal fluctuations of the tethered bead while keeping the tether intact.

6. Once the measurement plane is set and phase calibration is complete, assign the ROIs to tethered beads (manually or by a bead-finding algorithm).

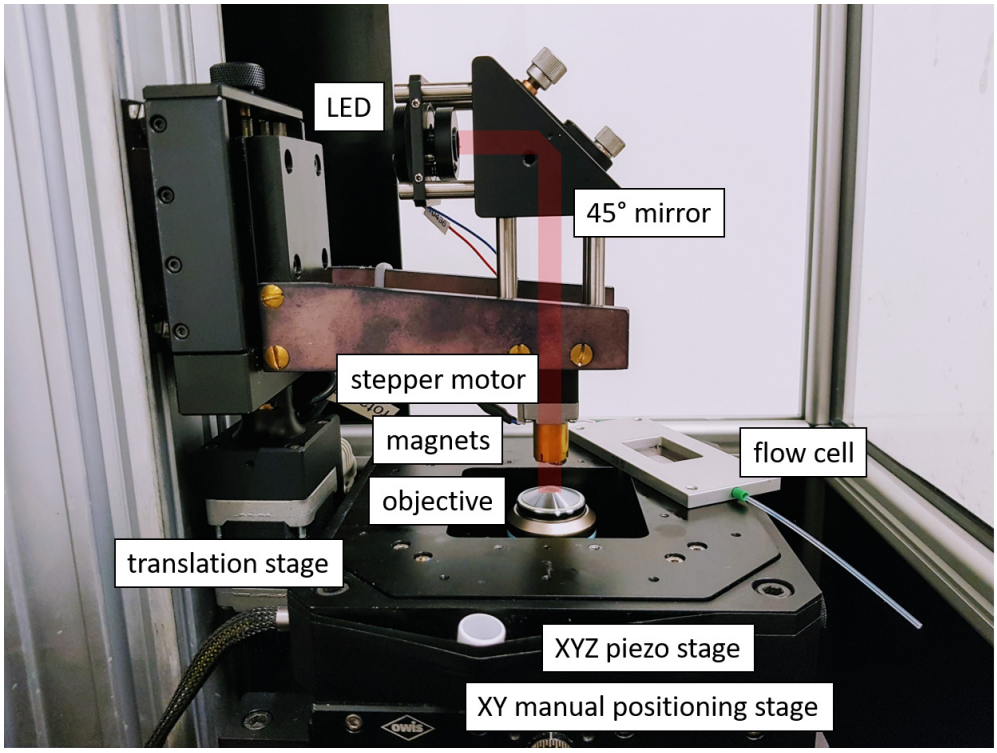



**Figure 2.2**
**The magnetic tweezers setup.** A collimated LED is mounted in the kinematic pitch/yaw adapter and coupled into a hollow-shaft stepper motor by a Ø1" mirror. The incoming light passes between a pair of magnets into the objective. The flow cell is mounted on top of an $xyz$ piezo stage, above the objective (here, the flow cell is shifted to the side for clarity). The piezo stage is mounted on top of an $xy$ manual positioning stage, for coarse movements. The magnets and objective are mounted onto two translation stages. Below the objective, the light is coupled into the camera through an infinity-corrected tube lens by means of a Ø2" mirror. The light path is depicted in red.

7. Start tracking the beads (*see* **Notes 36 and 37**).

### 2.3.7 Force calibration

1. For force calibration, prepare a flow cell containing torsionally free DNA tethers. Preferably, the tethers are relatively long with respect to the magnetic bead size (*see* **Note 38**). Start the measurements with the magnets far away, exerting negligible force. Move the magnets quickly down to the position that needs to be calibrated, keep the magnets fixed for 120 seconds (*see* **Note 39**), and move the magnets upwards. Append two stretches of negligible force before and after the experiment to provide a reference for drift correction. If possible, increase the sample rate $f_s$ of

the camera (*see* **Note 38**). Perform multiple calibration measurements at different magnet positions.

2. Load the data into data analysis software such as LabVIEW, Origin, Python or Excel. Correct the drift by subtracting a term proportional with time from the extension prior to further analysis (*see* **Note 40**).

3. Exclude anomalous events, such as sticking (*see* **Note 41**), double bead attachment to a single tether (*see* **Note 42**), or double tether attachment to a single bead (*see* **Note 43**).

4. Evaluate the $x$-coordinates (the direction of the magnetic field) in the time-domain where the magnet was exerting a constant force. Compute the single-sided power spectral density (PSD) in units of $\mu m^2_{rms}/Hz$.

5. Exclude all spectral components below 1 Hz, to discard $1/f$ noise and to minimize the effect of drift.

6. Fit the PDS to a blur and aliasing corrected PSD function to obtain the cut-off frequency to calculate the force

$$\mathrm{PDS}_{\mathrm{corrected}}(f, f_c) = \frac{\sigma^2}{f_s} + \sum_{n=-1}^{1} \frac{k_{\mathrm{B}}T}{\gamma\pi^2\left(f_c^2 + (f + nf_s)^2\right)} \left(\frac{\sin\left(W\pi(f + nf_s)\right)}{W\pi(f + nf_s)}\right)^2. \tag{2.1}$$

*See* **Note 44** for a derivation of Equation 2.1 and a full description of the parameters.

7. Plot force $F$ as a function of magnet position $h$, and fit the data with a bi-exponential decay [25]

$$F(h) = F_{\max}\left(0.7\exp\left(\frac{-h}{c_1}\right) + 0.3\exp\left(\frac{-h}{c_2}\right)\right). \tag{2.2}$$

This relation is unique for each combination of magnets and the type of magnetic beads. For our setup and 2.8 µm beads, we obtained $F_{\max} = 85$ pN, $c_1 = 1.4$ mm and $c_2 = 0.8$ mm.

### 2.3.8 Force spectroscopy experiments

1. For a dynamic force spectroscopy experiment on a torsionally free substrate, move the magnets down, increasing the force, and subsequently upwards, releasing the force. Move the magnets at a speed of 0.2 mm/s (*see* **Note 45**). Minimize the dwell time at the highest force to reduce the chance of a tether break. Several variations in trajectories are found in **Note 46**. Optionally, two stretches of low force can be appended before and after the experiment for drift correction.

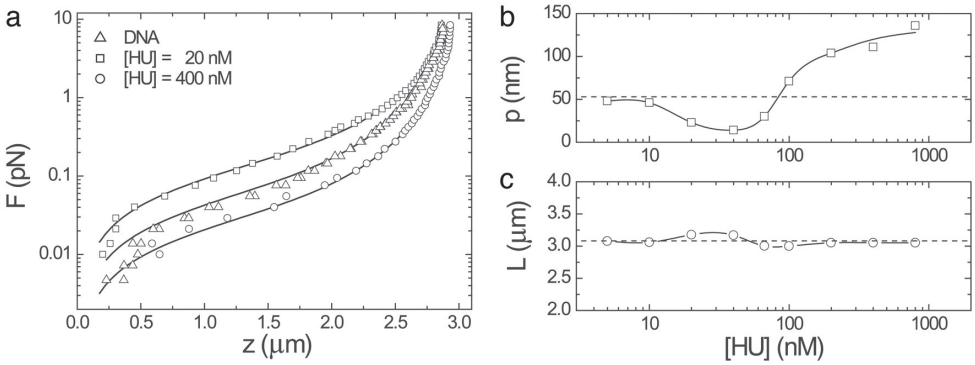



**Figure 2.3**
**The concentration of nucleoid-associated protein HU has a dual effect on the measured stiffness of a DNA tether.** (a) The force-extension (F (pN) vs. z (µm)) curves of DNA incubated with increasing concentration of HU. Solid lines represent fits to the extensible WLC model (Equation 2.3). The experiments reveal tether softening at nM HU concentration and tether stiffening at µM HU concentration. Experiments were performed in 60 mM of KCl and 20 mM HEPES (pH 7.9). (b and c) The persistence length p(nm) and contour length L (µm) as a function of HU concentration ([HU] (nm)). The dashed lines represent the values for bare DNA. Individual HU dimers introduce kinks in the DNA at nM HU concentration [35], resulting in a reduced persistence length. A rigid nucleoprotein filament is formed at high HU concentration. HU concentration does not affect the contour length. Reproduced from ref. [6] with permission from PNAS (Copyright (2004) National Academy of Sciences, U.S.A.)

2. For a dynamic force spectroscopy experiment on a torsionally constrained substrate, follow the same procedure as above. Optionally, pre-twist the torsionally constrained substrate before a dynamic force spectroscopy experiment: rotate the magnetic field while the stretching force upon the tether is kept low. Forces as low as 0.05 pN are sufficient for the bead to follow the rotation of the magnetic field. Rotate the magnets with a speed of 2 rotations per second (*see* **Note 47**).

### 2.3.9 Rotational spectroscopy experiments

1. For rotational spectroscopy on a torsionally constrained substrate, set the force by fixing the magnet height throughout the experiment. Rotate the magnets with a speed of 2 rotations per second. Apply positive twist as desired and relax the tether. Minimize the dwell time at the highest number of rotations (*see* **Note 48**). Apply an equal amount of negative twist at the same rotational speed, and relax the tether. Optionally, two stretches with fixed rotation can be appended before and after the experiment for drift correction.

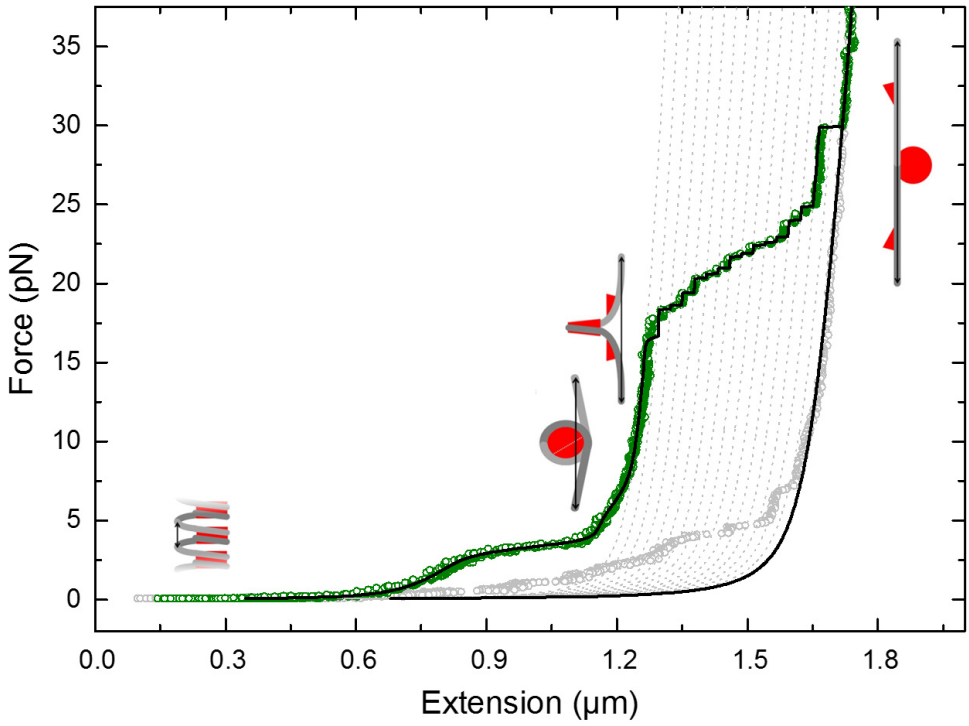



**Figure 2.4**
**Force spectroscopy on a eukaryotic chromatin fiber reveals different levels of compaction in 2 mM $MgCl_2$.** The green curve is the stretch curve; the grey curve is the release curve. The black line describes a fit of the data to the statistical mechanics model [11]. The chromatin was reconstituted on an array of 15 Widom 601 nucleosome positioning sequences spaced by 50 base pairs [36]. The low-force regime shows a non-cooperative transition at 3.5 pN, typical for a solenoid chromatin fiber, where the interactions between the nucleosomes are broken and the outer turn of DNA unwraps from the histone core, forming a beads-on-a-string structure. At approximately 6 pN, another transition takes place where this structure is slightly extended. Above 10 pN stepwise unwrapping indicates release of the final wrap of DNA from the histone core. The dashed lines represent discrete 25 nm steps. The chromatin fiber follows the force-extension curve of bare DNA at forces exceeding 25 pN and fits well with an extensible WLC model.

### 2.3.10 Data analysis of force spectroscopy experiments

1. Load the data into data analysis software such as LabVIEW, Origin, Python or Excel. Calculate force from magnet height using Equation 2.2. Correct the drift by subtracting a term proportional with time from the extension, such that the parts of the extension trace that were appended to the start and the end have the same height (*see* **Note 40**). Alternatively, a (linear) time-dependent term can be included the fitting model. Plot the force as function of extension (*see* **Note 49**).

2. Exclude anomalous events, such as sticking (*see* **Note 41**), double bead attachment to a single tether (*see* **Note 42**), or double tether attachment to a single bead (*see* **Note 43**).

3. Fit the data.

   (a) For bare DNA, fit an extensible WLC model to the data:

   $$z(F,t) = L\left(1 - \frac{1}{2}\sqrt{\frac{k_{\mathrm{B}}T}{FP}} + \frac{F}{S}\right) + z_0 + d \cdot t, \tag{2.3}$$

   where $z$ is the extension, $L$ the contour length, $P$ the persistence length, $F$ the force, $S$ the stretch modulus, $k_{\mathrm{B}}$ Boltzmann's constant, $T$ the absolute temperature, and $z_0$ an offset in extension (*see* **Note 50**). Optionally, a drift, comprising amplitude $d$ and time $t$ can be included.

   (b) For (bacterial) chromatin, fit appropriate models to the data. Generally, the effect of protein binding can be captured in a modified WLC model by adjusting fitting parameters (*see* **Note 51**). An example is shown in Figure 2.3, reproduced from Van Noort *et al.* [6], where the effect of HU-binding on the mechanical parameters of a DNA tether is shown. Analysis of partial nucleoprotein filaments can be performed when the mechanical parameters of homogeneous, fully saturated nucleoprotein filaments have been established. Then, changes in extension can be described as a linear combination of the extensions of parts of bare DNA and parts of nucleoprotein filament [9, 37]. An example of a complex, dynamic chromatin structure is plotted in Figure 2.4, showing force spectroscopy on a nucleosomal array reconstituted from tandem repeats of the Widom 601 sequence and eukaryotic histone proteins [36] (*see* **Note 52**).

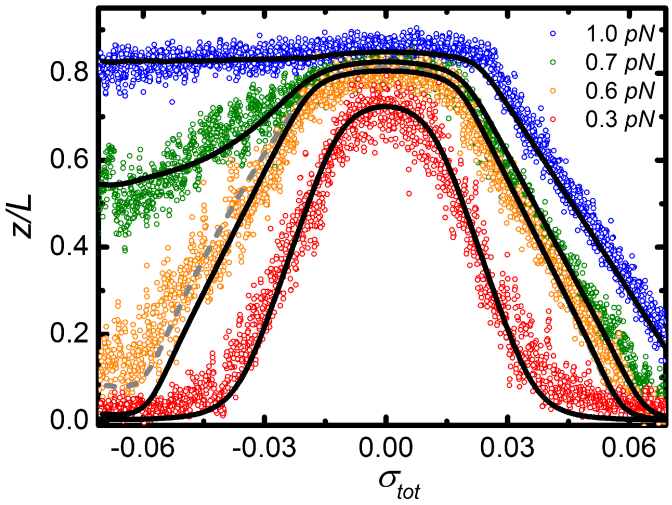



**Figure 2.5**
**The relative extension of a DNA tether as a function of linking number density shows an asymmetric response to twist at forces above 0.5 pN.** At low force, plectonemes are formed for both negative and positive twist, reducing the measured extension. Negative twist induces DNA melting at higher force, which keeps the extension constant. The lines represent calculated extensions following a three-state model that distributes twisted, plectonemic and melted regions in the DNA. Reproduced from ref. [31] with permission from Elsevier.

### 2.3.11 Data analysis of rotational spectroscopy experiments

1. Load the data into data analysis software. Correct the drift following the same procedure as that of the force spectroscopy experiments. Correct the offset in $z$ by computing the first percentile of the extension as a measurement of the surface and subtract it from all $z$-coordinates. Calculate linking number density from magnet rotation (*see* **Note 53**). Plot the relative extension as a function of linking number density.

2. Exclude disruptive events, following the same procedure as that of the force spectroscopy experiments (*see* **Notes 54-57**).

3. Fit the data.

   (a) For bare DNA, to a three-state model [31], which describes the coexistence of twisted, melted and plectonemic conformational states (*see* **Note 58**). An example is shown in Figure 2.5.

   (b) For (bacterial) chromatin, fit an appropriate model to the data. The torsional response of chromatin fibers can be modeled as a linear combination of twisted and wrapped DNA [38] (*see* **Note 59**).

## 2.4 Notes

**1** Although the torque is proportional to the strength of the magnetic field, the force in MT is proportional to the gradient of the magnetic field and therefore rapidly reduces with distance.

**2** In the measurement buffer: HEPES buffers the pH, NaCl sets the ionic strength, TWEEN 20 is a detergent and prevents aggregation of proteins, sodium azide is highly toxic and is added to prevent bacterial growth (when sodium azide is omitted, use sterile solutions or replace buffers frequently), and BSA is used as a crowding agent and retains flow cell passivation. In the passivation buffer, a high concentration of BSA is used to block non-specific binding of DNA and beads and tethers to the surface.

**3** For the DNA tethers, we generally use a pUC18-based DNA plasmid, carrying ampicillin resistance. We digest this plasmid with restriction enzymes BsaI and BseYI. For the handles of the torsionally constrained construct, we also use a pUC18 DNA plasmid as a template for PCR. The forward and backward primers were `5' CTC CAA GCT GGG CTG TGT 3'` and `5' GAT AAA TCT GGA GCC GGT GA 3'`. Reconstituted chromatin fibers were made using a plasmid containing 15 repeats of the Widom 601 positioning sequence spaced by 50 base pairs [36].

**4** Make LB medium by diluting 10 g tryptone, 10 g NaCl and 5 g yeast extract in 1000 mL Milli-Q water. Make LB agar plates by diluting 10 g tryptone, 10 g NaCl, 5 g yeast extract and 20 g agar in 1000 mL Milli-Q water.

**5** We advise to use Invitrogen Dynabeads™ M-270 Streptavidin or MyOne™ Streptavidin T1 paramagnetic beads. These beads have a relatively high iron content (MyOne: 26% / M-270: 14%) and monodispersed size distribution (CV $< 3\%$), which is important for reaching high forces, reliable bead tracking, and force calibration.

**6** When handling frozen competent cells, it is important to keep them cool. Frozen cells are very sensitive to temperature and will lose competence if not kept on ice.

**7** Very low concentrations, down to picograms per microliter, are sufficient.

**8** Use a reaction volume of at least 20 – 50 µL per µg DNA in a digestion.

**9** Digestion of (hundreds of) micrograms of DNA plasmid requires overnight incubation. We frequently observe undigested DNA plasmid in agarose gel electrophoresis if we shorten the incubation time.

**10** Inactivation temperatures vary for different restriction enzymes.

**11** The Promega Wizard SV Gel & PCR cleanup kit uses high concentrations of chaotropic agents to perturb the hydrogen bond network of the solvent, enabling DNA to bind to the silica of the cleanup kit [39–42].

**12** Multiple elutions with the spin column improve the yield. Furthermore, a longer incubation time increases the yield. For precious samples, the increased yield outweighs the increased waiting time.

**13** This step can be performed longer, up to overnight.

**14** Keep the reaction volumes small to reduce costs.

**15** The difference between dNTP and ddNTP is that the ddNTP is lacking an OH group at its 3' carbon, which is required for DNA polymerization. It prevents incorporation of more than one affinity tag per DNA strand as the strand cannot be extended.

**16** Ethidium bromide is an intercalating agent that fluoresces upon binding to DNA and illumination by a UV light source [43]. To visualize DNA in agarose gels it can be added to the electrophoresis gel or to the running buffer.

**17** BseYI can remain bound to DNA after digestion and alter the migration of DNA in agarose gels. To dissociate BseYI, add SDS to a final concentration of 0.5% or purify DNA before agarose gel electrophoresis, for instance with the Promega Wizard SV Gel & PCR cleanup kit.

**18** Minimize UV exposure of the sample to prevent DNA damage. Accordingly, block UV exposure of the gel with a non-transparent material such as aluminum foil. Use flanking DNA ladders to estimate the position to cut the DNA fragment from the gel. After cutting, verify that the desired band is cut out by imaging on the UV transilluminator.

**19** Dissolving the gel generally results in milliliters of DNA solution. To purify, load the spin column several times. Do not exceed the maximum amount of DNA that can be purified specified for the column.

**20** DNA plasmid pUC18 was used as a template for the primers to construct DNA handles.

**21** The PCR mixture contains 10% of biotin or digoxigenin labeled dUTP. The PCR product will thus have approximately 10 – 20 affinity tags incorporated.

**22** Use labeled dUTPs instead of ddUTPs, since chain termination is undesired.

**23** The small digestion products from the ends of the DNA (7 base pairs on the BseYI-side and 32 base pairs on the BsaI-side) are discarded by the Promega Wizard SV Gel & PCR cleanup kit, which has a cut-off at 100 base pairs.

**24** The digested pUC18 DNA (with 601-array) is approximately 7 times the length of the digested PCR products. To reach the approximate 1 : 1 : 1 molar ratio, use 1/7 : 1 : 1/7 weight ratio (multi-digoxigenin PCR product : digested pUC18 DNA (with 601-array) : multi-biotin PCR product).

**25** The unit definition of T4 DNA ligase can be quite ambiguous. In our protocol we use excessive amounts of T4 ligase for optimal yields.

**26** Purifying a DNA construct with multiple affinity tags will strongly reduce the yield as the affinity tags increase DNA adhesion to the column. The yield can be improved by multiple elutions and longer incubation times (up to hours).

**27** Mix thoroughly with, for instance, a flattened pipette tip to mix. Upon mixing, the PDMS starts to cure. Higher temperatures accelerate this process.

**28** Nitrocellulose coating of the cover slip makes the surface hydrophobic, enhancing AD binding.

**29** Incubate the flow cell with a high concentration of BSA to passivate the surface area. This step will reduce sticking of chromatin tethers during experiments.

**30** DNA can be replaced by pre-incubated protein-DNA complexes, for instance, reconstituted or native chromatin.

**31** During incubation of paramagnetic beads, keep the flow cell upright such that the beads move towards the coated cover slip by gravity.

**32** During flushing, turn the flow cell upside down such that unbound beads will detach from the surface.

**33** Our MT setup, depicted in Figure 2.2, is home-built and equipped with custom control and tracking software written in LabVIEW (available upon request). A pair of magnets are mounted 0.8 mm apart, at the center axis of a hollow shaft stepper motor that is mounted on a stepper motor-actuated translation stage. The magnets are oriented with their magnetization axis vertically, in opposite directions. For proper orientation, ensure that the magnets attract each other yet do not allow

rotation around the axis connecting the centers of the magnets. The microscope objective is mounted on a stepper motor-actuated translation stage to allow for coarse focusing. Stepper motors are controlled by a 6-axis two-phase stepper motor driver. A collimated LED is mounted in a kinematic pitch/yaw adapter above the hollow shaft motor, and provides a homogeneous illumination of the flow cell. We use an inverted microscope layout, including a tube lens and monochrome camera, to record the out-of-focus images of the Lorentz–Mie scattering pattern of the tethered bead [44]. The flow cell is mounted on an $xyz$ piezo stage, for accurate control of the focus position. The $xyz$ piezo stage is mounted onto an $xy$ positioning stage, for manual coarse translation of the flow cell. The microscope setup is surrounded by a box to minimize acoustic and thermal fluctuations and to protect the setup from dust.

**34** Before each experiment, the focus should be adjusted so that we always measure in the same plane.

**35** An automatic bead finder algorithm is currently implemented, which cross-correlates the field-of-view with a computer-generated reference image that resembles the beads. Beads are selected by setting a threshold on the cross-correlation amplitude.

**36** During a typical measurement, a series of frames are acquired and processed in real-time. The magnet position and rotation are recorded in synchrony by the same software that controls the frame grabber. Data collection is buffered in two parallel processes:

(a) **Image acquisition.** The camera captures full frames at a speed of 30 fps. The assigned ROIs and frame numbers are extracted and stored in a buffer.

(b) **Image analysis.** In parallel, the buffer is read out and analyzed. Each ROI is cross-correlated with a set of computer-generated reference images, using a FFT-based correlation algorithm. The maximum of the cross-correlation represents the $xy$-coordinates of the bead and is determined with sub-pixel accuracy. The phase of the cross-correlation at this location is also extracted, converted to a $z$-coordinate and written into a binary file. From the frame number and camera frequency a time column is computed and stored in the same file. After each measurement, the binary file is transformed into an ASCII-file that allows for versatile data processing. This file contains a time column, a column with the motor positions and four columns per bead: three spatial coordinates and the amplitude of the cross-correlation.

**37** Any appropriate bead tracking algorithm will suffice for this step. Use, for instance, the cross-correlation method [45], a center-of-mass calculation [46], directly fitting a Gaussian curve to the intensity profile [47], the quadrant-interpolation method [48] or the phase tracking method (introduced in Chapter 3 of this thesis).

**38** We used a 6955 base pairs DNA tether ($\sim$ 2.4 µm) to calibrate 1.0 µm and 2.8 µm diameter magnetic beads and recorded images with a frame rate of 200 Hz. Longer tethers provide a better signal-to-noise ratio, therefore, it is not uncommon that Lambda-DNA is used for calibration, which is 48,502 base pairs ($\sim$ 16.5 µm) in length. If appropriate camera corrections are applied (*i.e.* motion blur and aliasing, *see* **Note 44**), force calibration can be performed with shorter tethers.

**39** The required measurement time $\tau_m$ for a desired statistical accuracy $\epsilon$ is described in Equations 2.4 and 2.5 [25]

$$\tau_m \approx \frac{12\pi^2\eta RL}{F\epsilon^2} \qquad \text{for } F > 1 \text{ pN}, \tag{2.4}$$

$$\tau_m \approx \frac{8\pi^2\eta RPL}{k_{\mathrm{B}}T\epsilon^2} \qquad \text{for } F < 1 \text{ pN}, \tag{2.5}$$

where $\eta$ is the viscosity of water, $R$ the radius of the bead, $L$ the DNA contour length, $F$ the calibrated force, and $P$ the persistence length. Typically, the statistical accuracy $\epsilon$ is set to 0.05. We measure each magnet position for 120 seconds, which is well within limits.

**40** A plot of extension as a function of time may reveal drift as a steady increase or decrease of extension over time. When beginning and ending a trajectory at low force (or no rotations), at which the $z$-position of the bead is constrained by the bottom slide of the flow cell, compare the first percentile of the bead height of these two segments to quantify the drift. Calculate the first percentile rather than the average, since the surface poses a one-sided boundary to the bead motion. The first percentile will give a good estimation of the surface position, which is required for drift characterization. Typically, we measure less than 1 nm/s drift of the extension, yet this can accumulate to signification deviations of the measured force-extension curves.

**41** When the bead is in close proximity to the surface, non-specific interactions between the bead and the surface can cause the bead or part of the tether to stick to the flow cell bottom, immobilizing the bead. Small forces are generally sufficient to unstick the bead and resume the measurement as normal. As opposed to intramolecular rupturing or folding events, (un-)sticking events can be identified in the data as an abrupt step in the coordinates in three dimensions, rather than only in the $z$-direction.

**42** A double bead doubles the force exerted on the tether, therefore, transitions are measured at half the force at which they are expected (for instance, DNA overstretching measured at half the force). The measured persistence length is double of that measured with a single bead. Furthermore, the data is generally very noisy, since tracking becomes problematic with two beads in the same ROI.

**43** A double tether halves the force exerted on the tether. Therefore, transitions are measured at double the force as they are expected and add up (for instance, double the amount of stepwise unwrapping events when stretching chromatin, measured at double the force). The measured persistence length is half that of a single tether. A plot of the $x$-coordinates versus the $y$-coordinates of the bead during drift characterization is asymmetric.

**44** The magnetic force pulls the tethered bead in the $z$-direction and is counteracted by a restoring force from the tether. Brownian motion causes the bead to move away from its central position. The potential energy $E_p(x)$ in the direction of the field ($x$-direction) around the equilibrium position is described by:

$$E_p(x) = \frac{1}{2} k_x \delta x^2, \tag{2.6}$$

where $\delta x^2$ is the variance of the bead fluctuations. The effective trap stiffness $k_x$ follows

$$k_x = \frac{F}{z}, \tag{2.7}$$

where $F$ is the stretching force and $z$ the bead height. The equipartition theorem states that the energy per degree of freedom equals $\frac{1}{2} k_\mathrm{B} T$. The force therefore follows

$$F = \frac{k_\mathrm{B} T z}{\langle \delta x^2 \rangle}. \tag{2.8}$$

For an accurate experimental determination of the force however, it is necessary to look into the frequency $f$ dependence of the thermal fluctuations of $x$. Start by computing the PSD in the direction of the field at constant force. Since the PDS follows a Lorentzian curve

$$\mathrm{PDS}_\mathrm{thermal}(f, f_c) = \frac{k_\mathrm{B} T}{\gamma \pi^2 (f_c^2 + f^2)}, \tag{2.9}$$

which is fully defined by the lateral friction coefficient $\gamma$ and the cut-off frequency $f_c$

$$f_c = k_x / 2\pi\gamma. \tag{2.10}$$

The friction coefficient $\gamma$ depends on the bead radius $R$ and the viscosity $\eta$ (which increases with closer proximity to the surface, *see* **Note 60**)

$$\gamma = 6\pi\eta R, \tag{2.11}$$

resulting, together with equation 7, in

$$F = 12\pi^2 \eta R z f_c. \tag{2.12}$$

Thus, fitting $f_c$ from the PDS yields the force for a given magnet position. However, the finite integration time of the camera averages out some of the bead's movement, an effect known as blurring. To account for blurring, a correction term $C_{\text{blur}}$ must be applied [24–26, 49, 50]:

$$C_{\text{blur}} = \left( \frac{\sin(W\pi f)}{W\pi f} \right)^2, \tag{2.13}$$

where $W$ is the frame integration time. The finite sampling frequency of the camera causes another artefact, known as aliasing. Aliasing can be described by folding back the parts of the spectrum that exceed the sampling frequency $f_s$

$$\text{PDS}_{\text{alias}}(f, f_c) = \sum_n \text{PDS}_{\text{thermal}}(f + nf_s, f_c) C_{\text{blur}}(f + nf_s). \tag{2.14}$$

It usually sufficient to include only one aliasing term, *i.e.* $n = -1...1$ [24]. Finally, the PSD is offset by a tracking error of variance $\sigma^2$ that is independent of frequency [50]

$$\text{PDS}_{\text{tracking}} = \frac{\sigma^2}{f_s}, \tag{2.15}$$

in which $f_s$ is the sampling frequency. Overall, fitting

$$\text{PDS}_{\text{corrected}}(f, f_c) = \frac{\sigma^2}{f_s} + \sum_{n=-1}^{1} \frac{k_{\text{B}}T}{\gamma\pi^2 \left(f_c^2 + (f + nf_s)^2\right)} \left( \frac{\sin\left(W\pi(f + nf_s)\right)}{W\pi(f + nf_s)} \right)^2. \tag{2.16}$$

to the experimental spectrum of the lateral fluctuations will yield the cut-off frequency which is used to calculate the force. Equation 2.16 is the same expression as Equation 2.1.

**45** The loading rate increases exponentially with magnet speed. For forces below 5 pN, we typically have a loading rate that ranges between 0 - 0.75 pN/s. High loading rates shift the rate of non-equilibrium events such as breaking of bonds between antibodies [51].

**46** Several variations of magnet trajectories can be used:

(a) Repeat loops to gain insight into the reversibility of DNA folding.

(b) To approach a linear loading rate, break up the trajectory in several segments with decreasing magnet speeds.

(c) Start with a short force ramp up to 0.5 pN to set free loosely stuck beads.

**47** Torsionally constrained substrates build up torque while stretched, and consequently respond differently to force. These substrates can be pre-twisted before an experiment to study the role of supercoiling. DNA compacted by histone-like proteins for instance has an intrinsically twisted structure.

**48** Supercoiling reduces the bead height, which increases the probability of beads or tethers to stick to the surface. Minimizing the dwell time at the maximum number of rotations reduces sticking.

**49** Plot force as function of extension. Plotting it in this way is customary for optical tweezers and allows for easy comparison between methods. Because MT form a force clamp rather than a position clamp, this way of plotting may be counterintuitive.

**50** For tethers that have a length that is in range with the size of the bead, off-center attachment of the DNA to the bead may lead to a large error in the measured extension. Super-paramagnetic beads have a small, but finite permanent magnetic moment, giving them a preferred orientation when exposed to a magnetic field [52]. Since tethering happens in the absence of the magnetic field, application of force will turn the magnetic moment in the direction of the magnetic field, typically horizontally, while maximizing the height of the bead. When the DNA is attached at a location exactly at the circumference of the bead that has equal distance between the magnetic poles, it will rotate the bead, such that the attachment point will be at the bottom of the bead. Any location remote from this circle will still rotate the bead around the horizontal axis along the poles, but cannot spin the bead along the horizontal axis perpendicular to the poles. Consequently, there will be an offset between the height of the bead and the extension of the tether, and the measured extension is underestimated. This artefact is specific for MT and can be corrected by shifting the data in the $z$-direction, or by including this offset as a fitting parameter.

**51** Specific modes of protein binding yield distinctive patterns:

(a) Coating the DNA tether yields a stiffening of the DNA, which can be parameterized by an increased persistence length in the entropic regime at low force and/or and increased stretch modulus at high force. These two effects can only be differentiated when a large force range is probed. Examples are HU (at µM concentration) [6] and H-NS (at nM concentration) [9, 53, 54].

(b) Bending induces a kink in the DNA trajectory, which is apparent as a reduced persistence length [35]. The magnitude of this reduction scales with the kink angle and the number of kinks, which reflect the structure and number of proteins bound to the tether. Examples of chromatin proteins that introduce a kink are HU (at nM concentration) [6, 55] and IHF (at nM concentration) [56].

(c) Wrapping of DNA around proteins results in both a reduction of the persistence length, similar to bending and depending on the amount of DNA that is wrapped, and a reduction of the contour length. When exactly 1 wrap is formed, the persistence length is similar to that of bare DNA. The archetype of wrapping proteins is the eukaryotic histone octamer that wraps 1.7 turns of DNA [57–59]. Archaeal histones wrap smaller DNA lengths [60, 61].

(d) Bridging involves protein-induced formation of contacts between distant part of the DNA tether and results in a large reduction of the contour length. The DNA that is captured in the loop does not contribute to the extension of the tether. An examples of a bridging protein is H-NS, whose mode of binding depends on the concentration of monovalent and divalent ions [9, 53, 54].

(e) Stabilization of single-stranded DNA. Melting parts of the dsDNA tether by force and or torque yields sections of single stranded DNA. This leads to a 1.6 times extension of the contour length, as well as a reduction of the persistence length down to a few nanometers in absence of protein covering the single-strand DNA [37]. Single-strand binding proteins like SSB prevent annealing of force-induced, melted DNA, yielding force-extension curves that display a large hysteresis [62].

**52** Various force regimes reveal different chromatin conformations that can all be captured in a force-dependent linear combination of bare DNA and 4 different nucleosome structures [11]. Bacterial chromatin typically consists of proteins that bend, rather than wrap DNA, have a smaller footprint and are usually interspersed with different types proteins. Furthermore, they may have more varying sequence preferences. Nevertheless, a similar analysis, customized for different characteristics of individual proteins, can be used to retrieve a detailed understanding of chromatin folding based on force spectroscopy. When the chromatin conformation is not stable due to changes of protein concentration, force or salt, this may result in a change in extension. From such changes it is possible to extract binding and/or dissociation rates.

**53** Calculate the linking number density, using $\sigma = \Delta Lk/Lk_0$, where $Lk_0$ is the linking number of DNA and equals the contour length in base pairs

divided by the helical pitch of DNA, which is 10.4 base pairs. The change in linking number, $\Delta Lk$, equals the number of rotations of the magnets.

**54** A double bead affects rotational spectroscopy measurements similar to force spectroscopy measurements. The tether is more likely to break since the force that is kept constant during the measurement is double of that expected.

**55** A double tether inhibits rotational spectroscopy measurements similar to force spectroscopy measurements. Double tethers can easily be identified when twisting the tether. During the first turn, the bead is pulled down by a large step, as the two molecules get braided. Subsequent twist lowers the bead for both negative and positive rotations, as opposed to single tethers, for which the extension is independent of negative twist at forces exceeding 1 pN.

**56** The DNA tether can become nicked during sample preparation. A single nick is sufficient to release torque and consequently the tether cannot be used for rotational spectroscopy. These tethers can be discarded during an initial rotation experiment, or can serve as control for torsionally unconstraint tethers.

**57** Plot all $x$-coordinates versus all $y$-coordinates during rotation of the bead. The shape of this plot should describe a circle. The radius of the circle equals the attachment offset from the center of the bead.

**58** Since mechanical properties of twisted DNA are much more complex than those of torsionally free DNA, fitting the curves is challenging. The three-state model developed by Meng *et al.* [31] describes the coexistence of twisted, melted and plectonemic conformational states. The data and the model feature symmetric buckling at low forces ($F < 0.6$ pN), where both positive and negative twist is absorbed by plectonemes. At higher forces ($F > 0.6$ pN) the extension-twist curve becomes asymmetric, and negative twist is absorbed by local melting of the DNA.

**59** Quantitative analysis of the torsional stress in chromatin filaments is more involved. Bending proteins (such as HU of IHF) and wrapping proteins (such as HMf or eukaryotic histones) induce writhe in a torsionally constrained tether, while twist remains constant. The torsional stiffness of such fibers has not been measured, though quantification is highly desirable to test proposed models of twist-induced (de-)compaction of chromatin fibers and their role in transcription regulation. Eukaryotic tetrasomes have been shown to have a complex dynamic chirality [28, 29].

**60** Beads experience an increasing viscosity as the distance to the surface approaches the bead diameter [63]. The height dependence of the lat-

eral friction coefficient $\gamma$ is approximated by Faxén's law, described in Equation 2.17.

$$\gamma = \frac{\gamma_0}{1 - \frac{9R}{16h} + \frac{R^3}{8h^3} - \frac{45R^4}{256h^4} - \frac{R^5}{16h^3}} \tag{2.17}$$

where $\gamma_0 = 6\pi\eta R$ is the bulk friction coefficient, $\eta$ the viscosity, $R$ the radius of the bead, and $h$ the distance of the bead center to the surface.

# Bibliography II


[1] Martijn S Luijsterburg et al. "The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes". In: *Critical reviews in biochemistry and molecular biology* 43.6 (2008), pp. 393–418.

[2] Carlos Bustamante, Zev Bryant, and Steven B. Smith. "Ten years of tension: single-molecule DNA mechanics". In: *Nature* 421.6921 (2003), pp. 423–427. ISSN: 0028-0836. DOI: 10.1038/nature01405.

[3] S. Smith, L Finzi, and C Bustamante. "Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads". In: *Science* 258.5085 (1992), pp. 1122–1126. ISSN: 0036-8075. DOI: 10.1126/science.1439819.

[4] T. R. Strick et al. "The elasticity of a single supercoiled DNA molecule." In: *Science (New York, N.Y.)* 271.5257 (1996), pp. 1835–1837. ISSN: 0036-8075. DOI: 10.1126/science.271.5257.1835.

[5] Charlie Gosse and Vincent Croquette. "Magnetic tweezers: micromanipulation and force measurement at the molecular level." In: *Biophysical Journal* 82.6 (2002), pp. 3314–3329. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(02)75672-5.

[6] J. van Noort et al. "Dual architectural roles of HU: Formation of flexible hinges and rigid filaments". In: *Proceedings of the National Academy of Sciences* 101.18 (2004), pp. 6969–6974. ISSN: 0027-8424. DOI: 10.1073/pnas.0308230101.

[7] Marijn T J Van loenhout et al. "Dynamics of RecA filaments on single-stranded DNA". In: *Nucleic Acids Research* 37.12 (2009), pp. 4089–4099. ISSN: 03051048. DOI: 10.1093/nar/gkp326.

[8] Timothée Lionnet et al. "Single-molecule studies using magnetic traps". In: *Cold Spring Harbor Protocols* 7.1 (2012), pp. 34–49. ISSN: 19403402. DOI: 10.1101/pdb.top067488.

[9] Ci Ji Lim, Linda J. Kenney, and Jie Yan. "Single-molecule studies on the mechanical interplay between DNA supercoiling and H-NS DNA architectural properties". In: *Nucleic Acids Research* 42.13 (2014), pp. 8369–8378. ISSN: 13624962. DOI: 10.1093/nar/gku566.

[10] Ramon A Van Der Valk et al. "Genomic looping: a key principle of chromatin organization". In: *Journal of molecular microbiology and biotechnology* 24.5-6 (2014), pp. 344–359.

[11] 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.

[12] Jie Yan, Dunja Skoko, and John F. Marko. "Near-field-magnetic-tweezer manipulation of single DNA molecules". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 70.1 1 (2004). ISSN: 15393755. DOI: 10.1103/PhysRevE.70.011905.

[13] Bojk A. Berghuis et al. "High-throughput, high-force probing of DNA-protein interactions with magnetic tweezers". In: *Methods* 105 (2016), pp. 90–98. ISSN: 10959130. DOI: 10.1016/j.ymeth.2016.03.025.

[14] Noah Ribeck and Omar A. Saleh. "Multiplexed single-molecule measurements with magnetic tweezers". In: *Review of Scientific Instruments* 79.9 (2008). ISSN: 00346748. DOI: 10.1063/1.2981687.

[15] Iwijn De Vlaminck et al. "Highly parallel magnetic tweezers by targeted DNA tethering". In: *Nano Letters* 11.12 (2011), pp. 5489–5493. ISSN: 15306984. DOI: 10.1021/nl203299e.

[16] Jan Lipfert et al. "Freely orbiting magnetic tweezers to directly monitor changes in the twist of nucleic acids." In: *Nature communications* 2.May (2011), p. 439. ISSN: 2041-1723. DOI: 10.1038/ncomms1450. arXiv: arXiv:1011.1669v3.

[17] Jan Lipfert et al. "Magnetic torque tweezers: measuring torsional stiffness in DNA and RecA-DNA filaments". In: *Nature Methods* 7.12 (Oct. 2010), pp. 977–980. ISSN: 1548-7091. DOI: 10.1038/nmeth.1520.

[18] J. K. Fisher et al. "Thin-foil magnetic force system for high-numerical-aperture microscopy". In: *Review of Scientific Instruments* 77.2 (2006). ISSN: 00346748. DOI: 10.1063/1.2166509.

[19] Piercen M. Oliver, Jin Seon Park, and Dmitri Vezenov. "Quantitative high-resolution sensing of DNA hybridization using magnetic tweezers with evanescent illumination". In: *Nanoscale* 3.2 (2011), pp. 581–591. ISSN: 2040-3364. DOI: 10.1039/C0NR00479K.

[20] John S. Graham, Reid C. Johnson, and John F. Marko. "Concentration-dependent exchange accelerates turnover of proteins bound to double-stranded DNA". In: *Nucleic Acids Research* 39.6 (2011), pp. 2249–2259. ISSN: 03051048. DOI: 10.1093/nar/gkq1140.

[21] John S. Graham, Reid C. Johnson, and John F. Marko. "Counting proteins bound to a single DNA molecule". In: *Biochemical and Biophysical Research Communications* 415.1 (2011), pp. 131–134. ISSN: 0006291X. DOI: 10.1016/j.bbrc.2011.10.029. arXiv: NIHMS150003.

[22] Xi Long et al. "Mechanical unfolding of human telomere G-quadruplex DNA probed by integrated fluorescence and magnetic tweezers spectroscopy". In: *Nucleic Acids Research* 41.4 (2013), pp. 2746–2755. ISSN: 03051048. DOI: 10.1093/nar/gks1341.

[23] ID Vilfan et al. "Magnetic tweezers for single-molecule experiments". In: *Handbook of Single-Molecule Biophysics*. Springer, 2009, pp. 371–395.

[24] Aartjan J.W. W Te Velthuis et al. "Quantitative guidelines for force calibration through spectral analysis of magnetic tweezers data". In: *Biophysical Journal* 99.4 (2010), pp. 1292–1302. ISSN: 00063495. DOI: 10.1016/j.bpj.2010.06.008.

[25] Zhongbo Yu et al. "A force calibration standard for magnetic tweezers". In: *Review of Scientific Instruments* 85.12 (2014). ISSN: 10897623. DOI: 10.1063/1.4904148.

[26] Peter Daldrop et al. "Extending the range for force calibration in magnetic tweezers". In: *Biophysical Journal* 108.10 (2015), pp. 2550–2561. ISSN: 15420086. DOI: 10.1016/j.bpj.2015.04.011.

[27] Pooja Gupta, Jordanka Zlatanova, and Miroslav Tomschik. "Nucleosome assembly depends on the torsion in the DNA molecule: A magnetic tweezers study". In: *Biophysical Journal* 97.12 (2009), pp. 3150–3157. ISSN: 15420086. DOI: 10.1016/j.bpj.2009.09.032.

[28] Rifka Vlijm et al. "Nucleosome assembly dynamics involve spontaneous fluctuations in the handedness of tetrasomes". In: *Cell Reports* 10.2 (2015), pp. 216–225. ISSN: 22111247. DOI: 10.1016/j.celrep.2014.12.022.

[29] Rifka Vlijm et al. "Comparing the assembly and handedness dynamics of (H3.3-H4)2 tetrasomes to canonical tetrasomes". In: *PLoS ONE* 10.10 (2015). ISSN: 19326203. DOI: 10.1371/journal.pone.0141267.

[30] 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.

[31] He Meng et al. "Coexistence of Twisted, Plectonemic, and Melted DNA in Small Topological Domains". In: *Biophysical Journal* 106.5 (2014), pp. 1174–1181. ISSN: 00063495. DOI: 10.1016/j.bpj.2014.01.017.

[32] John F. Marko and Eric D. Siggia. "Stretching DNA". In: *Macromolecules* 28.26 (Dec. 1995), pp. 8759–8770. ISSN: 15205835. DOI: 10.1021/ma00130a008.

[33] C Bustamante et al. "Single-molecule studies of DNA mechanics". In: *Current opinion in structural biology* 10.3 (June 2000), pp. 279–85. ISSN: 0959-440X. DOI: 10.1016/S0959-440X(00)00085-3.

[34] John F. Marko. "Torque and dynamics of linking number relaxation in stretched supercoiled DNA". In: *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* 76.2 (2007), pp. 1–13. ISSN: 15393755. DOI: 10.1103/PhysRevE.76.021926.

[35] 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].

[36] P T T Lowary and J Widom. "New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning." In: *Journal of molecular biology* 276.1 (1998), pp. 19–42. ISSN: 0022-2836. DOI: 10.1006/jmbi.1997.1494. eprint: NIHMS150003.

[37] Steven B. Smith, Yujia Cui, and Carlos Bustamante. "Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules". In: *Science* 271.5250 (1996), pp. 795–798. ISSN: 0036-8075. DOI: 10.1126/science.271.5250.795. eprint: arXiv:1011.1669v3.

[38] Aurélien Bancaud et al. "Nucleosome Chiral Transition under Positive Torsional Stress in Single Chromatin Fibers". In: *Molecular Cell* 27.1 (2007), pp. 135–147. ISSN: 10972765. DOI: 10.1016/j.molcel.2007.05.037. arXiv: 0708.0156.

[39] Carton W. Chen and Charles A. Thomas. "Recovery of DNA segments from agarose gels". In: *Analytical Biochemistry* 101.2 (1980), pp. 339–341. ISSN: 10960309. DOI: 10.1016/0003-2697(80)90197-9.

[40] M. A. Marko, R. Chipperfield, and H. C. Birnboim. "A procedure for the large-scale isolation of highly purified plasmid DNA using alkaline extraction and binding to glass powder". In: *Analytical Biochemistry* 121.2 (1982), pp. 382–387. ISSN: 10960309. DOI: 10.1016/0003-2697(82)90497-3.

[41] Rene´ Boom et al. "Rapid and Simple Method for Purification of Nucleic Acids". In: *Journal of Clinical Microbiology* 28.3 (1990), pp. 495–503. ISSN: 0910-6340. DOI: 10.2116/analsci.25.941.

[42] Giovanni Salvi, Paolo De Los Rios, and Michele Vendruscolo. "Effective interactions between chaotropic agents and proteins". In: *Proteins: Structure, Function and Genetics* 61.3 (2005), pp. 492–499. ISSN: 08873585. DOI: 10.1002/prot.20626.

[43] J B LePecq and C Paoletti. "A fluorescent complex between ethidium bromide and nucleic acids. Physical-chemical characterization." In: *Journal of molecular biology* 27.1 (1967), pp. 87–106. ISSN: 00222836. DOI: 10.1016/0022-2836(67)90353-1.

[44] Sang-Hyuk Lee et al. "Characterizing and tracking single colloidal particles with video holographic microscopy." In: *Optics express* 15.26 (2007), pp. 18275–18282. ISSN: 1094-4087. DOI: 10.1364/OE.15.018275. arXiv: 0712.1738.

[45] Jeff Gelles, Bruce J. Schnapp, and Michael P. Sheetz. "Tracking kinesin-driven movements with nanometre-scale precision." In: *Nature* 331.6155 (1988), pp. 450–453. ISSN: 0028-0836. DOI: 10.1038/331450a0. arXiv: arXiv:1011.1669v3.

[46] G M Lee, A Ishihara, and K Jacobson. "Direct observation of brownian motion of lipids in a membrane." In: *Proceedings of the National Academy of Sciences of the United States of America* 88.14 (1991), pp. 6274–8. ISSN: 0027-8424. DOI: 10.1073/pnas.88.14.6274.

[47] Catherine M Anderson et al. "Tracking of cell surface receptors by fluorescence digital imaging microscopy using a charge-coupled device camera. Low-density lipoprotein and influenza virus receptor mobility at 4 degrees C." In: *Journal of cell science* 101 ( Pt 2 (1992), pp. 415–425. ISSN: 0021-9533.

[48] Marijn T.J. J Van Loenhout et al. "Non-bias-limited tracking of spherical particles, enabling nanometer resolution at low magnification". In: *Biophysical Journal* 102.10 (2012), pp. 2362–2371. ISSN: 00063495. DOI: 10.1016/j.bpj.2012.03.073.

[49] Wesley P Wong and Ken Halvorsen. "The effect of integration time on fluctuation measurements: calibrating an optical trap in the presence of motion blur". In: *Optics express* 14.25 (2006), pp. 12517–31. ISSN: 1094-4087. DOI: 10.1364/OE.14.012517. arXiv: 0607156 [physics].

[50] Astrid van der Horst and Nancy R Forde. "Power spectral analysis for optical trap stiffness calibration from high-speed camera position detection with limited bandwidth." In: *Optics express* 18.8 (2010), pp. 7670–7677. ISSN: 1094-4087. DOI: 10.1364/OE.18.007670.

[51] Gerrit Sitters et al. "Acoustic force spectroscopy." In: *Nature methods* 12.1 (2015), pp. 47–50. ISSN: 1548-7105. DOI: 10.1038/nmeth.3183.

[52] Daniel Klaue and Ralf Seidel. "Torsional stiffness of single superparamagnetic microspheres in an external magnetic field". In: *Physical Review Letters* 102.2 (2009), pp. 1–4. ISSN: 00319007. DOI: 10.1103/PhysRevLett.102.028302.

[53] Yingjie Liu et al. "A divalent switch drives H-NS/DNA-binding conformations between stiffening and bridging modes". In: *Genes and Development* 24.4 (2010), pp. 339–344. ISSN: 08909369. DOI: 10.1101/gad.1883510.

[54] Ramon A van der Valk et al. "Mechanism of environmentally driven conformational changes that modulate H-NS DNA bridging activity". In: *eLife* 6 (2017). ISSN: 2050-084X. DOI: 10.7554/eLife.27369.

[55] Kerren K. Swinger and Phoebe A. Rice. *IHF and HU: Flexible architects of bent DNA*. 2004. DOI: 10.1016/j.sbi.2003.12.003.

[56] Jie Lin et al. "Physical Organization of DNA by Multiple Non-Specific DNA-Binding Modes of Integration Host Factor (IHF)". In: *PLoS ONE* 7.11 (2012). ISSN: 19326203. DOI: 10.1371/journal.pone.0049885.

[57] 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.

[58] 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.

[59] 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.

[60] Bram Henneman and Remus T. Dame. "Archaeal histones: dynamic and versatile genome architects". In: *AIMS Microbiology* 1.1 (2015), pp. 72–81. ISSN: 2471-1888. DOI: 10.3934/microbiol.2015.1.72.

[61] Francesca Mattiroli et al. "Structure of histone-based chromatin in Archaea". In: *Science* 357.6351 (2017), pp. 609–612. ISSN: 0036-8075. DOI: 10.1126/science.aaj1849.

[62] K. Hatch et al. "Measurement of the salt-dependent stabilization of partially open DNA by Escherichia coli SSB protein". In: *Nucleic Acids Research* 36.1 (2008), pp. 294–299. ISSN: 03051048. DOI: 10.1093/nar/gkm1014.

[63] Erik Schäffet, Simon F. Nørrelykke, and Jonathon Howard. "Surface forces and drag coefficients of microspheres near a plane surface measured with optical tweezers". In: *Langmuir* 23.7 (2007), pp. 3654–3665. ISSN: 07437463. DOI: 10.1021/la0622368.

# Chapter 3

# Multiplexed Nanometric 3D Tracking of Microbeads Using an FFT-Phasor Algorithm

Many single-molecule biophysical techniques rely on nanometric tracking of microbeads to obtain quantitative information about the mechanical properties of biomolecules such as chromatin fibers. Their three-dimensional position can be resolved by holographic analysis of the diffraction pattern in wide-field imaging. Fitting this diffraction pattern to Lorentz–Mie scattering theory yields the bead position with nanometer accuracy in three dimensions but is computationally expensive. Real-time multiplexed bead tracking therefore requires a more efficient tracking method, such as comparison with previously measured diffraction patterns, known as look-up tables. Here, we introduce an alternative 3D phasor tracking algorithm, that provides robust bead tracking with nanometric localization accuracy in a $z$-range of over 10 µm under non-optimal imaging conditions. The algorithm is based on a 2D cross-correlation using FFTs with computer-generated reference images, yielding a processing rate of up to 10,000 regions of interest per second. We implemented the technique in magnetic tweezers and tracked the 3D position of over 100 beads in real-time on a generic CPU. The accuracy of the algorithm was extensively tested and compared to a look-up table approach using Lorentz–Mie simulations, avoiding experimental uncertainties. Its easy implementation, efficiency, and robustness can improve multiplexed biophysical bead tracking applications, especially where high throughput is required and image artefacts are difficult to avoid.

Brouwer T. B., Hermans N., van Noort S. J. T.: Multiplexed Nanometric 3D Tracking of Microbeads Using an FFT-Phasor Algorithm, **Biophysical Journal**, 118.9, 2020: 2245-2257.

## Significance statement

Microbeads are often used in biophysical single-molecule manipulation experiments and accurately tracking their position in three dimensions is key for quantitative analysis. Holographic imaging of these beads allows for multiplexing bead tracking but image analysis can be a limiting factor. Here we present a 3D tracking algorithm based on Fast Fourier Transforms that is fast, has nanometric precision, is more robust against common artifacts than the traditional LUT method and is accurate over 10s of micrometers. We show its real-time application for magnetic tweezers based force spectroscopy on more than 100 chromatin fibers in parallel, and anticipate that other bead-based biophysical essays can benefit from this simple and robust three-dimensional phasor algorithm.

## 3.1 Introduction

Single-molecule techniques overcome ensemble averaging and can resolve unique and rare events at the molecular level [1]. By manipulation of microbeads, single-molecule force spectroscopy techniques revealed the mechanical properties of biomolecules such as DNA or RNA with unprecedented detail [2–5]. In addition, the interactions with proteins, like the DNA compaction by histones in eukaryotic chromatin [6–11] and prokaryotic architectural proteins [12–16], supercoiling [17–20], and repair processes [21–23] were extensively studied with magnetic tweezers (MT) or optical tweezers (OT), Acoustic Force Spectroscopy (AFS) [24–26] or tethered particle motion (TPM) [13, 27, 28]. These bead manipulation techniques have also been used to quantify the mechanical properties of other biological structures, such as extracellular protein collagen [29–31], or even entire cells [32].

The beads not only constitute a micron-sized handle to manipulate the molecules of interest, they also function as a label, whose position reflects the extension or deformation of the studied biomolecule. In OT, the position of one or two beads is generally measured from the deflection of a focused laser beam that is projected on a quadrant split detector [34, 35], yielding nanometric accuracy and kHz bandwidth in three dimensions. MT, TPM, and AFS however, generally use wide-field imaging with CCD or CMOS cameras and real-time image processing for position measurements. Next to sub-pixel accuracy, many applications require high framerates to resolve fast conformational changes or to capture the full spectrum of thermal motion for accurate force calibration [36]. Cameras with kHz frame rates or tens of megapixel resolution are currently available for fast or large field-of-view imaging [37, 38]. With such high-end hardware, real-time processing to resolve the three-dimensional position of the beads becomes rate-limiting. Moreover, multiplexing the image processing puts large demands on the processing power of the CPU. In some applications the GPU is employed to achieve sufficient speed [39–41].

Holographic imaging and subsequent fitting of the images to Lorentz–Mie scattering theory (LMST) has been successfully used to convert movies of colloidal spheres into tracks of 3D coordinates [33]. Besides, one can accurately retrieve other physical characteristics that define the hologram, such as the bead radius and refractive index. Despite these advantages, bead tracking applications that are used in the single-molecule biophysics field generally use simpler, empirical methods to increase processing speed. A popular and fast method for bead tracking splits the tracking into 3 stages. First, the center of the bead is determined, either by computing the center of mass [42–44], or 1D or 2D cross-correlation with the mirrored intensity profile or predefined kernel [17, 35, 45–48]. Second, a radial intensity profile is computed. Third, this radial profile is compared to a previously calibrated look-up table (LUT) of radial profiles, and the z-coordinate is interpolated from the difference curve.

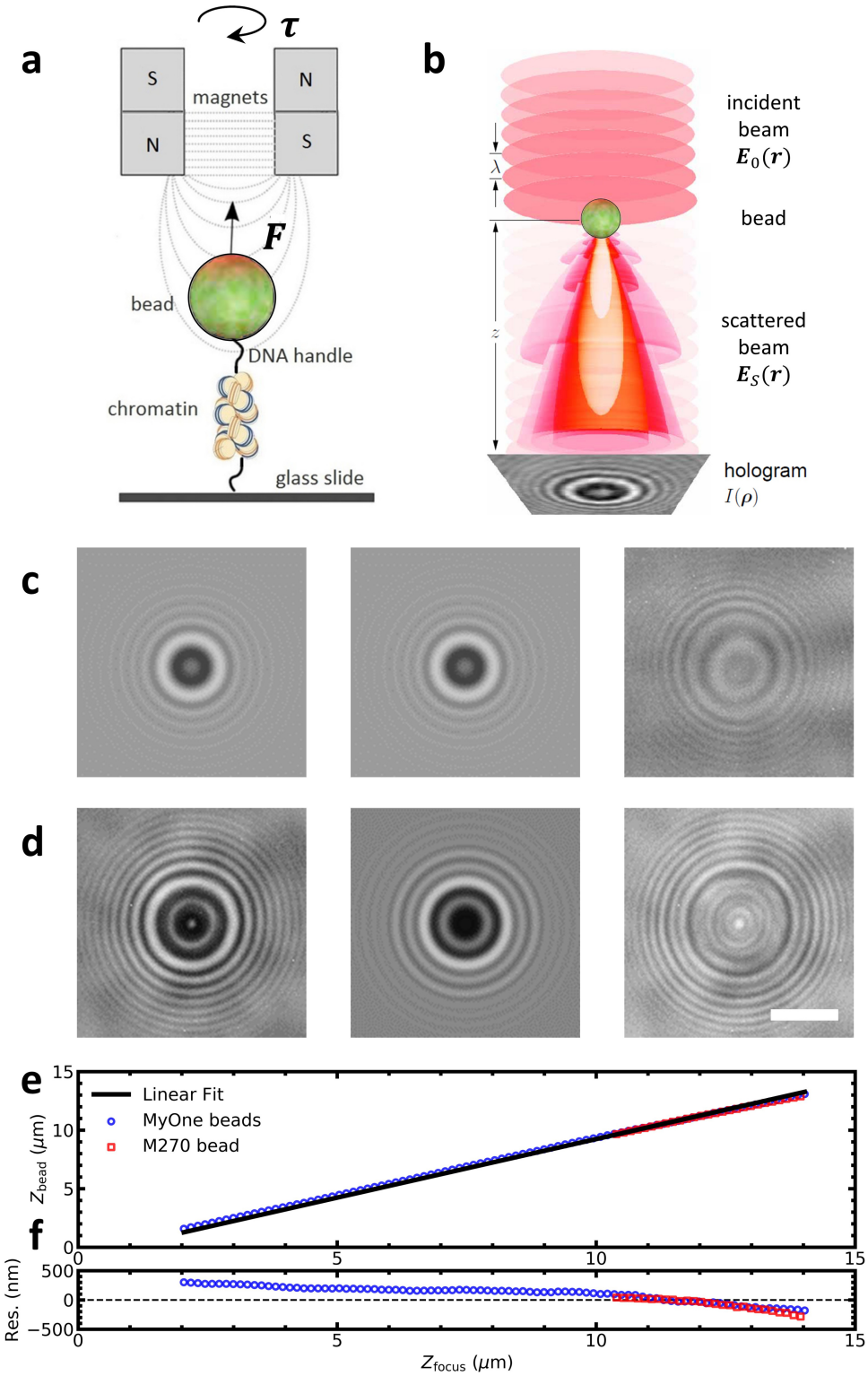

a
τ
S
N
magnets
N
S
F
bead
DNA handle
chromatin
glass slide
b
λ
z
incident beam $E_0(r)$
bead
scattered beam $E_S(r)$
hologram $I(\rho)$
c
d
e
Linear Fit
MyOne beads
M270 bead
$Z_{bead}$ (μm)
f
Res. (nm)
$Z_{focus}$ (μm)

**Figure 3.1**
**LMST cannot fully describe the holographic image of paramagnetic beads used in MT.** a) Schematic drawing of a typical MT experiment. A molecule tethers a paramagnetic bead to the bottom of a flow cell. The bead is manipulated by a pair of magnets exerting force ($\boldsymbol{F}$) and torque ($\boldsymbol{\tau}$). b) A holographic image $I(\boldsymbol{\rho})$ is recorded originating from the interference of an incident beam $E_0(r)$ with scattered the light $E_S(r)$. The diffraction pattern is analyzed to obtain the three-dimensional position of the bead with nanometer accuracy. This image was adapted with permission from Lee *et al.* © The Optical Society [33]. c) The diffraction pattern (left) of a 1.0 µm diameter paramagnetic bead (Dynabeads MyOne Streptavidin T1, Thermo Fisher Scientific) was fitted with LMST (center). The fit to yielded $x = -70 \pm 1$ nm, $y = 33 \pm 1$ nm, $z = 8300 \pm 100$ nm, $n_{\text{bead}} = 1.9 \pm 0.1$, $\alpha = 0.9 \pm 0.1$, $\beta = 57 \pm 1$, $\gamma = 57 \pm 1$ (fit±standard error). The values $a = 0.5$ µm and $n_{\text{medium}} = 1.33$ were fixed. The residual image (right) shows that some features could not be reproduced by LMST. d) The diffraction pattern of a 2.8 µm diameter paramagnetic bead (Dynabeads M270 Streptavidin, Thermo Fisher Scientific) yielded a worse fit: $x = -162 \pm 1$ nm, $y = 84 \pm 1$ nm, $z = 9500 \pm 100$ nm, $n_{\text{bead}} = 1.8 \pm 0.1$, $\alpha = 0.7 \pm 0.1$, $\beta = 49 \pm 1$, $\gamma = 67 \pm 1$. The values a = 1.4 µm and $n_{\text{medium}}$ = 1.33 were fixed. Scale bar: 3 µm. e) Two paramagnetic beads (MyOne in blue and M270 in red) were moved through focus, and the recorded holographic movies were fitted to LMST. For clarity, every fifth data point was plotted. The fits of the diffraction pattern (black lines) did not converge close to focus. Sufficiently far from the focus, the obtained bead height was proportional to $z_{\text{focus}}$. f) Residual of the linear fit of the bead height as a function of the focus height.

Quadrant interpolation minimized crosstalk between $x$, $y$, and $z$ coordinates, which significantly increased the tracking accuracy at the cost of being rather computationally intensive [49]. Cnossen *et al.* improved performance by shifting analysis to the GPU, which increased the speed and made it suitable multiplexed applications. This approach required specialized GPU hardware and advance software for analysis [50, 51].

Previously, we implemented the LUT bead tracking algorithm in our MT and used it to study the various transitions of chromatin fiber unfolding [6, 7, 52], as schematically depicted in Figure 3.1a. In these applications, we noticed that the dynamic range of the LUT method was sometimes insufficient, yielding an accuracy that depended on the bead height. As the composition of chromatin fibers may vary due to the quality of reconstitution [11], disassembly [6] or reflecting naturally occurring variations [53]. In our hands, the empirical LUT algorithm frequently flawed due to non-perfect imaging conditions, which led to discarding a large fraction of the beads, limiting the throughput. Tracking errors were enhanced by the increased field-of-view that is required for imaging multiple beads, which in practice yields more image artifacts, such as light gradients due to non-uniform illumination, astigmatism near the edges, or light obstructions by loose beads. Therefore, robustness against common image aberrations becomes increasingly important as the imaging settings can sometimes not be optimized for all beads.

Using the power of 2D Fast Fourier Transforms (FFT) to compute cross-

correlations with computer-generated reference images, we reduced the computational effort to three FFTs and skipped the generation and comparisons with radial profiles, which is the computationally most expensive part of traditional bead tracking. Instead, translations in the $z$-direction were captured into a single parameter, the phase, that we show to be proportional to the height. This new tracking algorithm, which we call 3D phasor tracking (3DPT), is simple, sufficiently fast and robust and meets all criteria for real-time multiplexed nanometric bead tracking experiments on a generic CPU.

## 3.2 Results

### 3.2.1 Tracking of super-paramagnetic beads using Lorenz-Mie scattering theory

The scattering of light by colloidal particles and its interference with the incident light is described in LMST. Figure 3.1b depicts the contrast mechanism of holographic imaging of a colloidal bead (adapted from [33]). The interference of the incident beam $E_0(r)$ with the light that is scattered off the bead $E_s(r)$, yields a circularly symmetric hologram $I(\rho)$ in the image plane. The center of the hologram corresponds to the $xy$-position of the bead. As the image plane moves away from the location of the bead, the interference pattern expands, leading to more and larger rings around the center of the bead.

We used LMST to fit the 6 parameters that describe the diffraction pattern: $x$, $y$, $z$, bead radius $a$, refractive index of the medium $n_m$, refractive index of the bead $n_p$, and three additional scaling parameters $\alpha$, $\beta$, $\gamma$ [54, 55] (*see* Materials and Methods). We tested two types of super-paramagnetic beads, either with a diameter of 1 µm (Dynabeads MyOne Streptavidin T1, Thermo Fisher Scientific, Figure 3.1c, left) or 2.8 µm (Dynabeads M270 Streptavidin, Thermo Fisher Scientific, Figure 3.1d, left), which are commonly used in MT. In both cases, we obtained a reasonable fit (Figure 3.1c and d, center), though spherical shapes in the residual images (Figure 3.1c and d right) indicate that there is a systematic discrepancy between the LMST fit and the experimentally obtained holographic images. The residuals were generally larger for the 2.8 µm beads than for the 1 µm. Though we did not further investigate this difference, we attribute it to the mixed composition of these beads, which may not fully be captured in a single refraction index used in LMST. Because the 1 µm beads are better described by LMST, we used these smaller beads in the remainder of this work.

For MT force spectroscopy applications, the $z$-coordinate is the most important parameter that can be extracted for the hologram, as it quantifies the extension of the tether. In Figure 3.1e, the fitted $z$-coordinate is plotted, when a bead that was fixed on a cover slide was moved linearly through the focus using a piezo stage. Fitting the obtained diffraction patterns with LMST at

corresponding height from focus, a fairly accurate position could be obtained over a range of more than 10 µm for the smaller and about 5 µm for the larger bead. The $xy$-coordinates also yielded reproducible results (data not shown). A closer look at the residual of fitting a linear curve to the measured versus applied height however, Figure 3.1f, shows that the systematic errors in the $z$-direction typically amounted hundreds of nanometers for our paramagnetic beads. The relative error can be smaller for smaller ranges, which makes it still useful for small tethers.

The main limitation of LMST fitting however, is the processing speed. Fitting the diffraction pattern of a single bead in a $100 \times 100$ pixels region-of-interest (ROI) took several seconds on our CPU. For off-line applications this may not be problematic, although implementation of off-line multiplexed bead tracking would imply storage and processing of very large data files. Real-time processing has the advantage that the experimenter can rapidly asses the quality of the measurements or make adjustments during the measurement. LMST fitting cannot achieve real-time processing with current computing power, which requires a more efficient tracking method.

### 3.2.2 3D Phasor Tracking

Here we introduce a novel method called 3D phasor tracking (3DPT), which exploits the circular symmetry of the LMST diffraction pattern, and its gradual expansion when the bead moves in the $z$-direction, to compute bead coordinates. Each step of the tracking algorithm is depicted in Figure 3.2. We calculated the $xy$-coordinate by cross-correlating the ROI with a computer-generated reference image, rather than its mirrored image. The complex reference image $I_n(r)$ resembles the holographic image, but consists of a single spatial frequency, characterized by period $k_n$:

$$I_n(r) = f_s(r) \cdot \left( \cos\left(2\pi \frac{r}{k_n}\right) + i \sin\left(2\pi \frac{r}{k_n}\right) \right), \tag{3.1}$$

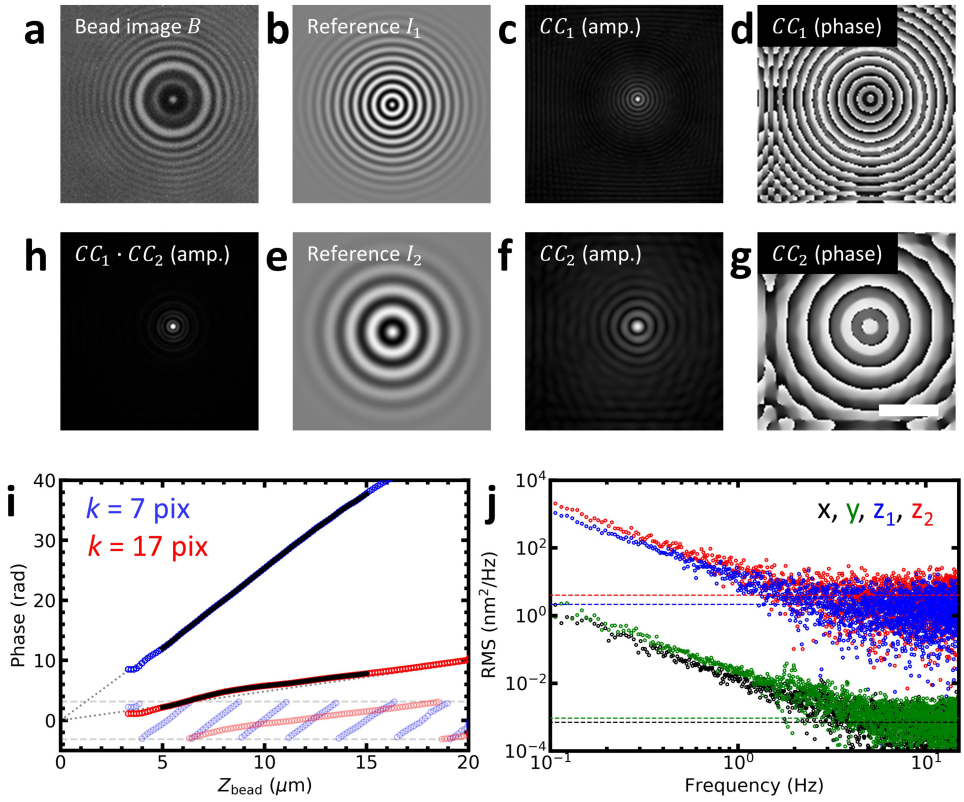
with $r$ the distance from the center of the ROI of size $s$. The reference image is spatially filtered with filter $f_s(r)$:

$$f_s(r) = \begin{cases} 1/2\,(\cos(2\pi r/s) + 1) & r \leq S/2 \\ 0 & r > S/2 \end{cases}. \tag{3.2}$$

Two examples of reference images for two different periods are shown in Figure 3.2b and e.

The experimentally measured diffraction pattern of the bead $I_{\text{bead}}$ is cross-correlated with the reference images yielding cross-correlation $CC_n$:

$$CC_n = \text{FFT}^{-1}\Big( f_B(k) \cdot \text{FFT}(I_n)^* \cdot \text{FFT}(I_{\text{bead}}) \Big), \tag{3.3}$$




## Figure 3.2

**The principle of 3DPT.** The holographic image (a) was cross-correlated with two complex reference images $I_1$ (b) and $I_2$ (e), computed using Equation 3.1 for period $k_1 = 7$ pix and $k_2 = 16$ pix. The cross-correlation yielded two complex images $CC_1$ and $CC_2$, displayed as amplitude (c, f) and phase (d, g). The amplitudes of $CC_1$ and $CC_2$ were multiplied, resulting in a sharp peak at the $xy$-position of the bead (h). The phases at the peak, $\varphi_1$ and $\varphi_2$, scaled approximately linearly with bead height $z$. Scale bar: 3 µm. i) The relation between $\varphi$ and $z$ was calibrated *a priori* using a measurement in which the focus is linearly shifted in time resulting in phases $\varphi_1$ (semi-transparent blue circles) and $\varphi_2$ (semi-transparent red circles). Subsequently, $\varphi_1$ and $\varphi_2$ were phase-unwrapped, eliminating $2\pi$ phase jumps (blue and red circles beyond the horizontal dashed lines). A linear increase in phase, starting at focus roughly describes the phase-height relation (grey dotted lines). The unwrapped phases $\varphi_1(z)$ and $\varphi_2(z)$ were fitted with a polynomial, starting 5 µm above focus (black lines). For clarity, every fifth data point was plotted. j) The tracking accuracy of 3DPT was calculated by spectral analysis, yielding an accuracy of $4 \pm 1$ nm$^2$/Hz for $z_1$, and $2 \pm 1$ nm$^2$/Hz for $z_2$. The accuracy in $x$ and $y$ was several orders of magnitude smaller: $0.7 \times 10^{-3} \pm 0.1$ nm$^2$/Hz for $x$, and $0.9 \times 10^{-3} \pm 0.1$ nm$^2$/Hz for $y$.

where a band-pass filter $f_B(k)$ is used in the frequency domain:

$$f_B(k) = \exp\left(-\frac{(k-k_n)^2}{2w^2}\right). \tag{3.4}$$

The computer-generated images of the reference signal and its filters are depicted in Supplementary Figure 3.S1.

Figures 3.2c-g show the amplitude and phase images for of the cross-correlation with a typical hologram, shown in Figure 3.2a, for periods $k_1 = 7$ pixels and $k_2 = 16$ pixels. The amplitude image of the $CC_n$ in both cases featured a single peak that represents the shift of the bead relative to the center of the image. Two amplitude images obtained by cross-correlating with reference images with two different periods were multiplied yielding a sharper peak at the $xy$-position of the bead (Figure 3.2h), which was measured with sub-pixel accuracy using polynomial interpolation in 2D.

The $z$-position of the bead was obtained from the phase $\varphi_n$ at the $xy$-position of the bead. The phase images featured concentric rings, around the bead center. In Figure 3.2i, we plotted the phase at the bead center of a fixed bead that was moved through the focus in the $z$-direction. Several micrometers above the focus, we obtained good cross-correlations with distinct peaks at the bead position. The phase was computed from the amplitude weighted average of a 10 pixel ROI around the maximum in the amplitude image. The resulting phase increased proportional to the bead height but was wrapped between -π and π. Unwrapping of the phase was trivial since the height of the bead increased linearly in time. The phase signal could thus be unwrapped unambiguously as shown in Figure 3.2i. We quantified this curve by fitting a polynomial function to the experimental $\varphi_n(z)$ data, which was inverted to compute the bead height as a function of the measured phase, $z(\varphi_n)$, during subsequent experiments.

Far from focus, $\varphi_n$ increased, to a good approximation, linearly with $z$. Empirically, we found for sufficiently defocused beads:

$$\varphi_n(z) \approx c k_n^2 z, \tag{3.5}$$

in which $c$ represents a calibration factor that depends on the magnification and the refraction index of the immersion medium, and did not change between experiments or bead sizes. This linear approximation is plotted in Figure 3.2i as dotted lines, and it can be seen that deviations become larger near focus and for larger periods, as near-field effects of the light scattering become more prominent.

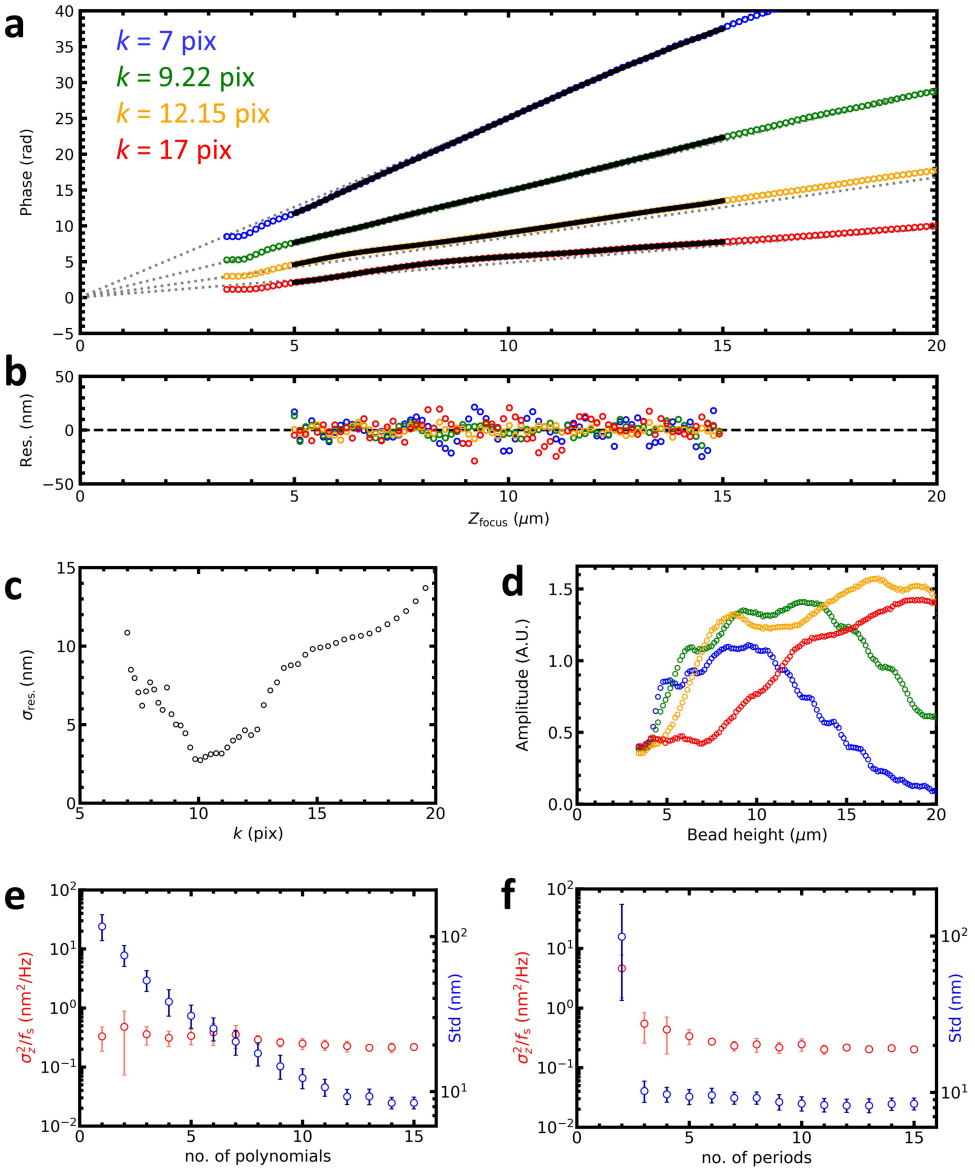
In tracking experiments in which there is no prior knowledge of the bead position, a single phase cannot be converted unambiguously into a unique $z$-position due to phase wrapping. A common solution to this phase unwrapping problem is the use of multiple frequencies [56–59], which was implemented

**Figure 3.3**
**Multiple reference images increase the tracking accuracy.** a) Phase calibration with 4 reference images (colored circles). $k_n$ was logarithmically sampled between 7 and 16 pixels. The curves were fitted with a $5^{\text{th}}$ order polynomial over a range of 10 µm (black lines). All curves converged in focus and approximated a straight line (grey dotted lines). b) The residuals of the polynomial fit for each reference image (colored circles). c) The tracking accuracy obtained with a single reference image varied with its period. The best accuracy was obtained for $k \approx 10$ pix. Below $k = 7$ pix, the correlation did not yield a distinct peak. d) The amplitude of the cross-correlation varied with $z$ and $k$. Larger periods were more prominent as the bead was shifted further out of focus. For clarity, every fifth point was plotted in graph a, b, and d. e) The tracking accuracy in the $z$-direction (red dots) slightly increased with the number of polynomials used during phase calibration. The errors significantly reduced up to approximately 5 polynomials. The average standard deviation of the residuals of the polynomial fit (blue dots) linearly decreased up to approximately 10 polynomials. f) The tracking accuracy in the $z$-direction (red dots) increased significantly with the number of reference images, up to approximately 10 reference images. Especially the use of 3 instead of 2 reference images was effective: an approximately 10-fold increase in accuracy was established. 15 polynomials were used during phase calibration. The average standard deviation of the residuals of the polynomial fit (blue dots) was not affected by the number of reference images when 3 or more images were used. The errors depicted in graph e and f indicated the standard deviation of 6 independent beads.

by computing at least two $CC_n$ images with different periods in the reference images. Provisional $z$-coordinates, which corresponded to the expected $z$-range of the bead, spaced by $2\pi$ in phase, were calculated. The set of $z$-coordinates corresponding to different spatial frequencies that showed the smallest variation was then selected and averaged to compute the final $z$-coordinate. Thus, by computing two, or more, reference images and subsequent cross-correlation with an experimental holographic image, it was possible to determine the three-dimensional position of the bead unambiguously.

### 3.2.3 Performance of 3DTP

The performance of the novel 3DPT method was tested in multiple ways. First, the power spectral density (PSD) of a time trace of an immobilized bead was computed (*see* Materials and Methods), as shown in Figure 3.2j. The tracking accuracy, expressed in $\sigma^2/f_s$, was determined as the plateau value of the PSD at frequencies over 2 Hz [34], since thermal drift and mechanical vibrations introduced low-frequency fluctuations that resulted in increased amplitudes at smaller frequencies ($1/f$ noise). The reference image with the highest spatial frequency ($k = 7$ pix) performed best: $\sigma_z^2/f_s = 2$ nm$^2$/Hz. For the lower spatial frequency ($k = 16$ pix) we obtained an accuracy of $\sigma_z^2/f_s = 4$ nm$^2$/Hz. The tracking accuracy in the $x$- and $y$-direction was several orders of magnitude higher: $\sigma_x^2/f_s = 0.7 \times 10^{-3}$ nm$^2$/Hz and $\sigma_y^2/f_s = 0.9 \times 10^{-3}$ nm$^2$/Hz. Thus, for a typical frame rate of 30 Hz, we can expect a tracking accuracy of 0.2, 0.2 and 10 nm for the $x-$, $y-$ and $z$-coordinate after cross-correlation with a single


a
k = 7 pix
k = 9.22 pix
k = 12.15 pix
k = 17 pix
Phase (rad)
b
Res. (nm)
Z_focus (μm)
c
σ_res. (nm)
k (pix)
d
Amplitude (A.U.)
Bead height (μm)
e
σ_z²/f_s (nm²/Hz)
Std (nm)
no. of polynomials
f
σ_z²/f_s (nm²/Hz)
Std (nm)
no. of periods

reference image.

Next, to illustrate the increased performance with an increased number of reference images, we plotted the phase calibration graphs obtained for four reference images, shown in Figure 3.3a. The $\varphi_n(z)$ curve converged to a single point following equation 5, which allowed us to unequivocally assign the focus offset. The deviations from linearity, starting 5 µm above focus, were fitted with a $5^{\text{th}}$ order polynomial over a range of 10 µm. The residuals of the fits give a good estimation of the dynamic range: the standard deviation, $\sigma_{\text{res}}$, was generally below 15 nm, as depicted in Figure 3.3b. We systematically tested the dependence of the accuracy of the 3DPT method by evaluating the standard deviation of the residuals as a function of the reference frequency. The lowest $\sigma_{\text{res}}$ was found for $k = 10$ pix. For $k < 7$ pix we did not obtain distinct correlation peaks, reflecting the diffraction-limited character of the holographic images. For $k > 10$ pix, $\sigma_{\text{res}}$ gradually increased up to 15 nm, which is still rather accurate for a dynamic range of 10 µm (Figure 3.3c). As could be expected from visual inspection of the holographic images, the further the bead was defocused, the less high-frequency information was obtained. This was reflected by the amplitudes of the cross-correlations, that were plotted in Figure 3.3d. Whereas for k=7 pix, the highest cross-correlation was obtained 10 µm above focus and vanished at 20 µm, larger periods peaked further from the focus. From Figure 3.3a-d it is clear that the dynamic range is not limited to 10 µm.

The periodic modulation in Figure 3.3b suggests that better accuracy could be obtained by fitting higher order polynomials. Increasing the order of the polynomials did decrease the average standard deviation of the residuals of the polynomial fit, although this was not reflected in the PSD analysis of the experimental data, shown in Figure 3.3e. This is probably due to the slow fluctuations that remain in the residual of the polynomial fit, which would be represented in the low-frequency part of the power spectrum.

The tracking accuracy in the $z$-direction could be increased by combining more reference images, at the cost of increased computational time. Figure 3.3f shows that the accuracy, as quantified from the PSD, increased about 10-fold when more reference images were used. The accuracy converged to $\sigma_z^2/f_s = 0.2$ nm$^2$/Hz for $n = 5$, implying 2.4 nm accuracy for 30 Hz imaging or 1 nm at 5 Hz. The average standard deviation of the residuals of the polynomial fit decreased similarly. For accuracy in the $x$- and $y$-direction, it was sufficient to use more than 2 reference images to achieve sub-nm accuracy at 30 Hz.

The computation speed was up to 10,000 diffraction patterns per second for $100 \times 100$ pix ROIs, even when 4 reference images were used. Note that the FFT of the reference images can be done prior to the tracking, so only $(n + 1)$ FFTs need to be computed in real-time for $n$ reference images. Subsequent processing, *i.e.* calculating the $xy$-position, computing the $z$-coordinate from the calculated phase and unwrapping the phase, was much faster. Thus, 3DPT

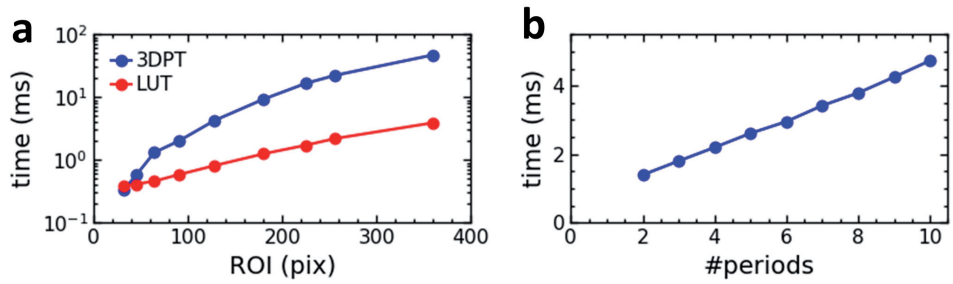



**Figure 3.4**
**3DPT can be computed within several milliseconds per ROI for typical imaging conditions.** a) Computation time increase rapidly with ROI size for 3DPT due to multiple 2D FFT computations. LUT tracking is generally several times faster. However for realistic ROI sizes, between 64 and 128 pixels, the processing time is less than 3 ms. These results were computed for 5 reference images in 3DPT and 64 radial profiles for LUT tracking. b) The time per ROI increases linearly with the number of reference images.

is a versatile, accurate and fast method for holographic tracking of microbeads with nanometer accuracy.

### Comparison with LUT tracking

In the most commonly used algorithm for camera based bead tracking, the height is interpolated from a look up table (LUT) of radial profiles that were pre-calibrated for a range of known bead heights. This provides a large increase of the processing rate as compared to LMST fitting. 3DPT is generally slower than LUT-tracking, as 2D FFTs scale with a ROI size of $N$ pixels as $N^2 \ln N$, and multiple 2D FFTs are required. LUT tracking on the other hand uses 1D FFTs, that scale as $N \ln N$, for $xy$-positioning and calculation of the radial profile scales with $N^2$. In Figure 3.4a, we compared the computation speed of 3DPT with a basic LUT-based tracking algorithm. Indeed, 3DPT is generally slower and computation time rises faster with larger ROIs. As the accuracy increases with the number of reference images, it is important to also quantify how this impacts the speed of the calculation. As expected, the time per ROI increases linearly with the number of reference images, *see* Figure 3.3b. For relevant ROI sizes, between 64 and 128 pixels, and 5 reference images, the processing time is less than 3 ms / ROI, which is slower than the LUT method, but sufficiently fast for real-time processing.

We tested both algorithms for accuracy using a set of typical imaging parameters (*see* Materials and Methods), by computing the difference between the input and output coordinates. Supplementary Figure 3.S2 shows for example that Poisson noise in the image hardly affected the accuracy and reproducibility of the $x$-, $y$- and $z$-coordinates for 3DPT, but introduced a systematic error

and increased variations for the $z$-coordinate in LUT-tracking. Note that the error in $z$ did not appear to be the result of inaccurate determination of the center of the bead, as $x$- and $y$-accuracy did not change with increasing image noise. Only a small fraction of the LUT coordinates (typically a few percent), featured large outliers, originated from wrongly assigned $x$ and $y$ coordinates. The simulated standard conditions did not include extreme excursions of the bead, or image distortions that happen frequently in experiments, so these errors can only be attributed to the limited range of validity of the 1D $xy$ tracking. More advanced LUT-tracking schemes [49] may alleviate this problem, at the cost of computational speed. Here we avoid the large impact of these outliers by presenting the tracking error in terms of median ± interquartile range, rather than average ± standard deviation. For 3DPT, we did not observe such outliers.

A large number of factors affect the tracking accuracy and these cannot always be varied in a systematic manner experimentally. Moreover, despite best efforts, it is hardly possible to exclude small variations in mechanical stability, bead size, illumination intensity and optical aberrations, which can have a major effect on the reproducibility of the tracking data. We therefore used LM theory to evaluate the effect of a number of parameters on the accuracy of both methods. Supplementary Figure 3.S2 shows the results of these simulations. Generally, both methods yielded accuracies and reproducibilities in the nanometer range, though 3DPT generally performed more accurately and yielded less fluctuations. $x$ and $y$ coordinates were more precise than the $z$-coordinate. Changing ROI sizes between 64 and 256 pixels hardly affected accuracy. Differences in magnification, represented by pixels size, did change the accuracy of 3DPT. However, this is expected as spatial frequencies in the image change with magnification, and these should be responded to by adjusting the chosen frequencies in 3DPT accordingly. When correctly chosen, tracking errors reduced to the nanometer range (data not shown). Changes of the coherence length of the illumination source and numerical aperture of the objective were not explicitly included in LMST. We approximated the combined effect of reducing these by applying a Hamming filter over the holographic image (Equation 3.11), resulting in a reduction of the number of diffraction rings. Images in which the width of this filter exceeded 2 µm yielded equally accurate results. Overall, this systematic comparison between LUT-tracking and 3DPT shows improved accuracy and reproducibility for the later under typical imaging conditions, at the cost of a 2-5 times increase in computation time.

### 3.2.4 Robustness of 3DTP

Large field of view imaging, which is a prerequisite for highly parallel tracking of multiple beads, often comes with image artefacts, as not the entire image

may be illuminated or imaged optimally. Supplementary Figure 3.S3 shows a typically field of view of our setup and highlights such artefacts. To evaluate the robustness of our tracking algorithm for common imaging artifacts, we simulated test data using LMST for 1.0 µm diameter beads that were moving randomly in three dimensions. We simulated the diffraction patterns 12 µm above focus, where the accuracy of LMST was the highest (Figure 3.1f). This approach allowed us to systematically introduce distortions and to compare tracking results with the known coordinates that were inserted into the LMST.

Figure 3.5a shows a simulated image in which we have introduced Poisson noise, representing the signal to noise of a typical camera. While the accuracy of the $x$- and $y$-coordinates was hardly affected and remained constant at 0.5 $nm^2/Hz$, the accuracy in the $z$-direction decreased from 0.05 to 10 $nm^2/Hz$ when the noise increased for 0 to 20 greyscale units in an 8-bit image. Typical experimental noise intensities ($\sim$ 33 greyscale units) resulted in 1 $nm^2/Hz$ accuracy in the $z$-direction, close to the experimentally obtained accuracy in Figures 3.3e-f.

The robustness of the algorithm was further tested in Figures 3.5b-e, in which we simulated several other distortions that are frequently observed in experimental imaging. Interlacing, which was prominent in analog CCD cameras, did not affect the tracking accuracy (Figure 3.5b). Distortions with lower spatial frequencies, such as light gradients, did decrease the accuracy of the calculated $z$-coordinate (Figure 3.5c). Astigmatism up to 15% resulted in moderate increases of the error (Figure 3.5d). Larger astigmatism was more problematic. However, such distortions should not occur in properly designed microscopes.

We also quantified the robustness of the tracking algorithm when the diffraction pattern was not fully contained within the ROI. Since there is always a trade-off between the size of the analyzed ROI and computation speed, it is in the interest of increasing throughput to reduce the ROI as much as possible. Moreover, restricting the ROI reduces the probability that other beads enter the ROI, which hampers accurate tracking. We observed that tracking accuracy was not affected until the bead was shifted by more than 30% out of the ROI (Figure 3.5e). In conclusion, our simulations showed that 3DPT was robust against many types of typical aberrations.

Finally, the performance of 3DPT was demonstrated experimentally by simultaneous measurement of the three-dimensional position of 10 immobilized 2.8 µm diameter paramagnetic beads. Drift and mechanical vibrations, which dominated our measurements, as reflected in the power spectral densities (Figure 3.2j), were characterized, isolated, and removed following the approach described in Materials and Methods. A representative time trace in three dimensions was shown in Figures 3.6a-c, yielded a standard deviation in the $x$- and the $y$-direction of 0.3 nm and in the $z$-direction 1.6 nm, roughly matching the simulated values.

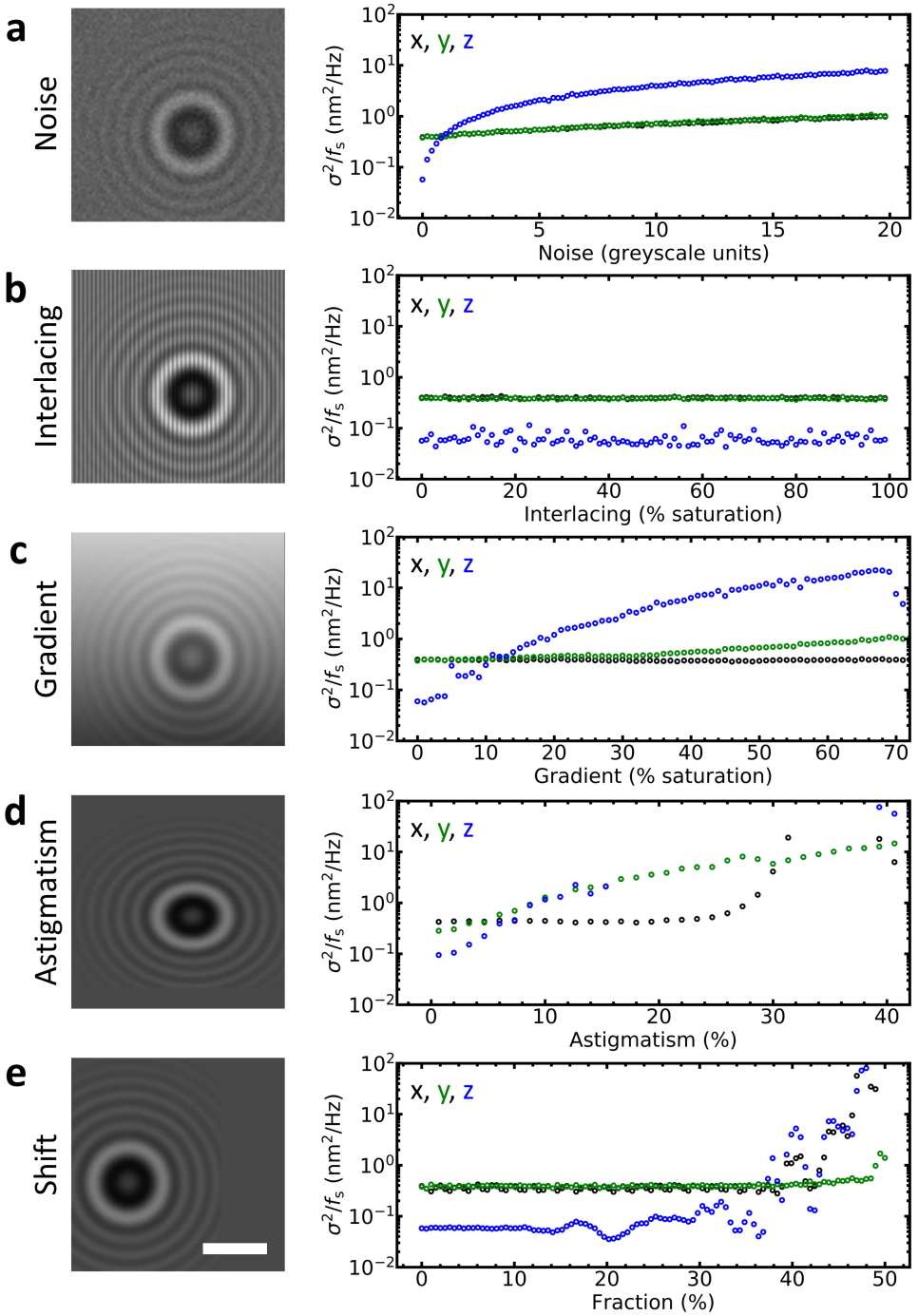

a
Noise
x, y, z
$\sigma^2/f_s$ (nm$^2$/Hz)
Noise (greyscale units)
b
Interlacing
x, y, z
$\sigma^2/f_s$ (nm$^2$/Hz)
Interlacing (% saturation)
c
Gradient
x, y, z
$\sigma^2/f_s$ (nm$^2$/Hz)
Gradient (% saturation)
d
Astigmatism
x, y, z
$\sigma^2/f_s$ (nm$^2$/Hz)
Astigmatism (%)
e
Shift
x, y, z
$\sigma^2/f_s$ (nm$^2$/Hz)
Fraction (%)

**Figure 3.5**
**3DPT is robust against image aberrations.** To evaluate the robustness of 3DPT, the diffraction patterns of 1.0 µm diameter paramagnetic beads were simulated with LMST. In the absence of image artifacts, the tracking accuracy of simulated images was $60 \times 10^{-3}$ nm$^2$/Hz in the $z$-direction. The diffraction patterns were superimposed with aberrations typically observed in MT and other bead tracking techniques. Note that the amplitude of the PSD remained below 100 nm$^2$/Hz in all cases, corresponding to 55 nm (0.5 pix) at a frame rate of 30 Hz, except for astigmatism exceeding 30%. Typical experimental values for our microscope were $\sim$ 3 greyscale units for Poisson noise, 0% for interlacing, $\sim$ 15% for light gradients, $\sim$ 2% for astigmatism, and $\sim$ 5% for shift. a) Poisson image noise mainly affected the accuracy of the $z$-coordinate and resulted in a 10-fold decrease in accuracy for an amplitude of 20 greyscale units. The tracking accuracy in the $x$- and $y$-direction only decreased a factor of 2. b) Interlacing did not affect the tracking accuracy. c) A light gradient in the ROI up to 20% still yielded an accuracy in the $z$-direction below 1 nm$^2$/Hz. The tracking accuracy was only slightly affected in the $y$-direction, the direction of the light gradient, while the $x$-direction was unaffected. When the light gradient exceeded 70%, no correlation peak was found. d) Astigmatism significantly affected tracking accuracy in three dimensions. e) The tracking method was unaffected when the bead center was shifted less than 30% out of the ROI. Exceeding a shift of approximately 40% resulted in the complete loss of tracking in both the $z$-direction and the $x$-direction (the direction that the bead was moved). Scale bar: 3 µm.

As an illustration of an application, we implemented our tracking algorithm to measure the unfolding of native chromatin, shown in Figure 3.6d (experimental details in [53]). Native chromatin is an example of a highly heterogeneous sample, where many molecules need to be measured to extract common features. We used 3DPT to measure the step size of the unwrapping of DNA from the histone core, using the method of the step size described by Kaczmarczyk [60] and summarized in Materials and Methods. As expected, 3DPT accurately revealed the characteristic 77 bp step size, corresponding to about 25 nm, shown in Figure 3.6e. Although the molecules were heterogeneous in composition and unfolded accordingly, the width of the individual steps occurring at forces above 5 pN could easily be resolved.

## 3.3 Discussion

Single-molecule biophysical techniques frequently employ bead tracking for the mechanical characterization of biomolecules. Here, we introduced 3DPT, a robust and non-iterative bead tracking algorithm for holographic imaging. It makes use of the circular symmetry of a holographic image of a bead and, similar to lock-in techniques, it selects a single spatial frequency for image analysis. Next to directly producing the $xy$-position, the height information is captured in a single parameter: the phase of the wave front of the diffraction pattern, which can be converted to a $z$-coordinate using calibration prior to the experiment. 3DPT was lightweight, robust against common aberrations, and yielded nanometer accuracy in three dimensions. We have implemented this

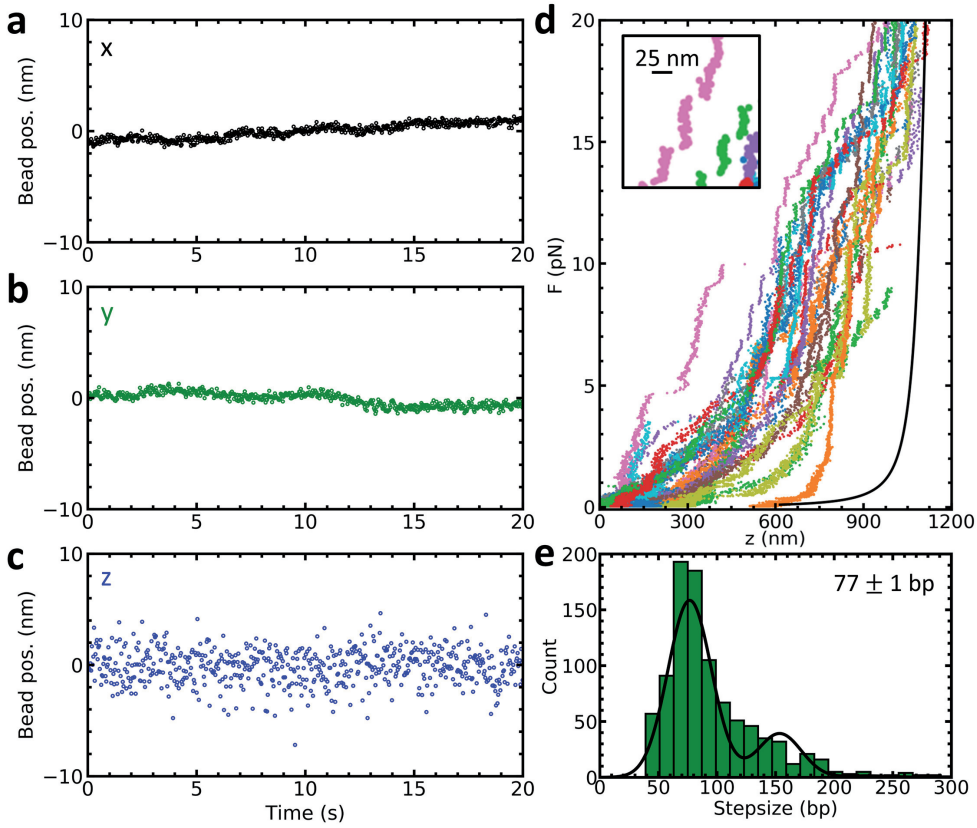



**Figure 3.6**

**Simultaneous nanometric tracking of multiple paramagnetic beads yields nanometer accuracy in three dimensions.** a-c) 10 immobilized 2.8 µm diameter paramagnetic beads (Dynabeads M270 Streptavidin, Thermo Fisher Scientific) were recorded for 20 seconds and analyzed using 3DPT using 15 polynomials during phase calibration and 15 reference images during tracking. Drift and mechanical vibrations were characterized, isolated, and removed following the approach described in Materials and Methods, and a typical trace was plotted in the $x$- (a), $y$- (b), and $z$-direction (c). Some residual drift that was present in all dimensions indicated imperfect immobilization of the beads. The distribution of the coordinates yielded a standard deviation of $\sigma_{xy} = 0.6$ nm in the $x$- and the $y$-direction and $\sigma_z = 1.6$ nm in the $z$-direction. d) 25 native chromatin molecules, reconstituted *in vivo* [53], were stretched and unfolded with magnetic tweezers. Although the complexes were heterogeneous, they unfolded in three distinct transitions [7]. The last transition, where the last singe turn of DNA unwraps from the histone core, takes place above $\sim 5$ pN (for native chromatin) and is recognized by its stepwise nature. e) The step size distribution was plotted in a histogram (bin width = 12 base pairs) and fitted with a triple Gaussian. The histogram contained data from 111 curves from which a selection was plotted in panel d. The step size was easily resolved with our approach, and measured $77 \pm 1$ base pairs, taking into account double and triple simultaneous steps.

algorithm in NI LabVIEW 2014 on a 10 core 2.8 GHz CPU and could track up to 10,000 diffraction patterns per second captured within a ROI of $100 \times 100$ pixels. 3DPT is robust against many common image artifacts, and can yield nanometer accuracy despite sub-optimal imaging conditions.

One could alternatively use 2D cross-correlation with a set of experimentally obtained reference images, which would include unknown variables such as optical aberrations, and bead dependent properties. As holographic images contain a range of spatial frequencies, the phase of such a cross-correlation is ill-defined and therefore only the amplitude can be used to determine the height, analogous with the 1D LUT method. Such an algorithm however, as compared to 3DPT, requires more 2D FFTs per calculation, making it much slower, and would also propagate errors in the reference images. We found that 3DPT is not only faster, but also more robust.

Due to drift and mechanical vibrations in our setup, we did not test whether 3DPT could resolve nanometer steps, as was demonstrated before [51]. Nevertheless, 25 nm steps, resulting from unwrapping single nucleosome were easily resolved. From PSD analysis, as well as unfiltered time traces, it is clear that 1 nm is close to the limit of the accuracy of our current experiments at 30 Hz bandwidth. 1 Hz averaging should be able to resolve nanometer changes in bead position. In the lateral direction, 3DPT performs an order of magnitude better. For many applications other than single-molecule force spectroscopy, the lateral resolution is as important as accuracy in the $z$-direction.

Traditional LUT based bead tracking algorithms proved to be faster than the more robust 3DPT and when imaging conditions can be optimized, they appear to yield more accurate tracking results. Computing the radial intensity profile is the most time-consuming step in LUT bead tracking algorithms and transferring this step to GPU can increase the processing speed [50]. We nevertheless do not expect a significant gain in speed by implementing 3DPT in GPU, as the current computation times are smaller than the time it takes for transferring the images from CPU to GPU. Avoiding computation of the radial intensity profile makes the tracking more robust: small errors in $xy$-position, as well as imaging aberrations, introduce large changes in this profile as compared to a reference image. In 3DPT, these artifacts predominantly reduce the amplitude of the cross-correlation, but have little effect on the position or the phase.

Because the computer-generated reference images that are used in 3DPT have a known center and lack noise and other artifacts, the algorithm is very robust. We therefore expect that 3DPT may be used for other applications than the tracking of spherical colloidal particles. This, however, should be tested for each particular application. In addition to bead tracking, we also used a cross-correlation with reference images for auto-focusing and for initial recognition of beads, relieving the operator from manually selecting beads. For multiplexed high-throughput applications, this can be a key advantage as the microscopy can be fully automated.

We could not obtain good fits of our super-paramagnetic beads to LMST, resulting in large tracking errors, typically hundreds of nanometers, which we tentatively attributed to the heterogeneous composition of the beads. This suggests that LMST fitting is a more viable option for, for example, TPM, AFS, and OT, that do not require magnetic beads. Though fitting imprecisions did not significantly affect the resulting position accuracy in the $xy$-direction, it was detrimental for tracking in the $z$-direction. Fits only converged in a limited range, especially in the case of the 2.8 µm beads. For MT, LMST fitting therefore may not only impede real-time processing, but it may also be inadequate for applications that require more than several micrometers range.

In previous work, the Grier group used a laser to create holographic images [33]. Most biophysical single-molecule studies (including ours), however, used a collimated LED to illuminate the sample. Due to the limited spatial coherence of a LED, the images do not feature speckle patterns, that were subtracted in studies using LMST fitting [61]. The limited coherence also reduces the range of the diffraction pattern, and laser illumination in combination with image background subtraction may further improve the accuracy of 3DPT by generating more contrast in the holographic image.

Multiplexing becomes increasingly important in single-molecule biophysics as more complex and more heterogeneous samples are investigated. A good example is our previous study, in which we performed force spectroscopy on natively assemble chromatin fibers [53]. Multiplexing allowed us to pick out hundreds of chromatin fibers that each unfolded differently, reflecting variations in composition and folding. Another example is a study on RNA polymerase pausing, in which multiplexing served to observe sufficient rare events to analyze the statistics [62]. Because 3DPT can easily replace traditional radial profile comparisons with LUTs, it may be adopted by many experimentalists.

Overall, its easy implementation, robustness, and nanometer accuracy over a wide $z$-range makes 3DPT an ideal method for multiplexed particle tracking applications. It may improve all single-molecule techniques that rely on bead tracking such as MT, TPM, AFS, and OT. By enhancing their throughput, it will help to realize one of the truly unique promises of single-molecule biophysics: detecting rare events in a large population of molecules with unprecedented resolution.

## 3.4 Materials and Methods

### 3.4.1 Magnetic tweezers setup

A multiplexed MT setup equipped with a NIKON CFI Plan Fluor objective MRH01401 (NA = 1.3, 40×, Oil, NIKON Corporation, Tokyo, Japan) was used to track paramagnetic beads. Samples were measured in custom-built flow cells mounted on a multi-axis piezo scanner P-517.3CL (Physik Instrumente GmbH & Co. KG, Karlsruhe, Germany). A field-of-view of $0.5 \times 0.5$ mm$^2$ was captured on 25 Mpix Condor camera (cmv5012-F30-S-M-P8, CMOS Vision GmbH, Schaffhausen, Switzerland) using an infinity-corrected tube lens ITL200 (Thorlabs, Newton, USA). The camera was read out by a PCIe-1433 frame grabber (National Instruments, Austin, USA) integrated with a T7610 PC (Dell, Round Rock, USA) equipped with a 10-core Intel Xeon 2.8 GHz processor (E5-2680 v2, Intel, Santa Clara, USA) and 32GB DDR3 memory. The setup measured the full frame at 30 frames per second. Each pixel measured 112 nm in this configuration. The flow cell was illuminated with a 100 mW 645 nm LED-collimator-packaged (LED-1115-ELC-645-29-2, IMM Photonics GmbH, Unterschleißheim, Germany).

### 3.4.2 Software

All MT control software, LMST fitting tools, and tracking software were written in LabVIEW (National Instruments, Austin, USA). The tracking code and the simulation program is available on Github (`https://github.com/JvN2/3DPT`).

### 3.4.3 Spectral analysis of tracking accuracy

A solution 1 pg/µl 2.8 µm diameter paramagnetic beads was deposited onto a cover slide and heated to 95°C for several minutes to melt the beads to the glass. Subsequently, the cover slide with immobilized beads was mounted into the flow cell and placed onto the setup. Immobilized beads were tracked for 120 seconds. To obtain $\sigma^2/f_s$, the PSD was calculated and fitted with a horizontal line for $f > 5$ Hz.

### 3.4.4 Lorentz–Mie scattering theory

LMST fitting was implemented following [33, 54, 55]. The diffraction pattern $I_{\mathrm{LMST}}(\rho)$ results from the interference between the incident field $E_0(r)$ and the field scattered off the particle $E_S(r)$:

$$I_{\mathrm{LMST}}(\rho) = |E_S(r) + E_0(r)|^2 \Big|_{z=0}. \tag{3.6}$$

The incident field was described by:

$$E_0(\rho) = u_0(\rho)\exp(ikz)\hat{\epsilon}, \tag{3.7}$$

where the incident field was uniformly polarized in the $\hat{\epsilon}$-direction so that amplitude $u_0(\rho)$ at position $\rho = (x, y)$ in plane $z = z_p$ of the particle is equal to that in the focal plane $z = 0$. The wavenumber of the propagating wave was $k = 2\pi n_m/\lambda$, where $n_m$ was the refractive index of the medium and $\lambda$ was the wavelength of the light in vacuum.

The scattered field was described by:

$$E_S(\rho) = \alpha \cdot \exp(-ikz_p)u_0(\rho)f_s(\rho - \rho_p), \tag{3.8}$$

where $f_s(\rho)$ was the LMST function which depended on bead radius $a$, $n_p$, $n_m$, and $\lambda$. Scaling coefficient $\alpha \approx 1$, and accounted for variations in the illumination.

The diffraction pattern was normalized in the $z = 0$ plane by $\beta$:

$$\frac{I_{\mathrm{LMST}}(\rho)}{|u_0(\rho)|^2} \equiv \beta \cdot I_{\mathrm{LMST}}(\rho) = \beta\left[1 + 2\Re\{E_S(r) \cdot E_0^*(r)\} + |E_S(r)|^2\right]. \tag{3.9}$$

The normalized image scaled with the calculated Mie scattering pattern $f_s(r)$ by:

$$I_{\mathrm{LMST}}(\rho) \approx \beta\left[1 + 2\alpha\Re\Big\{f_s(r - r_p) \cdot \hat{\epsilon}\exp(-ikz_p)\Big\} + \alpha^2\left|f_s(r - r_p)\right|^2\right] \cdot f_\gamma(r), \tag{3.10}$$

where $f_\gamma(r)$ is a Hamming filter of width $\gamma$ which represented the decay of the diffraction pattern from the center due to the limited spatial coherence of the light source:

$$f_\gamma(r) = \left(\cos\left(\frac{\pi}{\gamma}r\right) + 1\right)^2 \tag{3.11}$$

Fitting Equation 3.10 yielded the physical parameters $a$, $n_p$, $n_m$, $x$, $y$, $z$, and scaling coefficients $\alpha$, $\beta$, and $\gamma$. The three-dimensional position and radius $a$ of the bead was typically fitted with nanometer precision. The refractive index $n_p$ was reproducible between beads within one part in a thousand [33].

### 3.4.5 LUT-tracking

A standard LUT algorithm was implemented as follows: in both $x$ and $y$ directions, the central 5 lines of a ROI were summed. These line traces were correlated with their corresponding mirrored trace. The bead center was assigned to the maximum correlation, as interpolated by quadratic fitting. Using this bead center, a radial intensity profile of half the ROI size was computed. The $z$-coordinate was assigned to the minimum of the Root Mean

Squared Difference with a set of radial profiles that were previously computed for a range of bead heights, again interpolated by fitting a quadratic curve. For comparison with 3DPT, we used the following standard parameters: ROI size = 100 pixels, number of radial profiles = 64, $z$-range for calibration was 0 to 15 µm, $x$ and $y$ offset = 0 µm, Poisson noise = 0 greyscale units. As standard tracking simulation consisted of 256 frames, bead diameter = 1 µm, a linear $z$-ramp from 4 to 7 µm, $x$ and $y$ offset = 0 µm, Poisson noise = 5 greyscale units and $x$-, $y$-, and $z$-positions that were drawn from a normal distribution with a width of 100 nm.

### 3.4.6 Robustness simulations

We simulated movies of beads randomly moving in three dimensions using LMST using the following parameters: $a = 0.5$ µm, $n_p = 1.9$, $n_m = 1.33$, $\alpha = 1.0$, $\beta = 54$, $\gamma = 45$, and $\lambda = 645$ nm. This approach yielded realistic diffraction patterns of 1.0 µm diameter paramagnetic beads. Parameters $n_p$, $\alpha$, $\beta$, and $\gamma$ were average values obtained from fitting the experimental patterns of 186 separate beads (data not shown). For these simulations, the phase was calibrated using 15 polynomials and 15 reference periods. For every simulation, we computed 3600 holograms ($150 \times 150$ pixels) of beads randomly moving in three dimensions ($dx$, $dy = 0 \pm 5$ nm, $dz = 12,000 \pm 5$ nm). The simulations were equivalent to a 120 seconds measurement on a 30 Hz camera. The simulated beads were tracked using 3DPT, and the time traces were converted to PSD to extract $\sigma^2/f_s$, similar to experimental data.

Five common aberrations were superimposed on the simulated diffraction patterns. Poisson noise was added to the diffraction pattern. Interlacing was simulated by multiplying every other row of pixels with a gain factor. 100% interlacing corresponded to a gain of 2. Light gradients were simulated by adding a slope in the $y$-direction of the diffraction pattern. 100% light gradients corresponded to a curve which rose to a maximum background intensity of 255. Astigmatism was simulated by resampling the columns in the $y$-direction over a smaller number of pixels. 100% astigmatism corresponded to an aspect ratio of 2. Shift out of the ROI was simulated by moving the bead in the $x$-direction until the average bead center was shifted by 50% of the ROI.

### 3.4.7 Characterization of mechanical vibrations and drift

To characterize 3DPT accuracy experimentally, time traces of 10 immobilized beads were recorded simultaneously. The effects of mechanical vibrations and drift, were largely removed by averaging these traces and subtracting it after applying a 1 Hz low-pass filter. The standard deviations of the resulting time trace yield the experimental tracking accuracy.

### 3.4.8 Quantification of the step size of unwrapping of DNA from the histone core

The length of the stepwise unwrapping of DNA from native nucleosome cores was measured using a method developed by Kaczmarczyk, described in detail in [60]. In short: each data point in the force-extension curve was compared to the theoretical extension of a given a contour length of free DNA, following a Worm Like Chain. The theoretical standard deviation for each point was computed using equipartition theorem and the derivative of the force-extension relation of the wormlike chain and supplemented by the tracking error. Next, the $z$-score and corresponding probability that the data point belonged to this contour length was calculated. The probabilities for all data points at given contour length were summed and this procedure was iterated for all contour lengths between 0 and the contour length of the DNA substrate. Peaks in the plot of the summed probability as a function of contour length were attributed to a stable state of unfolding of the chromatin fibers and distances between neighboring peaks reflect single unwrapping events.

## 3.5 Acknowledgments

The authors would like to thank Artur Kaczmarczyk for valuable discussions during the development of the algorithm. This work is part of the research program VICI with project number 680-47-616, which is (partly) financed by the Netherlands Organization for Scientific Research.

## 3.6 Supplementary Material

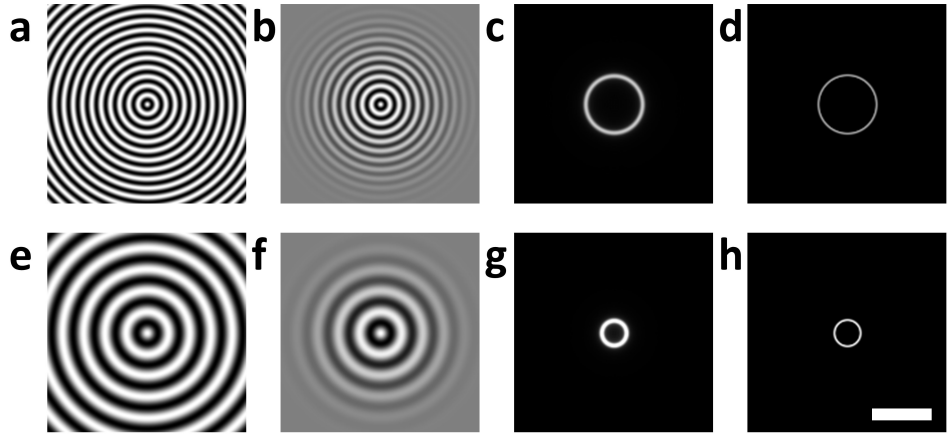
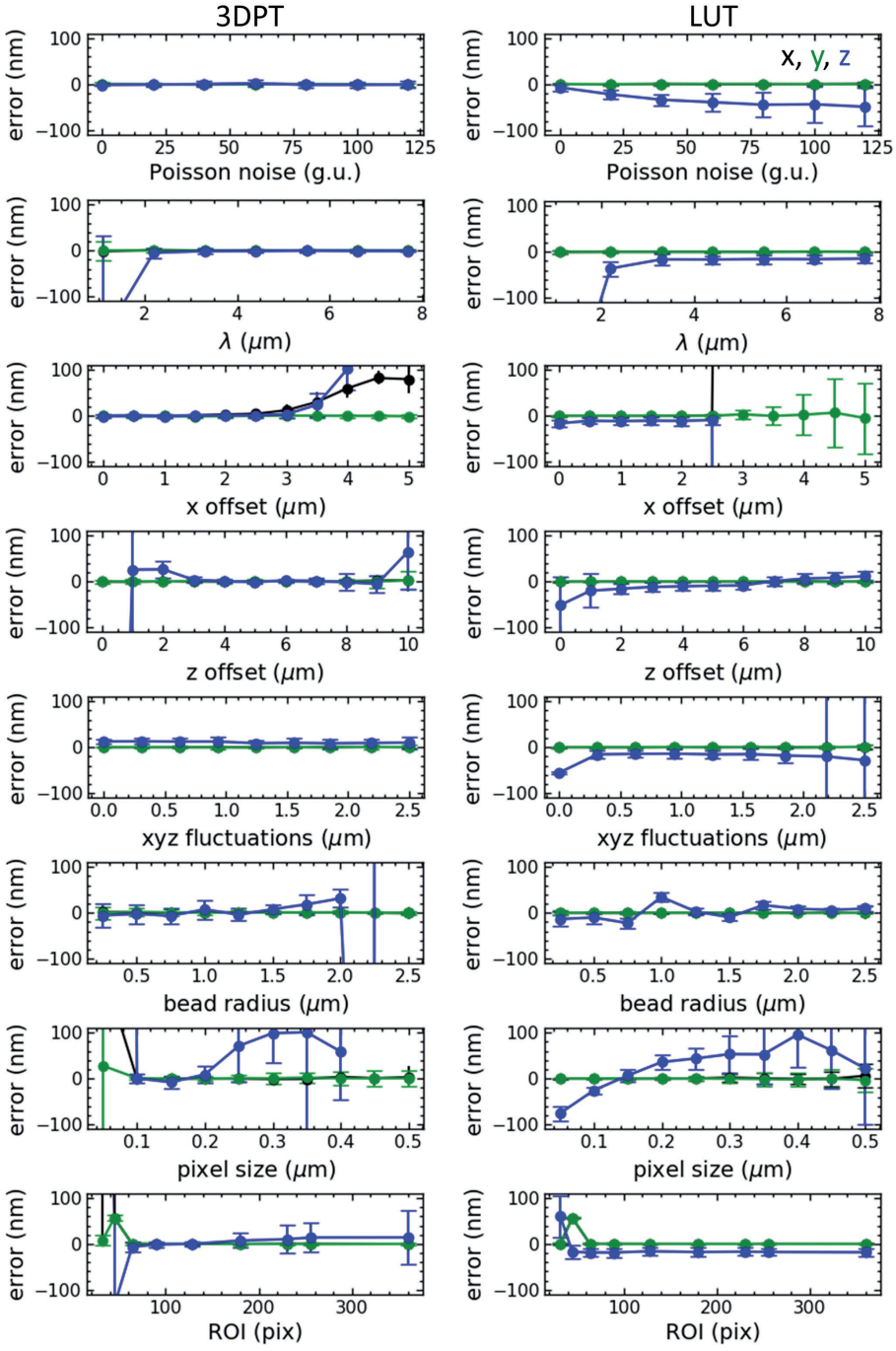



**Figure 3.S1**
**Reference images and corresponding band-pass filters for 3DPT.** A reference image containing a single spatial frequency of period $k_A = 7$ pix (a) was filtered by a Hamming filter (b). The reference image was transformed into Fourier space (c), and filtered by a frequency band-pass filter (d). Panel e-h depicted a reference image containing a spatial frequency of period $k_B = 16$ pix. Scale bar: 3 µm.


3DPT
LUT
x, y, z
error (nm)
Poisson noise (g.u.)
λ (μm)
x offset (μm)
z offset (μm)
xyz fluctuations (μm)
bead radius (μm)
pixel size (μm)
ROI (pix)

### Figure 3.S2

**Comparison between 3DPT and LUT accuracy.** The accuracies as a function of several parameters that define a tracking experiment were computed using LMST simulations. We used the following standard parameters: ROI size = 100 pixels, pixel size 0.11 µm, number of radial profiles = 64, $z$-range for calibration was 0 to 15 µm, $x$ and $y$ offset = 0 µm, Poisson noise = 0 greyscale units. A standard tracking simulation consisted of 256 frames, a bead diameter of 1 µm, a linear $z$-ramp from 4 to 7 µm, $x$ and $y$ offset = 0 µm, Poisson noise = 5 greyscale units and $x$, $y$, and $z$ positions that were drawn from a normal distribution with a width of 100 nm. Points represent median difference between input coordinates and computed coordinates. Error bars represent inter quartile ranges of the fluctuations within a tracking simulation. Means and standard deviations were disproportionally affected by rare outliers in the LUT tracking. $x$ and $y$ data generally overlap and demonstrate the reproducibility of these simulations. Graphs on the left represent 3DPT results, graphs on the righ trepresent LUT-tracking results. Poisson noise hardly affected 3DPT results, but introduced a systematically smaller bead height using LUT tracking. Coherence length of the illumination source and NA of the objective were effectively included in the width of the Hamming filter $\gamma$, and did not affect tracking accuracy of 3DPT when larger than 2 µm. For LUT, a similar trend was found, but the $z$-coordinate was again underestimated. Offsets in the imaging plane did not affect accuracy up to 3 µm for 3DPT, and was inhibiting LUT tracking. In the $z$-direction, 3DPT was accurate between 3 and 9 µm above focus. Here LUT tracking appeared robust over a larger range, but accuracy was a bit decreased. Simultaneous fluctuations in all 3 directions, mimicking more realistic experiments, did not affect 3DPT accuracy within 2.5 µm, whereas LUT tracking produced large variations in accuracy beyond 2 µm. For bead radii larger than 2 µm we observed reduced accuracy in 3DPT, which resulted from a more intricate diffraction pattern in close vicinity of the bead (data not shown). To remedy this, we computed images at 8 µm above focus, rather than 4 µm, which resulted in less convolved holograms. As a result 3DPT performed well again, up to beads with a radius larger than 2 µm. LUT tracking proved to be more flexible in terms of bead size, though slightly less accurate. Image magnification also affected tracking accuracy for 3DPT. This is not surprising as the spatial frequencies in the holographic image directly scale with magnification. The chosen frequencies were optimized for a magnification corresponding to 0.11 µm per pixel. Larger magnifications yielded poorer accuracy. However, this can be remedied by scaling the spatial frequencies in the reference images accordingly. Somewhat surpassingly, the LUT method also performed less with increasing magnification. This can probably be attributed to the reduced number of fringes that are captured in a fixed sized ROI. Finally, ROI sizes above 64 pixels were sufficient for both methods to yield accurate results. Note however that 3DPT became less accurate for very large ROIs, presumably because of the increased influence of Poisson noise when more pixels are included. Apparently LUT tracking is better at averaging out these fluctuations. Overall, both methods can achieve similar accuracies over a large range of parameters. 3DPT appeared more robust, yielding $< 10$ nm accuracy over a broader range of parameters, and featured smaller systematic errors.

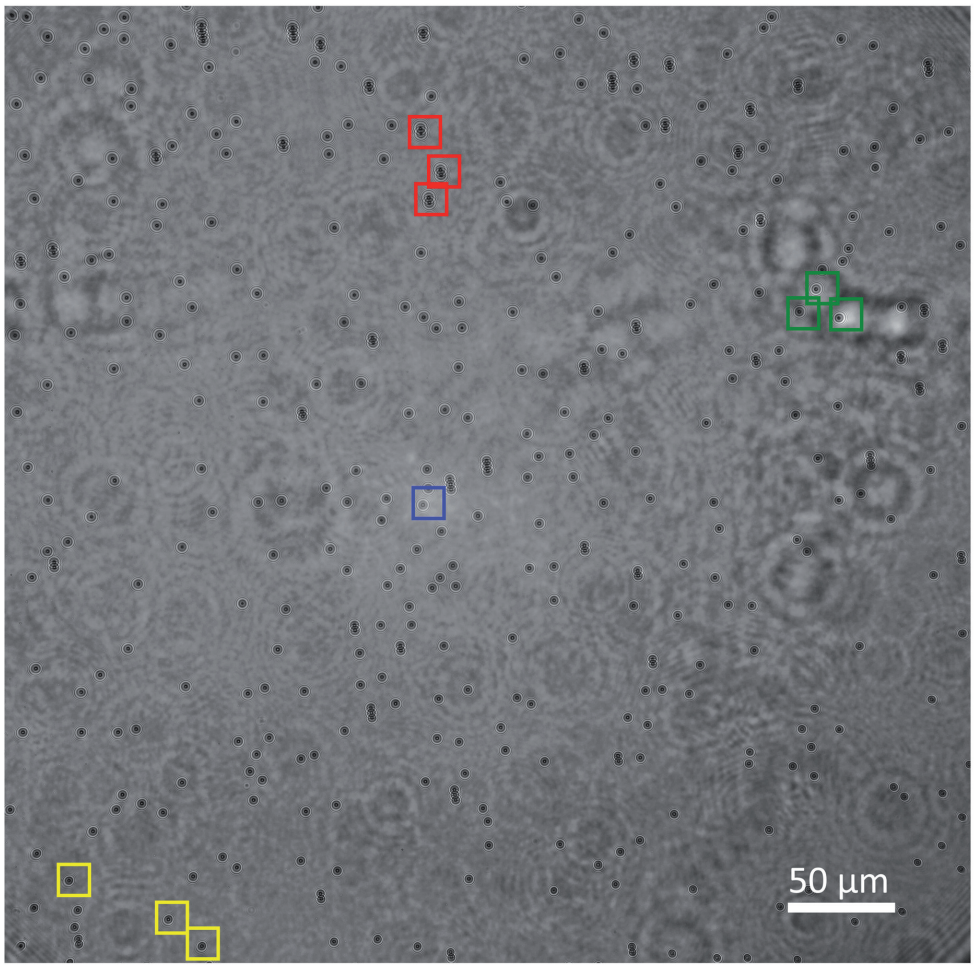



**Figure 3.S3**
**Typical image artefacts in large area imaging.** This image shows a full 4096 × 4096 pixels field of view of tethered beads in our MMT microscope. Though holographic images of the beads are well resolved throughout the image, inhomogeneous illumination and image distortions deform the bead images. In particular spherical aberrations dominate at the edges, leading to elongated diffraction patterns (yellow). The center of the field of view was somewhat over exposed, reducing the contrast in the central ROIs (blue). Loose beads, well above the focus plane, obstruct illumination and cause a gradient in the background (green). 3DPT is quite tolerant for such distortions (*see* Figure 3.5 in the main text). Aggregated beads (red) however, not only affect tracking, but also result in increased forces that are applied to the tether. These beads should be discarded from further analysis. Surprisingly, bead aggregation were not easily identified from 3DPT traces, which were remarkably accurate and error free for small aggregates and should therefore be identified directly in the image. Overall, 3DPT could accurately track all beads in this image.

# Bibliography III


[1] A Nagy K. Neuman, Keir C Neuman, and Attila Nagy. "Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy". In: *Nature methods* 5.6 (2008), pp. 491–505. ISSN: 1548-7091. DOI: 10.1038/NMETH.1218. eprint: NIHMS150003.

[2] Carlos Bustamante, Zev Bryant, and Steven B. Smith. "Ten years of tension: single-molecule DNA mechanics". In: *Nature* 421.6921 (2003), pp. 423–427. ISSN: 0028-0836. DOI: 10.1038/nature01405.

[3] G J Wuite et al. "An integrated laser trap/flow control video microscope for the study of single biomolecules." In: *Biophysical journal* 79.2 (Aug. 2000), pp. 1155–67. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(00)76369-7.

[4] C G Baumann et al. "Ionic effects on the elasticity of single DNA molecules." In: *Proceedings of the National Academy of Sciences of the United States of America* 94.12 (1997), pp. 6185–90. ISSN: 0027-8424. DOI: 10.1073/pnas.94.12.6185.

[5] Steven B. Smith, Yujia Cui, and Carlos Bustamante. "Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules". In: *Science* 271.5250 (1996), pp. 795–798. ISSN: 0036-8075. DOI: 10.1126/science.271.5250.795. eprint: arXiv:1011.1669v3.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] Behrouz Eslami-Mossallam, Helmut Schiessel, and John van Noort. "Nucleosome dynamics: Sequence matters". In: *Advances in colloid and interface science* 232 (2016), pp. 101–113.

[11] Artur Kaczmarczyk et al. “Probing Chromatin Structure with Magnetic Tweezers”. In: *Nanoscale Imaging*. Springer, 2018, pp. 297–323.

[12] Ramon A van der Valk et al. “Mechanism of environmentally driven conformational changes that modulate H-NS DNA bridging activity”. In: *eLife* 6 (2017). ISSN: 2050-084X. DOI: 10.7554/eLife.27369.

[13] Bram Henneman et al. “Quantitation of DNA-binding affinity using Tethered Particle Motion”. In: *Bacterial Chromatin*. Springer, 2018, pp. 257–275.

[14] J. van Noort et al. “Dual architectural roles of HU: Formation of flexible hinges and rigid filaments”. In: *Proceedings of the National Academy of Sciences* 101.18 (2004), pp. 6969–6974. ISSN: 0027-8424. DOI: 10.1073/pnas.0308230101.

[15] Martijn S Luijsterburg et al. “The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes”. In: *Critical reviews in biochemistry and molecular biology* 43.6 (2008), pp. 393–418.

[16] Thomas B Brouwer et al. “Unraveling DNA Organization with Single-Molecule Force Spectroscopy Using Magnetic Tweezers”. In: *Bacterial Chromatin*. Springer, 2018, pp. 317–349.

[17] T. R. Strick et al. “The elasticity of a single supercoiled DNA molecule.” In: *Science (New York, N.Y.)* 271.5257 (1996), pp. 1835–1837. ISSN: 0036-8075. DOI: 10.1126/science.271.5257.1835.

[18] Terence Strick et al. “Twisting and stretching single DNA molecules”. In: *Progress in biophysics and molecular biology* 74.1-2 (2000), pp. 115–140.

[19] He Meng et al. “Coexistence of Twisted, Plectonemic, and Melted DNA in Small Topological Domains”. In: *Biophysical Journal* 106.5 (2014), pp. 1174–1181. ISSN: 00063495. DOI: 10.1016/j.bpj.2014.01.017.

[20] Franziska Kriegel, Niklas Ermann, and Jan Lipfert. “Probing the mechanical properties, conformational changes, and interactions of nucleic acids with magnetic tweezers”. In: *Journal of structural biology* 197.1 (2017), pp. 26–36.

[21] Ineke Brouwer et al. “Human RAD52 captures and holds DNA strands, increases DNA flexibility, and prevents melting of duplex DNA: implications for DNA recombination”. In: *Cell reports* 18.12 (2017), pp. 2845–2853.

[22] Ineke Brouwer et al. “Two distinct conformational states define the interaction of human RAD51-ATP with single-stranded DNA”. In: *The EMBO journal* 37.7 (2018).

[23] Ineke Brouwer et al. “Sliding sleeves of XRCC4–XLF bridge DNA and connect fragments of broken DNA”. In: *Nature* 535.7613 (2016), p. 566.

[24] Gerrit Sitters et al. “Acoustic force spectroscopy.” In: *Nature methods* 12.1 (2015), pp. 47–50. ISSN: 1548-7105. DOI: 10.1038/nmeth.3183.

[25] Douwe Kamsma et al. “Tuning the Music: Acoustic Force Spectroscopy (AFS) 2.0”. In: *Methods* 105 (2016), pp. 26–33. ISSN: 10959130. DOI: 10.1016/j.ymeth.2016.05.002.

[26] Douwe Kamsma and Gijs JL Wuite. “Single-molecule measurements using acoustic force spectroscopy (AFS)”. In: *Single Molecule Analysis*. Springer, 2018, pp. 341–351.

[27] Dan Song et al. “Tethered particle motion with single DNA molecules”. In: *American Journal of Physics* 83.5 (2015), pp. 418–426. ISSN: 0002-9505. DOI: 10.1119/1.4902187.

[28] Ramon A Van Der Valk et al. “Genomic looping: a key principle of chromatin organization”. In: *Journal of molecular microbiology and biotechnology* 24.5-6 (2014), pp. 344–359.

[29] Yu-Long Sun et al. “Stretching type II collagen with optical tweezers.” In: *Journal of biomechanics* 37.11 (Nov. 2004), pp. 1665–9. ISSN: 0021-9290. DOI: 10.1016/j.jbiomech.2004.02.028.

[30] Marjan Shayegan et al. “Probing multiscale mechanics of collagen with optical tweezers”. In: *Optical Trapping and Optical Micromanipulation X*. Vol. 8810. International Society for Optics and Photonics. 2013, 88101P.

[31] Naghmeh Rezaei et al. “Using optical tweezers to study mechanical properties of collagen”. In: *Photonics North 2011*. Vol. 8007. International Society for Optics and Photonics. 2011, 80070K.

[32] Douwe Kamsma et al. “Single-Cell Acoustic Force Spectroscopy: Resolving Kinetics and Strength of T Cell Adhesion to Fibronectin”. In: *Cell Reports* 24.11 (Sept. 2018), pp. 3008–3016. ISSN: 2211-1247. DOI: 10.1016/J.CELREP.2018.08.034.

[33] Sang-Hyuk Lee et al. “Characterizing and tracking single colloidal particles with video holographic microscopy.” In: *Optics express* 15.26 (2007), pp. 18275–18282. ISSN: 1094-4087. DOI: 10.1364/OE.15.018275. arXiv: 0712.1738.

[34] Astrid van der Horst and Nancy R Forde. “Power spectral analysis for optical trap stiffness calibration from high-speed camera position detection with limited bandwidth.” In: *Optics express* 18.8 (2010), pp. 7670–7677. ISSN: 1094-4087. DOI: 10.1364/OE.18.007670.

[35] Astrid van der Horst and Nancy R Forde. “Calibration of dynamic holographic optical tweezers for force measurements on biomaterials.” In: *Optics express* 16.25 (Dec. 2008), pp. 20987–1003. ISSN: 1094-4087.

[36] Zhongbo Yu et al. “A force calibration standard for magnetic tweezers”. In: *Review of Scientific Instruments* 85.12 (2014). ISSN: 10897623. DOI: 10.1063/1.4904148.

[37] Peter Daldrop et al. “Extending the range for force calibration in magnetic tweezers”. In: *Biophysical Journal* 108.10 (2015), pp. 2550–2561. ISSN: 15420086. DOI: 10.1016/j.bpj.2015.04.011.

[38] Bob M Lansdorp et al. “A high-speed magnetic tweezer beyond 10,000 frames per second”. In: *Review of Scientific Instruments* 84.4 (2013), p. 044301.

[39] Noah Ribeck and Omar A. Saleh. “Multiplexed single-molecule measurements with magnetic tweezers”. In: *Review of Scientific Instruments* 79.9 (2008). ISSN: 00346748. DOI: 10.1063/1.2981687.

[40] Iwijn De Vlaminck et al. “Highly parallel magnetic tweezers by targeted DNA tethering”. In: *Nano Letters* 11.12 (2011), pp. 5489–5493. ISSN: 15306984. DOI: 10.1021/nl203299e.

[41] Iwijn de Vlaminck et al. "Magnetic forces and DNA mechanics in multiplexed magnetic tweezers". In: *PLoS ONE* 7.8 (2012). ISSN: 19326203. DOI: 10.1371/journal.pone.0041432.

[42] Richik N Ghosh and Watt W Webb. "Automated detection and tracking of individual and clustered cell surface low density lipoprotein receptor molecules". In: *Biophysical journal* 66.5 (1994), pp. 1301–1318.

[43] G M Lee, A Ishihara, and K Jacobson. "Direct observation of brownian motion of lipids in a membrane." In: *Proceedings of the National Academy of Sciences of the United States of America* 88.14 (1991), pp. 6274–8. ISSN: 0027-8424. DOI: 10.1073/pnas.88.14.6274.

[44] Lisa Gottesfeld Brown. "A survey of image registration techniques". In: *ACM computing surveys (CSUR)* 24.4 (1992), pp. 325–376.

[45] Akihiro Kusumi, Yasushi Sako, and Mutsuya Yamamoto. "Confined lateral diffusion of membrane receptors as studied by single particle tracking (nanovid microscopy). Effects of calcium-induced differentiation in cultured epithelial cells". In: *Biophysical journal* 65.5 (1993), pp. 2021–2040.

[46] Jeff Gelles, Bruce J. Schnapp, and Michael P. Sheetz. "Tracking kinesin-driven movements with nanometre-scale precision." In: *Nature* 331.6155 (1988), pp. 450–453. ISSN: 0028-0836. DOI: 10.1038/331450a0. arXiv: arXiv:1011.1669v3.

[47] M K Cheezum, W F Walker, and W H Guilford. "Quantitative comparison of algorithms for tracking single fluorescent particles." In: *Biophysical journal* 81.4 (2001), pp. 2378–2388. ISSN: 00063495. DOI: 10.1016/S0006-3495(01)75884-5.

[48] Charlie Gosse and Vincent Croquette. "Magnetic tweezers: micromanipulation and force measurement at the molecular level." In: *Biophysical Journal* 82.6 (2002), pp. 3314–3329. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(02)75672-5.

[49] Marijn T.J. J Van Loenhout et al. "Non-bias-limited tracking of spherical particles, enabling nanometer resolution at low magnification". In: *Biophysical Journal* 102.10 (2012), pp. 2362–2371. ISSN: 00063495. DOI: 10.1016/j.bpj.2012.03.073.

[50] J. P. Cnossen, D. Dulin, and N. H. Dekker. "An optimized software framework for real-time, high-throughput tracking of spherical beads". In: *Review of Scientific Instruments* 85.10 (2014). ISSN: 10897623. DOI: 10.1063/1.4898178.

[51] David Dulin et al. "High Spatiotemporal-Resolution Magnetic Tweezers: Calibration and Applications for DNA Dynamics". In: *Biophysical Journal* 109.10 (2015), pp. 2113–2125. ISSN: 15420086. DOI: 10.1016/j.bpj.2015.10.018.

[52] Fan Tso Chien and Thijn Van Der Heijden. "Characterization of nucleosome unwrapping within chromatin fibers using magnetic tweezers". In: *Biophysical Journal* 107.2 (2014), pp. 373–383. ISSN: 15420086. DOI: 10.1016/j.bpj.2014.05.036.

[53] Nicolaas Hermans et al. "Toehold-enhanced LNA probes for selective pull down and single-molecule analysis of native chromatin". In: *Scientific reports* 7.1 (2017), p. 16721.

[54] Gustav Mie. "Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen". In: *Annalen der physik* 330.3 (1908), pp. 377–445.

[55] Jianqi Shen. "Algorithm of Numerical Calculation on Lorentz Mie Theory". In: *PIERS Online* 1.6 (2005), pp. 691–694. ISSN: 1931-7360. DOI: 10.2529/PIERS050127104820.

[56] Mark D Pritt and Jerome S Shipman. "Least-squares two-dimensional phase unwrapping using FFT's". In: *IEEE Transactions on geoscience and remote sensing* 32.3 (1994), pp. 706–708.

[57] Dennis C Ghiglia and Mark D Pritt. *Two-dimensional phase unwrapping: theory, algorithms, and software*. Vol. 4. Wiley New York, 1998.

[58] Dennis C Ghiglia and Louis A Romero. "Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods". In: *JOSA A* 11.1 (1994), pp. 107–117.

[59] Donald J Bone. "Fourier fringe analysis: the two-dimensional phase unwrapping problem". In: *Applied optics* 30.25 (1991), pp. 3627–3632.

[60] A Kaczmarczyk. "Nucleosome stacking in chromatin fibers probed with single-molecule force-and torque-spectroscopy". PhD thesis. Leiden University, 2019.

[61] Fook Chiong Cheong, Bhaskar Jyoti Krishnatreya, and David G Grier. "Strategies for three-dimensional particle tracking with holographic video microscopy". In: *Optics express* 18.13 (2010), pp. 13563–13573.

[62] David Dulin et al. "Elongation-competent pauses govern the fidelity of a viral RNA-dependent RNA polymerase". In: *Cell reports* 10.6 (2015), pp. 983–992.

# 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

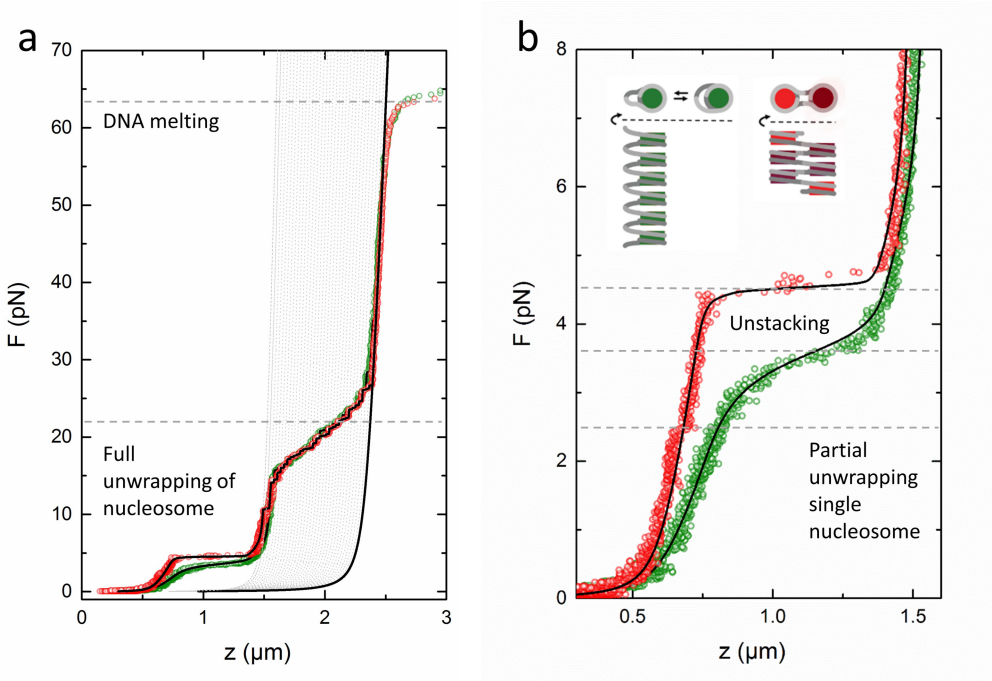



**Figure 4.1**
**Experimental force-extension curves of individual folded 167 NRL (red) and 197 NRL (green) chromatin fibers show a difference in stretching and unfolding at forces below 5 pN.** a) At 65 pN the curves exceed the extension of a Worm Like Chain (WLC) as the DNA overstretches. Multiple 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

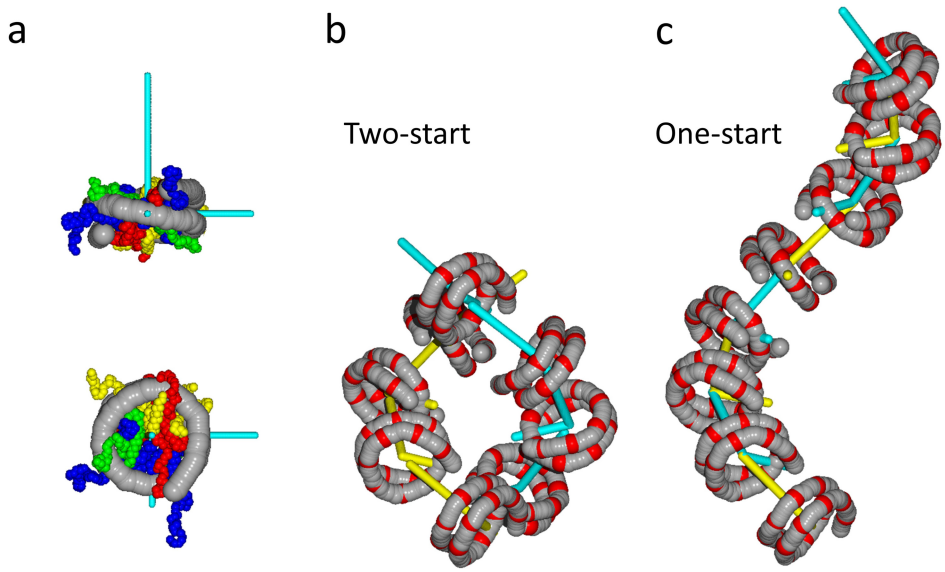



**Figure 4.2**
**Two additional layers of step parameters define the folded chromatin fiber.** a) A nucleosome frame (cyan) was created based on the coordinates of the 14 base pairs that are fixed in the nucleosome (colored red in panels b and c), as determined from the local minima of the average B-factors 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

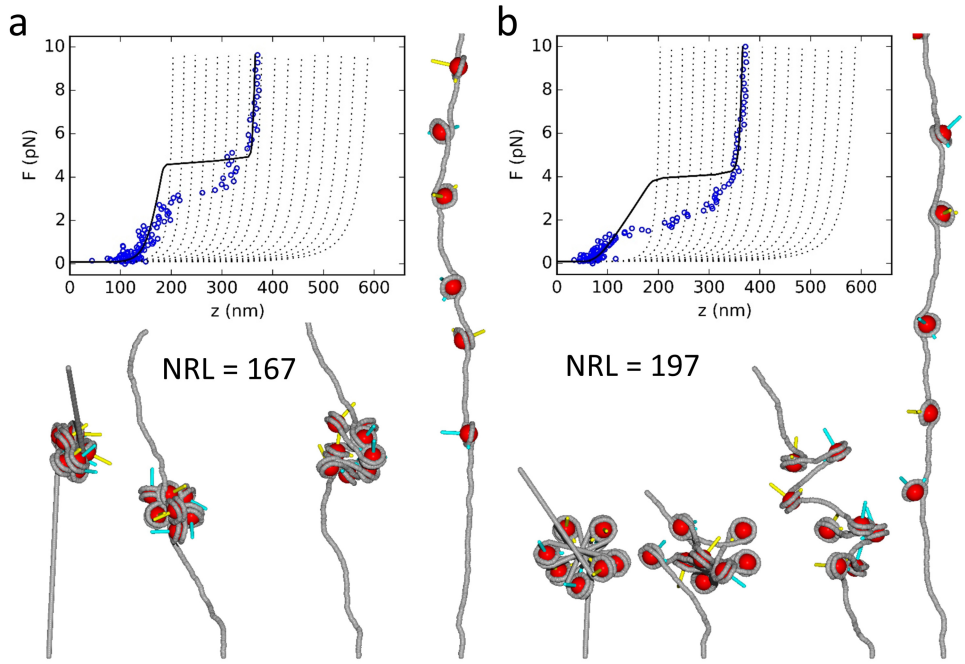



**Figure 4.3**
**MC simulations of arrays of non-interacting nucleosomes that readily unfold when the force is increased up to 10 pN.** a) A simulated force-extension curve of an 1800 base pairs 167 NRL fiber containing 8 nucleosomes. b) Simulated force-extension curve of an 1800 base pairs 167 NRL fiber 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.

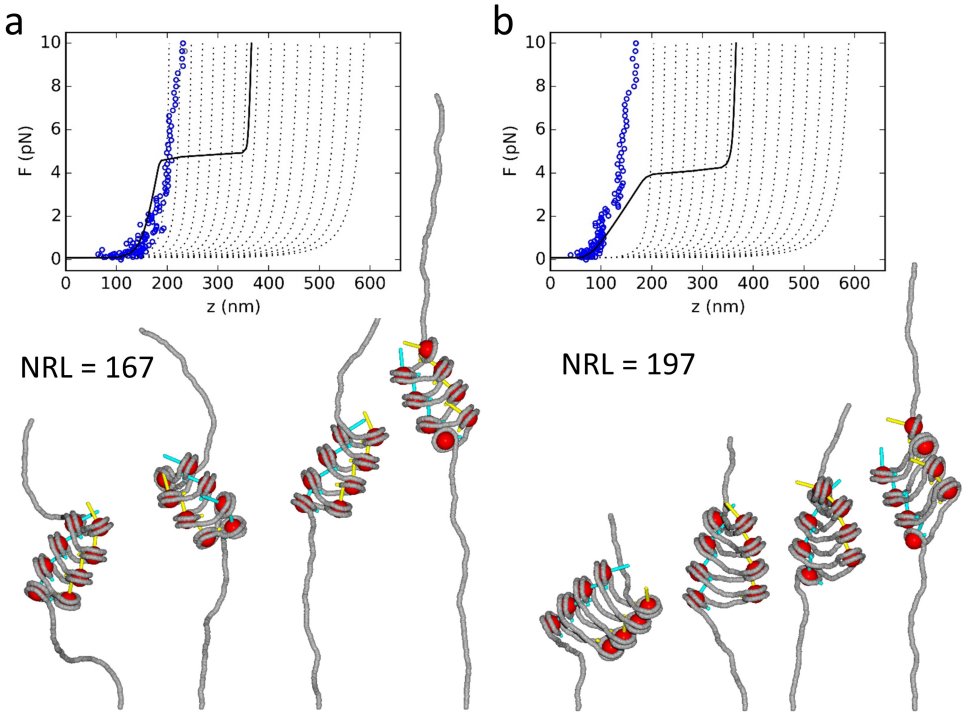



**Figure 4.4**
**MC simulations of the stretching of 2-start fibers show a high stiffness.** a) A simulated force-extension curve of a 2-start 167 NRL fiber containing 8 nucleosomes. b) The force-extension curve of a 2-start 197 NRL fiber. The snapshots below the curves show the structures of the fiber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

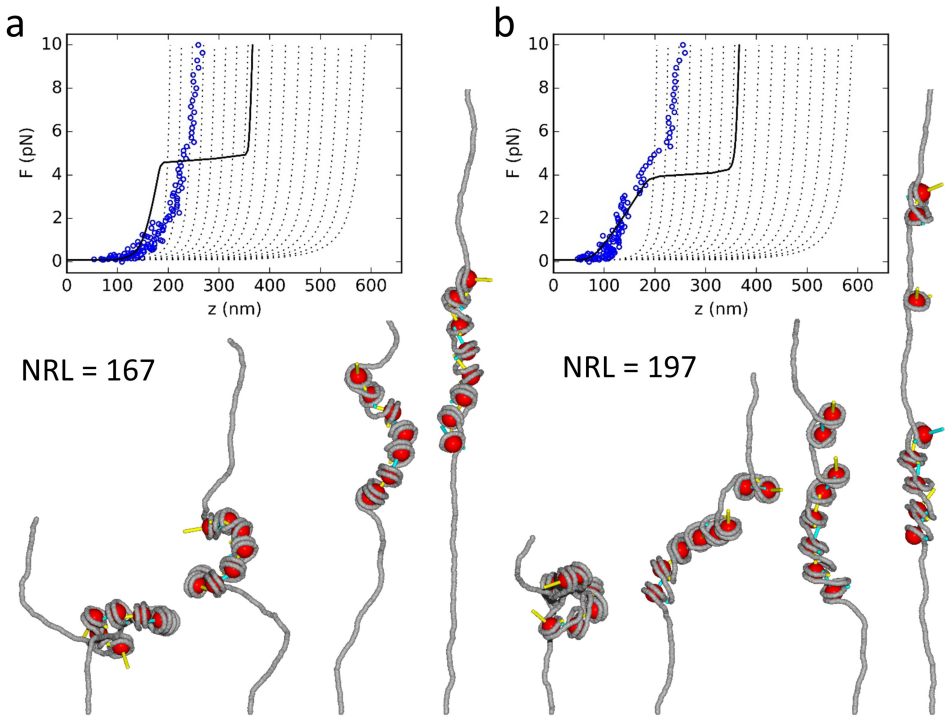



**Figure 4.5**
**MC simulations of the stretching of 1-start fibers feature a lower stiffness.** a) A simulated force-extension curve of the 1-start 167 NRL fiber containing 8 nucleosomes. b) The force-extension curve of a 1-start 197 NRL fiber. The snapshots below the curves show the structures of the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

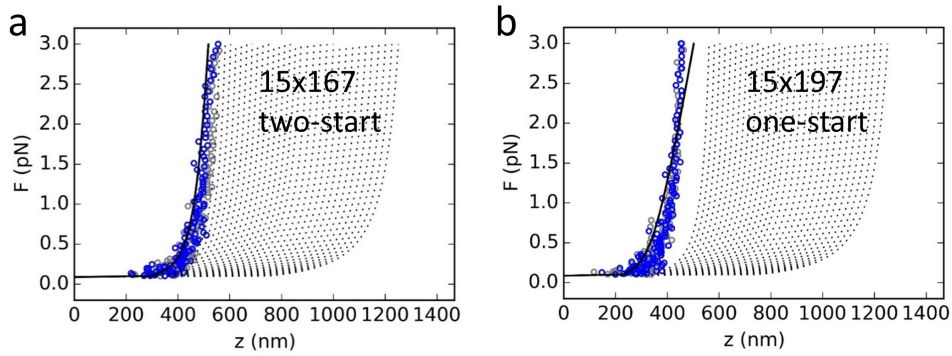



**Figure 4.6**
**MC simulations of 4000 base pairs fibers containing 15 nucleosomes reproduce experimental data up to the unstacking transition.** a) A 167 NRL fiber in a 2-start conformation. b) A 197 NRL fiber in a 1-start conformation. Blue dots represent pulling curves, grey dots the subsequent release 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`.

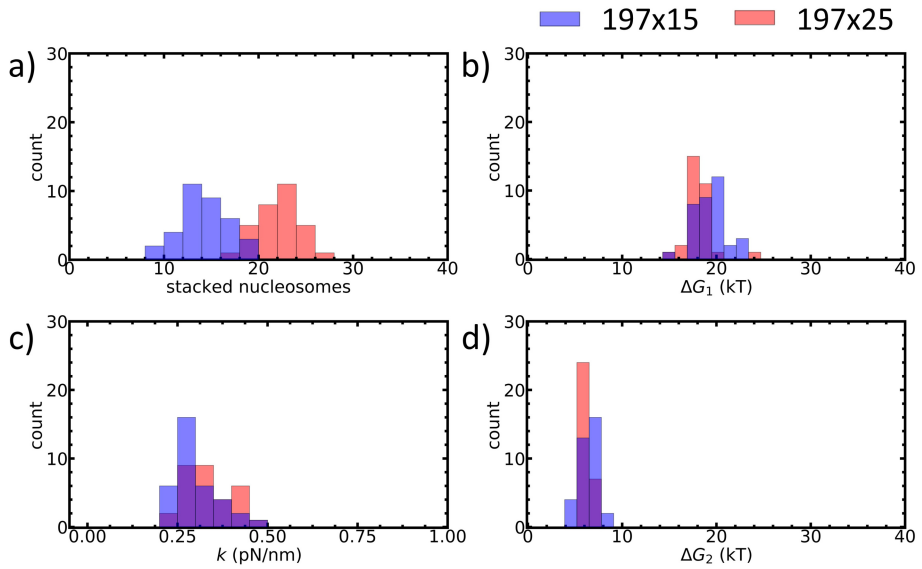



**Figure 4.S1**
**Experimental force-extension curves can be captured in 3 parameters that are independent of the number of nucleosomes.** Force-extension curves of more than 30 different fibers having either 15 (blue) or 25 (red) 197 NRL repeats of the 601 nucleosome positioning sequence were fitted 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.

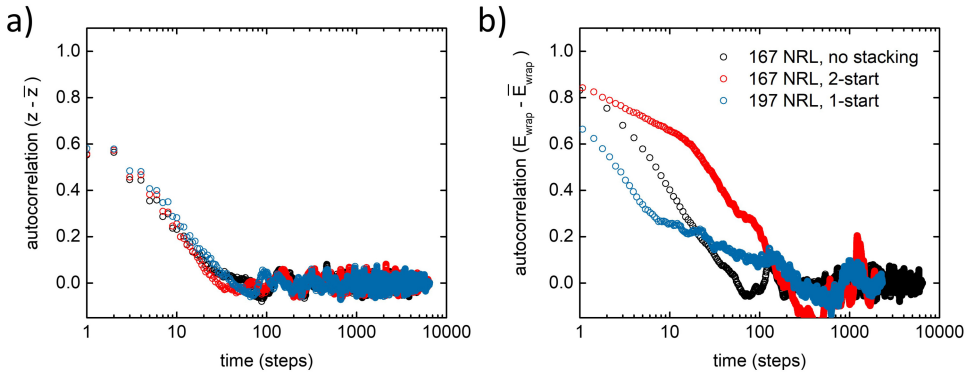



**Figure 4.S2**

**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.

# Chapter 5

# A Critical Role for Linker DNA in Higher-Order Folding of Chromatin Fibers

Nucleosome-nucleosome interactions drive folding of nucleosomal arrays into dense chromatin fibers, though resolving their structure and dynamics remains challenging. A better physical account of the folding of chromatin fibers is necessary to understand the role of chromatin in regulating DNA transactions. Here, we studied the unfolding pathway of regular chromatin fibers as a function of single base pair increments in linker length, using both rigid base pair Monte Carlo simulations and single-molecule force spectroscopy. Both computational and experimental results reveal a periodic variation of the folding energies due to the limited flexibility of the linker DNA. We show that twist is more restrictive for nucleosome stacking than bend, and find the most stable stacking interactions for linker lengths of multiples of 10 base pairs. We analyzed nucleosomes stacking in both 1- and 2-start topologies and show that stacking preferences are determined by the length of the linker DNA. Impaired stacking was observed in fibers with 21-28 base pairs of linker DNA. In fibers with 45-55 base pairs of linker DNA, DNA flexibility is sufficient to accommodate stacked nucleosomes in folded fibers. Moreover, we present evidence that the sequence of the linker DNA also modulates nucleosome stacking and that the effect deletion of the H4 tail depends on the linker length. Importantly, these results imply that nucleosome positioning *in vivo* not only affects the phasing of nucleosomes relative to DNA but also directs the higher-order structure of chromatin.

## Significance statement

The role of chromatin organization in transcription regulation is widely appreciated, but a mechanistic insight is largely lacking. Here, we systematically varied the nucleosome spacing in chromatin fibers and measured their mechanical stability. Fiber properties strongly depended on single base pair changes in linker length as well as the sequence of the linker DNA. Therefore, chromatin structure should be addressed locally, by defining nucleosome-nucleosome interactions as a function of individual linker lengths and sequence, rather than assuming global homogeneity in nucleosome positions. These findings may have large consequences as the accessibility of our genome may be tuned, for example by chromatin remodeling complexes, by small changes of nucleosome positions that modulate the higher order structure of chromatin.

This chapter is based on: Brouwer T. B., Pham C. Kaczmarczyk A., de Voogd W. J., Botto M., Vizjak P., Müller-Planitz F., and van Noort S. J. T.: A Critical Role for Linker DNA in Higher-Order Folding of Chromatin fibers, *submitted*.

## 5.1 Introduction

In eukaryotic cells, DNA is compacted into arrays of nucleosomes, forming chromatin fibers. The structure of these fibers affects the accessibility of DNA [1] and consequently plays an important role in regulating DNA transactions such as transcription, replication, and repair [2–4]. Nucleosome stacking interactions drive chromatin fibers into compact higher-order structures, generally referred to as 30-nm fibers. The structure of these fibers has proven difficult to resolve and have been described in terms of disordered fibers, without regular features, and 1-start and 2-start models, in which neighboring or next-neighboring nucleosomes form one or two stacks of nucleosomes that fold in a super helix. The condensed structure of the left-handed 2-start fiber, composed of tetra-nucleosomal units, was resolved with cryo-EM [5], and was highly similar to the crystal structure of a tetra-nucleosome [6]. Analogous 2-start structures have been reported based by RICC-sequencing [7] and Micro-C [8] approaches. However, both the generality of such a 2-start topology for 30-nm fibers [9–11] and their occurrence and function *in vivo* have been heavily debated [12]. Whereas highly condensed, regular fibers have been reconstituted, folded and characterized *in vitro*, high-resolution EM imaging [13, 14] and super-resolution optical microscopy [15] of eukaryotic nuclei rather suggest a disordered packing of nucleosomes in native chromatin fibers. Pending further high-resolution structural data, the discrepancy between these reports impedes a mechanistic understanding of chromatin folding and its role in the eukaryotic genome metabolism.

The primary structure of chromatin, *i.e.* the location of nucleosomes along the DNA, has been mapped with exquisite detail using nucleosome-sequencing approaches [16, 17]. Alignment of transcription start sites (TSS), yields decidedly regular, equally spaced nucleosome distributions near the TSS, and other barriers in the genome [18]. Recently, Baldi and coworkers reported that such regular spacing of nucleosomes is more universal and is also prominent in silent promotors [19]. This implies variations in phasing of otherwise regular arrays of nucleosomes. Importantly, nucleosome spacing, quantified by the average nucleosome repeat length (NRL), varies with the functional conformation of chromatin [20]. This implies that subtle, local regulation of the NRL may play an important role in genome function, possibly by modulation of the higher-order structure of chromatin.

When reporting quantized linker lengths of $10n + 5$ base pairs in Saccharomyces cerevisiae, Widom speculated that this might be the result of constraints imposed by the higher-order structure of chromatin [21]. However, the overall high transcription activity in yeast suggests an absence of condensed chromatin. Higher organisms and mature cells tend to have longer linker lengths and favor $10n$ lengths, as summarized by Perišić *et al.* [22]. Though modern high-resolution genome-wide methods allow for mapping of nucleosome positions with single base pair accuracy [16], analysis of nucleosome-nucleosome

interactions not only requires positional mapping of multiple nucleosomes on the same DNA substrate, it also necessitates resolving nucleosome-nucleosome interactions simultaneously. This has not been possible and structural studies on native chromatin typically average over large populations of chromatin fragments [12–15], obscuring the details of higher-order folding. The structural consequences of variations of the NRL have to date only been evaluated coarsely.

*In silico* studies of chromatin fibers, which allow a detailed energetic evaluation of single chromatin fiber structures, highlight the role of linker DNA – the DNA between adjacent nucleosomes – in the geometry and elasticity in compacting chromatin. A multitude of condensed fiber topologies has been computed for different linker lengths when using straight linker DNA and fixed nucleosome angles [23]. Mesoscale modeling yielded zigzag structures for short NRL chromatin fibers and solenoid-like features for longer NRLs [22]. The implications of small changes in NRL, however, seem to be best modeled by more refined rigid base pair models. By fixing nucleosome orientations in a 2-start topology, Norouzi and Zhurkin reported a 10 base pair periodicity in fiber compaction [24–26] and quantified the energetic penalty for non-optimal NRLs.

To experimentally resolve the role of NRL in chromatin higher-order structure, tandem arrays of the Widom 601 nucleosome positioning sequence have been instrumental [27]. Though any NRL can be cloned into a regular array [28] and reconstituted with nucleosomes, experimental work has almost exclusively focused on $10n$ NRLs, yielding high-resolution X-ray structures of 157 [29] and 167 [6] NRL fibers, single-particle cryo-EM reconstructed structures of 177 and 187 NRL fibers [5], as well as quantification of the fiber diameter of 177-237 NRL fibers in 10 base pair steps by EM [9]. It is likely that this focus on $10n$ NRLs not only reflects earlier reports on NRLs found *in vivo*. It may also be that the better folding properties of these fibers make it easier to handle them experimentally, but little experimental data are available for intermediate NRLs.

Our previous force spectroscopy studies, which uniquely probe single chromatin fibers without staining, fixation or surface deposition, showed that unfolding of 167 NRL fibers under physiological conditions could best be interpreted by 2-start structures, whereas unfolding of 197 NRL fibers could better be described with a 1-start unfolding model [10, 11, 30]. Many of these $10n$ NRLs have also been scrutinized using other biochemical and biophysical techniques, such as sedimentation analysis, Atomic Force Microscopy (AFM) and single-molecule force spectroscopy. Most notably, Grigoryev *et al.* used sedimentation analysis, AFM and EM to quantify chromatin compaction as a function of small NRL increments [31]. A 10 base pair periodic dependence of chromatin folding for NRLs in the range of 165–177 base pairs was found and the highest level of condensation was found for 20 and 30 base pairs linker DNA. However, a systematic experimental analysis of chromatin fiber structure

with single base pair increments of the NRL is lacking.

Here, we combine rigid base pair Monte Carlo (rbMC) simulations with single-molecule force spectroscopy to probe a series of chromatin fibers and uncover the detailed implications of NRL changes on fiber folding with single base pair increments. The rbMC simulations allow detailed evaluation of chromatin structures with different topologies and quantitative comparison of the energetic cost of linker DNA distortions that nucleosome stacking imposes. Previously, we have shown that rbMC simulations, in which histone-DNA and histone-histone interactions are implemented by simple harmonic potentials, can reproduce experimental force-extension curves of folded fibers fairly well [32]. In this study, we used magnetic tweezers to measure the mechanical properties of a large set of reconstituted fibers and combined the results with the insight obtained from simulations. Furthermore, we quantified nucleosome-nucleosome stacking energies and investigated how much the H4 tail contributes to nucleosome stacking in short and long NRL fibers. Finally, we tested whether the sequence of the linker DNA could affect the mechanical properties of chromatin. Overall, we find that the mechanical constraints imposed by the linker DNA define if and with which other nucleosomes the nucleosome stacks, and thereby direct nucleosomal arrays into specific higher-order structures.

## 5.2 Results

### 5.2.1 rbMC simulations yield the detailed conformation of linker DNA

To quantify the geometrical constraints on nucleosome stacking in folded chromatin fibers, we performed a set of rbMC simulations in which we evaluated three structures: a fiber devoid of nucleosome-nucleosome interactions (0-start), a solenoid fiber (1-start), and a zig-zag fiber (2-start). Nucleosome stacking was implemented by nucleosome step parameters in 6 degrees-of-freedom (DoF), analogous to base pair step parameters [32]. Figure 5.1 shows snapshots of states of NRL 167 and NRL 197 fibers. The fibers were gently stretched by 3 pN force to provide an unobstructed view on fiber folding. As expected, and particularly visible in Figure 5.1b, the 0-start structures were most elongated and featured some unwrapping of the nucleosomal DNA. Nucleosome stacking largely inhibited such unwrapping in 1-start and 2-start fibers. As reported before [32], the linker DNA was more bent in 1-start than in 2-start structures, but nucleosome stacking was better accommodated in 1-start fibers. The harmonic stacking and wrapping potentials may require further refinement to properly discriminate between the different topologies. Nevertheless, the rbMC simulations allow for a detailed, qualitative comparison of the linker DNA in different structures and are well suited to evaluate changes upon increasing the linker length.

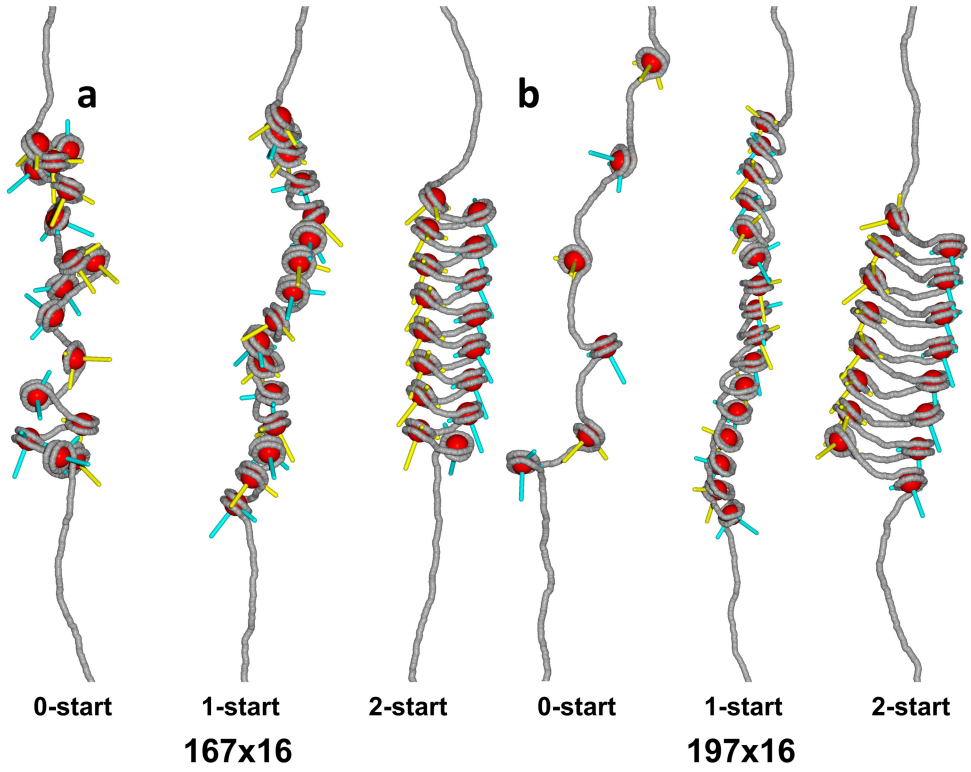



**Figure 5.1**
**Nucleosome stacking defines the higher-order structure of chromatin fibers.** The spatial orientation of linker DNA (grey) and nucleosomes (red) is depicted in three chromatin fiber structures: 0-start, 1-start, and 2-start. In this simulation, 3 pN of force was applied to stretch the fiber structures [32]. Nucleosome orientation was indicated by the blue and yellow rods, pointing to the dyad of the nucleosome and the direction along the superhelical axis. a) The relatively short linker DNA in NRL 167 fibers constrains nucleosome stacking in chromatin fibers. The 1-start fiber required strong linker DNA bending to accommodate nucleosome stacking. The linker DNA in the 2-start fiber was less bent. b) NRL 197 fibers can more easily facilitate the linker bending imposed by 1-start fibers. The linker DNA of the 2-start fiber featured only moderate bends.

In Figures 5.2 and 5.3 we show the changes in the conformation of the linker DNA for the three topologies as the linker length is increased in small steps. The structures were reconstructed from the average step parameters of the DNA of the first of three nucleosomes after 10,000 Monte Carlo iterations at zero force. To pinpoint large deformations of the linker DNA, the deviations in the step parameters from unperturbed DNA were converted to energies per base pair and depicted in red (*see* Materials and Methods). In the absence of stacking, the linker DNA was approximately straight. Excluded volume effects, incorporated in our simulations, did not impose enough constrains to

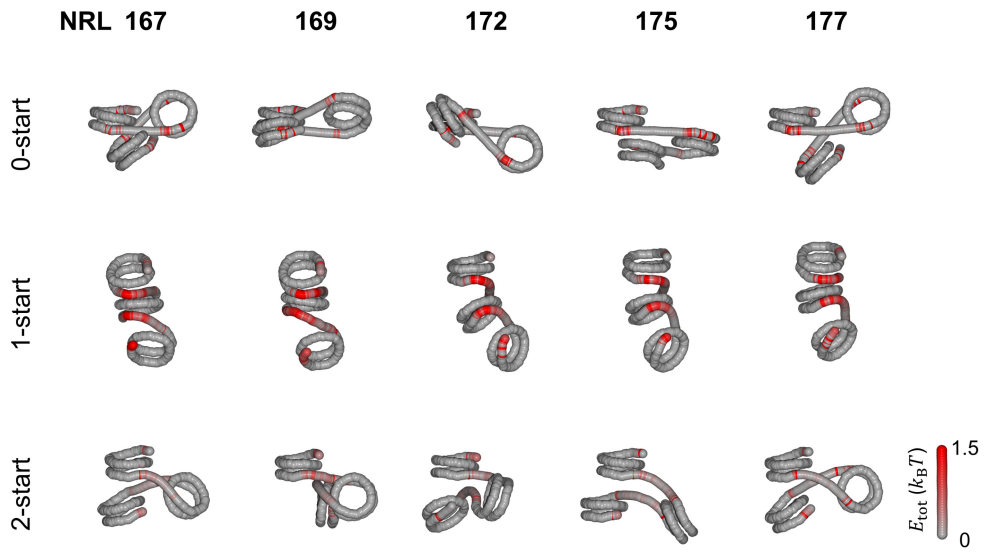



**Figure 5.2**
**Bending of short (20-30 base pairs) linker DNA imposes large energy penalties for 1-start fibers and compromises stacking in 2-start fibers.** The average conformation of the DNA of three adjacent nucleosomes, including their linker DNA, was calculated for the 0-start, 1-start, and 2-start structures. The tension in the individual base pairs in the linker DNA is represented by the red color, which shows the bending energy compared to the starting, minimal energy structure. When no stacking interactions were simulated, the linker DNA was on average straight. In this conformation, nucleosome breathing was observed at the outer contact points of each nucleosome, resulting in red base pairs that reflect an increased standard deviation of their coordinates as compared to the start conformation. The bending of the linker DNA in a 1-start structure required strong deformation of the linker DNA, indicated by large energy penalties. NRL 167 fibers showed approximately straight linker DNA in a 2-start structure. Increasing the linker length imposed deviations from optimal stacking, which was maximal for NRL 172. Increasing the linker length to NRL 177 recovered proper stacking.

impose large deformations of the linker DNA in any of the linker lengths of the 0-start fiber. The base pairs depicted in red near the end of the nucleosomes reflect unwrapping of the nucleosomal DNA. As the linker length increases, the orientation of adjacent nucleosomes rotates, as every additional base pair twists the adjacent nucleosome by approximately 36 degrees.

When neighboring nucleosomes were forced to stack into a 1-start fiber, bending of the linker DNA imposed large energy penalties, depicted in Figure 5.2 by the bright red color of the linker DNA. Nucleosome pairs were positioned in very similar orientations, showing that stacking in the 1-start structure could be achieved for all linker lengths. Little difference in terms of structure and

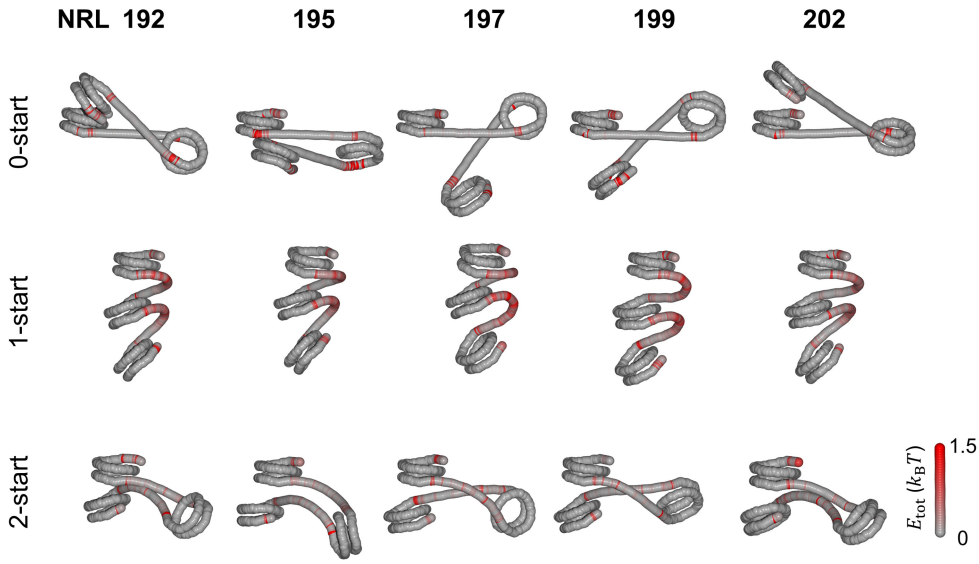



**Figure 5.3**
**Longer (45-55 base pairs) linker lengths can better accommodate linker DNA stress in 1-start fibers.** 0-start structures with straight DNA featured nucleosome breathing. The increased linker DNA significantly reduced the stress in linker DNA imposed by the 1-start stacking as compared to 167-177 NRL fibers. A large variety of nucleosome orientations was found in 2-start structures in which linker DNA was slightly deformed.

in terms of energy penalties was observed as the linker length increased from 20 to 30 base pairs. Perhaps surprisingly, there was little unwrapping of the nucleosomal DNA in the 1-start fibers, except for the nucleosomal DNA in the first half of the bottom nucleosome, which was not constrained by stacking.

When nucleosomes were forced into a 2-start fiber, the linker DNA was straighter, yielding lower energy penalties. Like in the 1-start state, nucleosome unwrapping was repressed in the 2-start fiber. The relative orientation of the first and the third nucleosomes, however, varied strongly as the linker DNA length increased from 20 to 30 base pairs. Especially the NRL 169, NRL 172 and NRL 175 structures deviated considerably from the optimal stacking geometry. Apparently, nucleosome stacking was more restricted by the linker DNA in 2-start fibers than in 1-start fibers, resulting in a large variety of higher-order structures. Chromatin fibers with NRL 167 and NRL 177 had the straightest linker DNA and appeared better stacked. In summary, the rbMC simulations of fibers with 20-30 base pairs of linker DNA shown in Figure 5.2 featured different geometries in which the 1-start fibers featured most stress in the linker DNA, while 2-start fibers compromised stacking to reduce linker DNA stress.

For linker lengths between 45 and 55 base pairs, we observed similar trends,

as shown in Figure 5.3. The energy penalties for chromatin with longer linker DNA were, however, considerably lower since stress was distributed over more base pairs. In the absence of nucleosome stacking, chromatin fibers with longer linkers sustained straight DNA and nucleosome breathing. In 1-start fibers, the stress appeared to accumulate at distinct locations that were spaced by 10 base pairs, but nucleosomes positions and orientations were quite similar in the series of simulations. The linker DNA of 2-start fibers was only gradually bent and the stress appeared to be well distributed along the linker DNA. The nucleosomes appeared similarly stacked, independent of the linker length although the relative orientation of the two staked nucleosomes varied a lot, which may have a large impact on the resulting higher-order structures.

### 5.2.2 Decomposition of energy distributions as a function of NRL highlights twist as a decisive factor in fiber folding

To quantify the effect of linker length in more detail, we calculated the average folding energy per nucleosome. Each energy contribution, relative to the 0-start fiber structure, was calculated by summing the energy terms over all base pairs between two dyads. In Figure 5.4, the contributions of each degree of freedom are plotted as a function of NRL. Overall, the energy contributions of the linker DNA decreased with linker length, as stacking could be better accommodated when the resulting stress is distributed over longer linker DNA. The translational step parameters (shift, slide, rise) showed little difference between the 1-start and 2-start states. The deformation of the linker DNA was more prominent in the rotational step parameters (tilt, roll, twist): folding fibers into a 1-start conformation costs up to 8 $k_{\mathrm{B}}T$ for tilt and 10 $k_{\mathrm{B}}T$ for roll. The stress in the linker DNA in 2-start fibers was significantly less, starting at a few $k_{\mathrm{B}}T$ for tilt and roll in NRL 167 and approaching zero for longer NRLs. The two states differ most in terms of twist: whereas 2-start fibers featured a small but gradual decay in twist energy as the linker length increases, 1-start fibers displayed a 10 base pair periodic modulation of the twist energy. The most favorable linker lengths at $10n$ base pairs were up to 10 $k_{\mathrm{B}}T$ lower in energy than the most unfavorable ones.

Next to the energy contributions due to changes in the base pair steps, we also plotted the changes in wrapping and stacking energy in Figure 5.4. Unwrapping of nucleosomal DNA contributed only marginally to energy differences between the three structures. Nucleosome stacking lowered the energy per nucleosome by approximately 23 $k_{\mathrm{B}}T$, which is about 2 $k_{\mathrm{B}}T$ less than the maximum stacking energy, showing that the simulations resulted in a compromise between optimal stacking and minimization of the stress in the linker DNA. Interestingly, closer inspection of the stacking energy revealed a 10 base pair periodicity for 2-start fibers (inset Figure 5.4h). It therefore seems that 2-start fibers prefer to compromise stacking when the linker length is not optimal, whereas 1-start

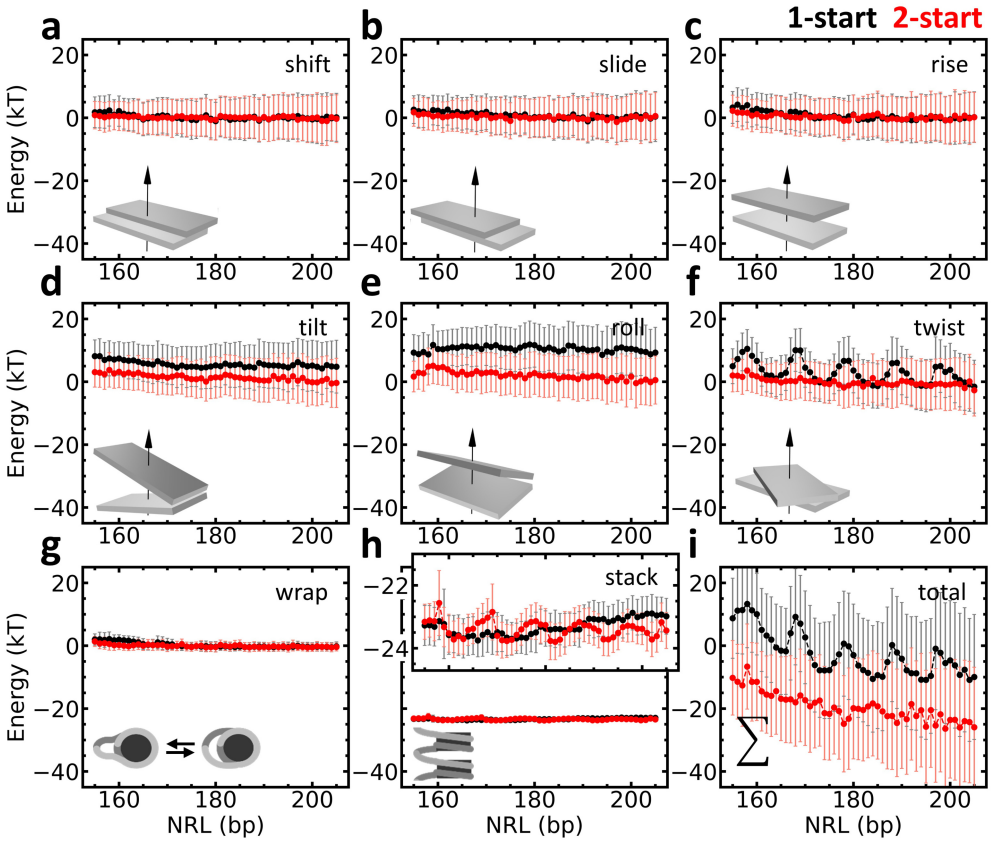



**Figure 5.4**

**Rigid base pair Monte Carlo simulations decompose the energy contributions of fiber folding into linker DNA deformation, DNA unwrapping, and nucleosome stacking.** rbMC simulations were used to quantify the total energy penalty for each base pair step associated with nucleosome stacking. The energy contributions relative to the 0-start fiber were split in shift (a), slide (b), rise (c), tilt (d), roll (e), twist (f) and nucleosome wrapping (g) and stacking (h) parameters. Increasing NRLs resulted in decreased the energies of most DoF. The exception was the twist energy (f) of 1-start fibers, which was strongly modulated by NRL and oscillated with a period of 10 base pairs and an amplitude of approximately 10 $k_BT$. The 2-start fiber was more affected in the stacking energy than in twist energy, and also oscillated with 10 base pairs periodicity, but with a significantly smaller amplitude. In general, energies of the 1-start fiber were higher than the 2-start fiber, especially in the rotational DoF. However, since standard deviations were large, the 1-start fiber could occasionally be favored over the 2-start fiber for high NRL. i) The sum of all contributions also exhibits a periodic average modulation of the energy with NRL, but differences between 1- and 2-start folding are comparable to the standard deviation of the simulated energies.

fibers rather accommodate stacking by twisting the linker DNA.

In summary, the total energy per nucleosome pair was dominated by contributions of the rotational step parameters. The 2-start state appeared to have the lowest energy for all NRLs. Note, however, that the difference between the two conformations reduced with linker length and, perhaps more importantly, that the standard deviations that were computed exceeded the difference in energy between the two topologies. Though the trends that we observed are clear and highlight torsional stress rather than bending of linker DNA as the defining aspect of fiber folding, we want to emphasize that the harmonic approximation of the stacking energy may not be sufficient for absolute quantitative comparison (*see* Discussion). Nevertheless, based on these simulations we expect a strong dependence of fiber folding on the linker length, featuring a 10 base pair periodicity of its stability.

### 5.2.3 Single-molecule force spectroscopy reveals a 10 base pair periodic dependence of fiber folding parameters on NRL

To experimentally examine the role of linker DNA on chromatin fiber structure in a systematic manner, we performed a series of force spectroscopy experiments on chromatin fibers with increasing NRLs. A series of 22 tandem arrays of 16 601 positioning sequences were cloned using the method of Wu *et al.* [28] (*see* Materials and Methods). The linker length was varied in single base pair increments in range NRL 167-177 (*short* linker DNA) and NRL 192-202 (*long* linker DNA). Both series cover a one helical period of the DNA, which should result in a full turn of each nucleosome with respect to its neighbor in 0-start fibers.

Figure 5.5a schematically shows the unfolding of a nucleosome that is embedded in a folded fiber, in which several characteristic force-induced intermediate conformations occur [30]. A typical experimental force-extension curve of a chromatin fiber is shown in Figure 5.5b. At forces below 3-5 pN, the nucleosomes are predominantly stacked in a folded fiber (conformation I in Figure 5.5a shows a 1-start fiber, but the same model, with adjusted parameters can account for 2-start fibers). When the force increases, the nucleosomes unstack into a partially unwrapped conformation (conformation II). Further unwrapping yields a singly wrapped conformation in which 80 base pairs are wrapped around the histone core (conformation III). At forces above 10 pN, the singly wrapped nucleosomes unfold into a fully unwrapped conformation that has an extension that equals that of bare DNA (conformation IV). This last transition is not in equilibrium, as opposed to the first two transitions, and results in distinct steps of approximately 25 nm for each nucleosome [30].

Three exemplary force-extension curves of chromatin fibers with NRL 167, 168, and 169 are plotted in Figure 5.5c. Compared to the NRL 167 fiber, the NRL 168 and NRL 169 fibers were easier to stretch and ruptured at smaller

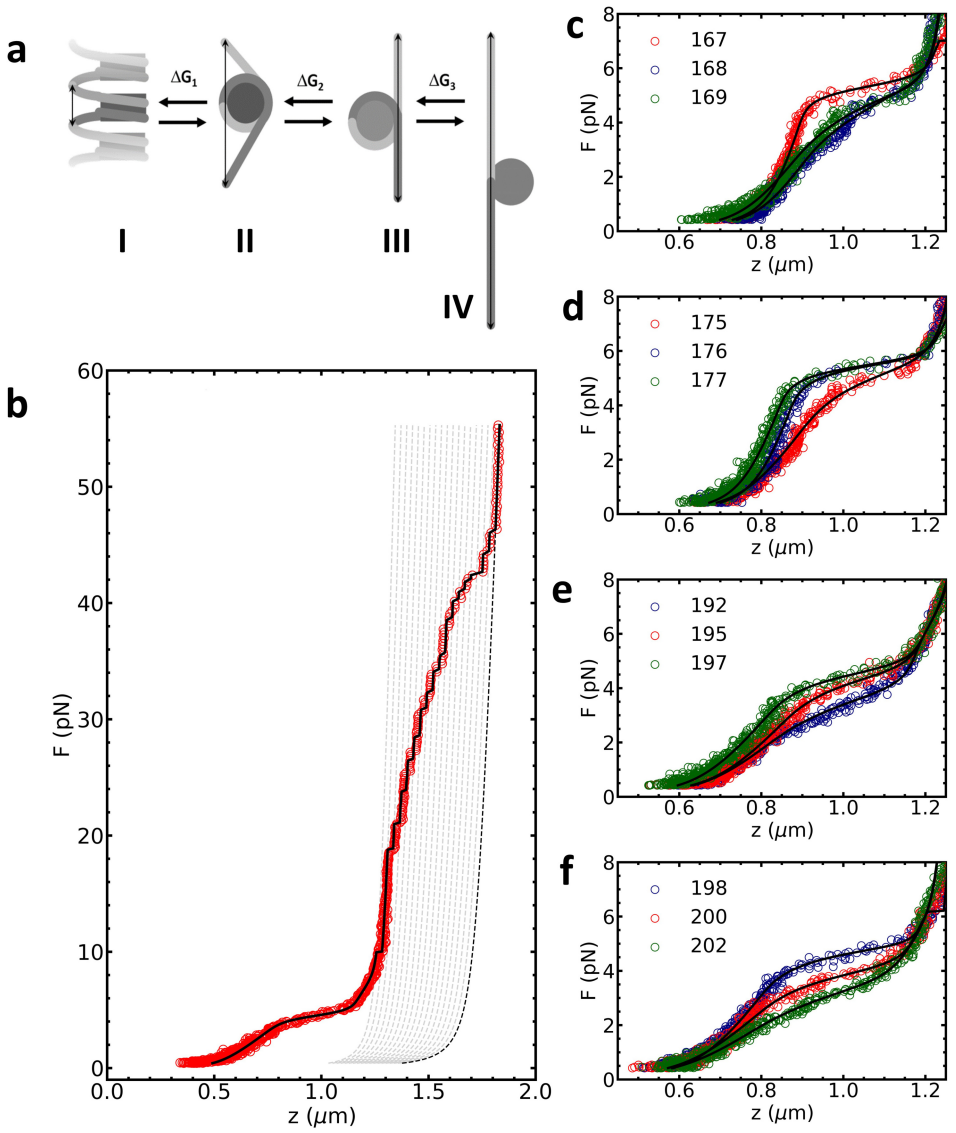

a
ΔG1
ΔG2
ΔG3
I
II
III
IV
b
F (pN)
z (μm)
c
167
168
169
d
175
176
177
e
192
195
197
f
198
200
202

**Figure 5.5**

**Force spectroscopy experiments show that single base pair changes in linker length have large effects on fiber folding.** a) chromatin fibers unfold in three steps and switches between four conformations: I) Folded fibers consist of stacked nucleosomes. II) Unstacking yields a partially unwrapped nucleosome and costs rupture energy $\Delta G_1$. III) Subsequent unfolding yields singly wrapped nucleosomes and costs interaction energy $\Delta G_2$. IV) At high force, the DNA is completely unwrapped from the histone core at the cost of wrapping energy $\Delta G_3$. b) Fitting a typical force-extension curve of a chromatin fiber (NRL 197) with the statistical mechanics model yielded 18 stacked nucleosomes, fiber stiffness $k = 0.3 \pm 0.1$ pN/nm, rupture energy $\Delta G_1 = 23 \pm 1\ k_\mathrm{B}T$, and partial interaction energy $\Delta G_2 = 8 \pm 1\ k_\mathrm{B}T$. c) A qualitative comparison of chromatin fibers with NRL 167, NRL 168, and NRL 169 revealed that fibers with NLR 167 unfolded cooperatively and had the highest $k$ and $\Delta G_1$. Increasing the linker length by a single base pair to NRL 168 distorted fiber folding, reflected by the loss of cooperativity and the sharp drop in $k$ and $\Delta G_1$. Increasing the linker length to NRL 169 did not have a subsequent effect. $\Delta G_2$ was the same for all NRL. d) For approximately 10 base pairs longer linker lengths the reverse trend was observed. NRL 175 depicted gradual unfolding, low stiffness and reduced rupture force. Increasing the linker length to NRL 176 restored the shape of the curve to the one observed for NRL 167. Further increase of the linker length to NRL 177 had no effect. Again, $\Delta G_2$ remained unaffected by NRL. e) A different trend was observed for longer linker DNA. Chromatin fibers with NRL 192 had a relatively low $k$ and $\Delta G_1$, which steadily increased for NRL 195 and NRL 197. f) NRL 198, NRL 200 and NRL 200 showed a decline in $k$ and $\Delta G_1$. Fitted parameters of all curves in c-f) can be found in Table 5.S1 of the supplementary material.

forces. The larger stiffness and rupture force of the NRL 167 fiber were recovered when the NRL approached 177 base pairs (Figure 5.5d). For the longer NRLs, we obtained a similar, but more gradual trend (Figures 5.5e and f). For both short and long linkers, linker lengths near $10n$ base pairs were stiffer and ruptured at larger forces than intermediate NRLs. Thus, single base pair increments of the NRL had a large impact on the fiber's mechanical response to force.

The mechanical properties of each of the 4 unfolding conformations, as well as the equilibrium transitions between them, were fitted to a statistical mechanics model that is based on the mechanical properties of DNA, the strength of nucleosome-nucleosome interactions, the histone-DNA interactions, and the composition of the chromatin fiber [30]. Details are described in Materials and and Methods. Here, we focus on the parameters that describe the rupture of the folded chromatin fiber. For all measurements, we found that the experimentally obtained force-extension curves in the force regime below 3 pN could be described by the combination of a worm-like chain (WLC) for the free DNA handles, and a Hookean spring with stiffness $k$ for the chromatinized part of the tether. The limited force range and the relatively small extension over which the fiber is stable may explain the success of this linear approximation.

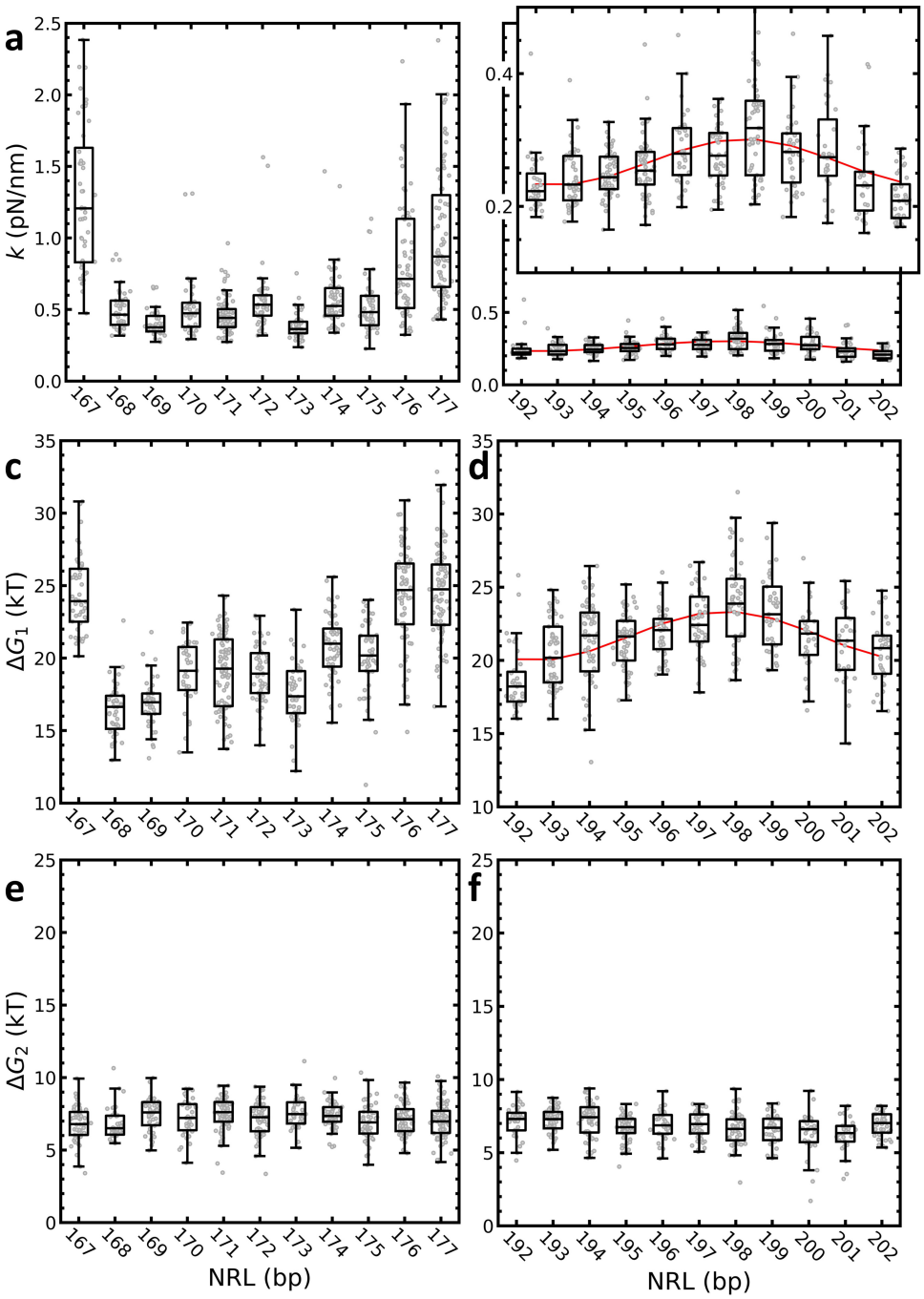
Because we found that fibers reconstituted with tandem arrays of 601 sequences still featured heterogeneity in terms of the number of reconstituted nucleosomes, we manually counted the number of 25 nm steps at $f > 10$ pN and attributed these to the number of nucleosomes. Between fibers, this

**Figure 5.6**
**Base pair step parameters feature a 10 base pair NRL periodicity.** a) For short NRLs, $k$ peaked at NRL 167, NRL 176, and NRL 177. Strikingly, the stiffness dropped sharply between NRL 168 and NRL 175, illustrating the sensitivity of $k$ for linker DNA length. b) For long NRLs, $k$ described a gradual modulation which could be captured by a sinusoid, peaking at NRL 198 ($k = 0.3 \pm 0.1$ pN/nm). c) $\Delta G_1$ showed a similar trend as $k$; for short linker lengths peak values were found for NRL 167, NRL 176, and NRL 177 which dropped in between. d) Also for long linker lengths $\Delta G_1$ oscillated with a period of 10.4 base pairs and peaked at NRL 198, similar to $k$. e-f) As expected for non-interacting nucleosomes, $\Delta G_2$ was independent of NRL. Parameters were based on $48 \pm 14$ (mean $\pm$ std) fibers per NRL. Fitted Parameters of all populations can be found in Table 5.S2 of the supplementary material.

number varied between 12 and 20, but was typically centered around 16 for 16 repeats of the 601 positioning sequence. The highly reproducible value of the interaction energy $\Delta G_2$ for all NRLs implies that this correction worked well. When comparing fibers with identical linker sequence but a different number of repeats (both NRL 167 and NRL 197), the distribution of fitting parameters overlapped (*see* for example Figure 5.8), which implies that the variation in the composition of individual chromatin fibers did not affect the fit results.

For different NRL fibers, the obtained fitting parameters, *i.e.* stiffness $k$, rupture energy $\Delta G_1$, and interaction energy $\Delta G_2$, showed a large variation that should be attributed to differences in fiber folding. Fitting the curves in Figures 5.5c-f yielded the highest stiffness $k = 1.1 \pm 0.1$ pN/nm and rupture energy $\Delta G_1 = 23 \pm 1\ k_{\mathrm{B}}T$ for NRL 167, in agreement with our qualitative assessment. For NRL 168 and NRL 169, both the stiffness and the rupture energy dropped, which was ultimately recovered for NRL 176 base pairs. The chromatin fibers with longer linker lengths had a 2-3 times smaller stiffness than that of the short NRLs. The rupture energy $\Delta G_1$ varied between 15 and 25 $k_{\mathrm{B}}T$. $\Delta G_2$ was unaffected by NRL, as would be expected since at this stage the fibers are stretched far enough to prevent interactions between nucleosomes. The difference in contour length between the partially wrapped and the singly wrapped nucleosomes was fixed at 13 base pairs, which would correspond to 2 histone-DNA contacts of each about 3.5 $k_{\mathrm{B}}T$, in fair agreement with previous estimations [10, 33]. For the intermediate NRLs we obtained good fits and parameters that are consistent with these trends. Thus, all effects of changing the NRL could be captured in two fitting parameters: the fiber stiffness and the nucleosome-nucleosome rupture energy. Representative examples of experimental curves for all NRLs and their fits are plotted in Supplementary Figure 5.S1.

To quantify variations between individual fibers from the same batch, we measured and fitted the force-extension curves of at least 35 different fibers for each of the 22 batches of different NRL that we reconstituted. The resulting fit parameter distributions are plotted as decorated box plots in Figure 5.6. Next


a
k (pN/nm)
2.5
2.0
1.5
1.0
0.5
0.0
0.4
0.2
0.5
0.0
167 168 169 170 171 172 173 174 175 176 177
192 193 194 195 196 197 198 199 200 201 202
c
d
ΔG1 (kT)
35
30
25
20
15
10
e
f
ΔG2 (kT)
25
20
15
10
5
0
NRL (bp)
NRL (bp)

to the trends described above, we observed a qualitative difference between the short and the long NRL series. Whereas the long NRLs show a gradual periodic modulation of the average stiffness and rupture energy $\Delta G_1$, that we fitted with a sinusoidal function with a period of 10.4 base pairs, the short NRL series featured a much more abrupt drop (NRL 167 to NRL 168) and recovery (NRL 175 to NRL 176) of $k$ and $\Delta G_1$. Apparently, an increase of 1 base pair in NRL has a larger effect on short NRLs than on long NRLs. We tentatively attribute the abruptness of the transition and the relatively small values of both stiffness $k$ transition and rupture energies $\Delta G_1$ in between to inhibition of nucleosome stacking. The rupture energy $\Delta G_1$ would, in absence of stacking interactions, correspond to the partial unwrapping of nucleosomal DNA from non-interacting nucleosomes rather than unstacking. Interactions between histone tails and the (linker-) DNA could add to the stability of such non-interacting nucleosomes. Independent of the interpretation of the fit parameters, these single-molecule force spectroscopy experiments and subsequent analysis clearly reveal a 10 base pair periodic modulation of fiber folding with NRL, with a larger impact for short NRLs than for long NRLs.

### 5.2.4 H4 tails are essential for 2-start fiber folding, but not for 1-start fiber folding

To obtain further insight into the balance between nucleosome stacking and linker DNA deformation, we tested the effect of deleting the N-terminal tail of histone H4, resulting in so called globular H4 (gH4). Deletion of this tail would reduce the stacking between nucleosomes in folded fibers. We reconstituted chromatin fibers with recombinant Drosophila melanogaster (*Dm*) histones and compared the fiber unfolding characteristics of fibers reconstituted with both wild-type (WT) and gH4 histone with those reconstituted with human (*Hs*) histones, as used in the previous paragraphs. As shown in Figures 5.7a-b, the force-extension curves for both NRL 167 and NRL 197 fibers featured the characteristic unfolding plateau at 3-5 pN, and all curves fitted well to our unfolding model. Fibers reconstituted with *Dm* WT histones yielded a smaller $k$ and a lower rupture energy $\Delta G_1$ than WT *Hs* histones. The interaction energy, $\Delta G_2$, was only marginally reduced. Thus, chromatin fibers containing WT *Hs* nucleosomes folded slightly better than fibers reconstituted with WT *Dm* histones.

Deletion of the H4 tail impeded folding of NRL 167 fibers, as shown in Figure 5.7a. The fibers unfolded more gradually and at a lower force than with *Dm* WT histones. Fitting the model yielded a lower $k$ and $\Delta G_1$: $k$ dropped from $0.8 \pm 0.4$ to $0.5 \pm 0.2$ pN/nm ($p < 0.001$), and $\Delta G_1$ dropped from $21 \pm 2$ to $17 \pm 2$ $k_\mathrm{B}T$ ($p < 0.001$). Both the stiffness and the rupture energy of the first transition had values that were similar to those of NRL 168 fibers, for which we attributed the reduced values to inhibited nucleosome stacking. In addition,

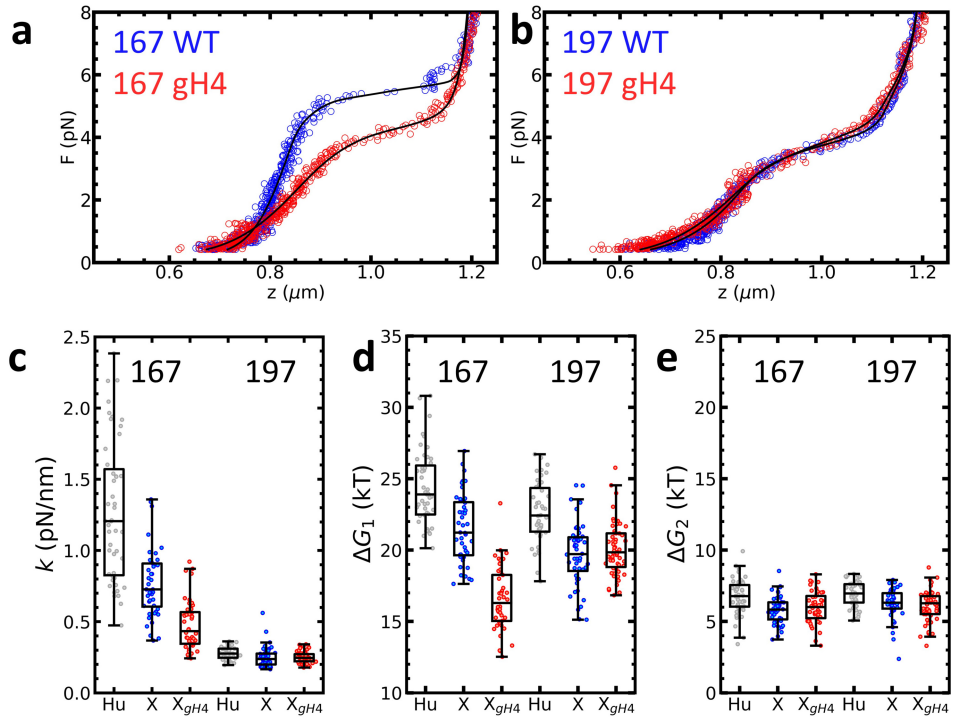


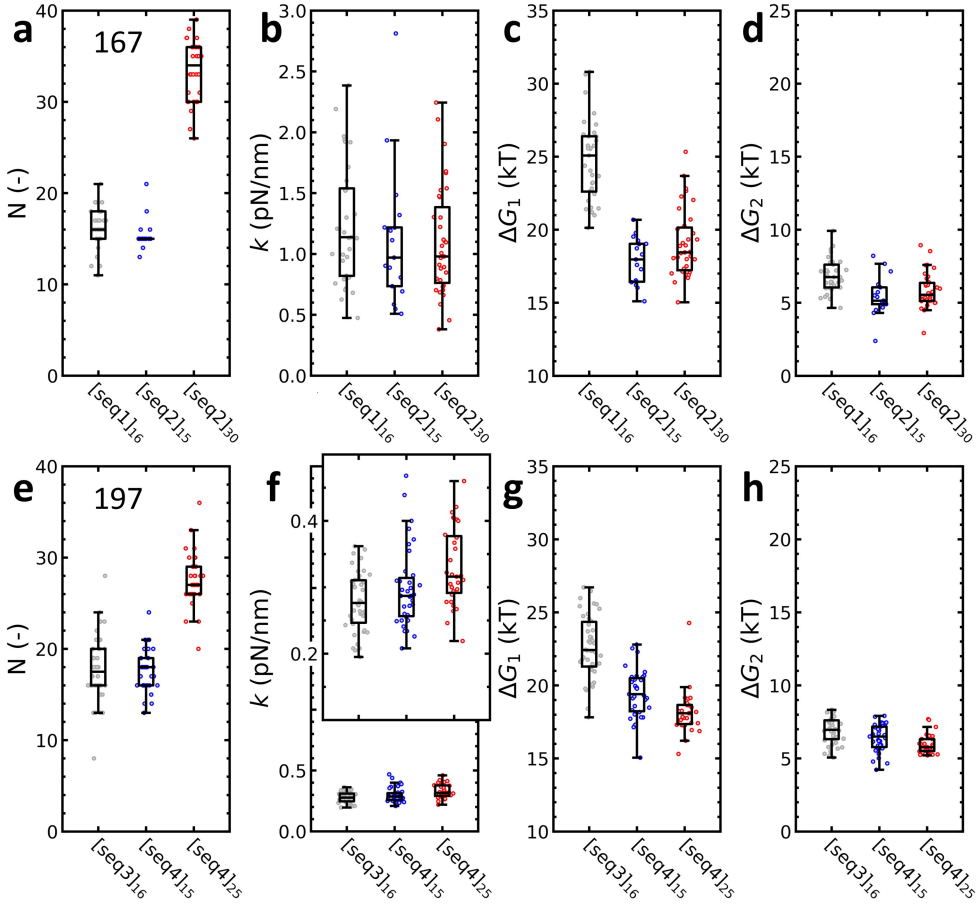

**Figure 5.7**
**Deletion of the H4 tail strongly affected folding in NRL 167, but not in NRL 197 chromatin fibers.** a) Force spectroscopy experiments on chromatin fibers reconstituted with *dm* WT and gH4 mutants. Fibers reconstituted from *dm* WT histones (blue dots) unfolded similarly as those with *h* WT histones. Chromatin fibers reconstituted with *dm* gH4 mutant histones (red dots) unfolded at a lower force. Both stiffness $k$ and $\Delta G_1$ dropped significantly, suggesting an alternative higher-order structure. Surprisingly, deletion of the H4 tail had no observed effect on the mechanical properties of NRL 197 fibers. Fitted parameters (black lines) can be found in Table 5.S3 of the supplementary material. b) In NRL 167 fibers, *dm* gH4 reduced $k$ significantly compared to WT histones ($p < 0.001$), while $k$ of NRL 197 fibers was unaffected ($p = 0.98$). $\Delta G_1$ was similarly modulated: *h* WT histones featured a slightly higher $\Delta G_1$ than *dm* WT histones, both for NRL 167 and NRL 197. Deletion of the H4 tail had a strong effect on $\Delta G_1$ for NRL 167, which dropped significantly ($p < 0.001$). For NRL 197 the unstacking energy was not different ($p < 0.35$). $\Delta G_2$ remained constant irrespective of NRL nor the presence of the H4 tail ($p = 0.43$ for NRL 167, $p = 0.41$ for NRL 197). The much lower $\Delta G_1$ for *dm* gH4 chromatin with NRL 167 in comparison with NRL 197 ($p < 0.001$) implies the absence of higher-order structure. Irrespective of NRL, *h* WT chromatin fibers had a slightly higher $k$, higher $\Delta G_1$, and higher $\Delta G_2$ than *dm* WT fibers. The average and standard deviation are provided in Table 5.S4 of the supplementary material. The size of the datasets was $47 \pm 6$ (mean $\pm$ std).

the change from cooperative to uncooperative rupture of the fiber, as apparent from the more gradual rupturing plateau, is characteristic for the absence of nucleosome stacking in a 2-start structure. As expected, $\Delta G_2 = 6 \pm 1\ k_{\mathrm{B}}T$ was not affected by H4 tail deletion ($p = 0.43$). These results indicate that nucleosome stacking in NRL 167 fibers was primarily mediated by the H4 tail, and support the interpretation that short non-10$n$ linker chromatin fibers are best described by a 0-start state, in which nucleosomes do not stack.

Surprisingly, folding of NRL 197 fibers was unaffected by H4 tail deletion, as shown by the overlapping curves in Figure 5.7b. Both WT and gH4 fibers unfolded gradually at approximately 3 pN. Fitting the model to gH4 fibers yielded $k = 0.3 \pm 0.1$ pN/nm and $\Delta G_1 = 20 \pm 2\ k_{\mathrm{B}}T$. These values were the same for WT H4 histones ($p = 0.98$ and $p = 0.35$). $\Delta G_2$ was also unaffected by H4 tail deletion ($p = 0.41$). Note that in fibers reconstituted with gH4, $\Delta G_1$ was larger in NRL 197 fibers than in NRL 167, indicating that despite the lack of the H4 tail, other nucleosome-nucleosome interactions contributed to an increased rupture energy. As opposed to the NRL 167 fibers, stiffness $k$ was unaffected, supporting the interpretation that additional contacts mediate the stacking of nucleosomes in NRL 197 chromatin fibers. Thus, 197 NRL fibers are organized differently than 167 NRL fibers, and 197 NRL fibers reach similar compaction when H4 tails are deleted.

### 5.2.5 Not only linker length, but also linker sequence modulates fiber folding

Since our rbMC simulations suggest that deformation of the linker DNA defines nucleosome stacking, and knowing that the mechanical properties of DNA depend on its sequence, we expect that the sequence of the linker DNA may play an important role in the stability of chromatin fiber folding. Two batches of chromatin fibers with NRL 167, consisting of different linker DNA sequences but a similar number of repeats were reconstituted, stretched, and fitted with the statistical mechanics model. They were compared with another batch of chromatin fibers with identical linker sequence, but with a different number of repeats of the 601 sequence. The fitted parameters are depicted in figure 8a-d. A similar approach was followed for three batches of chromatin fibers with NRL 197, depicted in Figures 5.8e-h. We observed that the difference in linker DNA sequence (*see* Materials and Methods) did not affect the measured stiffness of the fiber, indicating the same folding. All fits yielded approximately the same $\Delta G_2$. The rupture energy $\Delta G_1$, however, differed significantly: approximately 6 $k_{\mathrm{B}}T$ ($p < 0.001$). Despite that the mechanical properties of the folded fibers did not depend on the linker DNA sequence, the rupture force did. The difference in $\Delta G_1$ should predominantly be attributed to the different sequence-dependent mechanical properties of the linker DNA. This modulation of stability of the folded fiber could play a role in epigenetic regulation *in vivo*. However, a full




## Figure 5.8

**The unstacking energy of both NRL 167 and NRL 197 fibers, but not the stiffness, depends on the linker DNA sequence.** Force-extension curves of three populations of NRL 167 and NRL 197 fibers with either a different linker DNA sequences or a different number of repeats (*see* Materials and Methods) were fitted to Equation 5.15, yielding the mechanical parameters plotted in the decorated boxplot. The number of 601 repeats (a, e) (167: $p < 0.005$, 197: $p < 0.005$) did not affect stiffness $k$ (b, f) (167: $p = 0.43$, 197: $p < 0.005$) or interaction energy $\Delta G_2$ (d, h) (167: $p < 0.005$, 197: $p < 0.005$), which featured a narrow and overlapping distribution. Linker DNA sequence, however, strongly affected the unstacking energy $\Delta G_1$ which differed over 6 $k_B T$ between populations (167: $p < 0.005$, 197: $p < 0.005$). The $p$-values were calculated with one-way ANOVA, which is rather strict for small sample size. Visual inspection, however, reveals that the mentioned trends are apparent.

understanding of how the linker DNA sequence directs higher-order folding requires studying linker sequence variations in a more systematic manner, which will be explored in future work.

## 5.3 Discussion

The role of the linker DNA on the topology of chromatin fibers has been debated extensively, but generally in a coarse and qualitative manner. Here, we combined computational modeling with experimental single-molecule force spectroscopy and studied the effects of single base pair increments on chromatin fiber folding. Using rbMC, we showed that the energy cost of deforming the linker DNA to enable nucleosome stacking decreased with increasing linker length, as the resulting stress could be distributed over a larger number of base pairs. The linker DNA energy in stacked nucleosomes was generally dominated by contributions of the rotational step parameters. We showed that the twist of the linker DNA is an important factor in fiber folding, demonstrated by the 10 base pair periodicity in 1-start fibers, where the most favorable states were approximately 10 $k_{\mathrm{B}}T$ lower in energy than the most unfavorable ones. A similar 10 base pair periodicity was observed in nucleosome stacking in 2-start fibers. We conclude that 1-start fibers preferred to twist the linker DNA, where 2-start fibers compromised stacking when the length of the linker was not optimal.

Experimentally, we measured a 10 base pair modulation in stiffness and stacking energy of folded fibers for both short and long linker lengths. In fibers with short linker DNA, the modulations inhibited stacking between nucleosomes, which is apparent from the reduced stiffness and rupture energy. For linker DNA between 45 and 55 base pairs, the modulations were smoother, indicating that longer linker DNA can more easily accommodate the higher-order structure imposed by nucleosome stacking. The experimental data qualitatively matched well with the rbMC simulations, despite the somewhat naive implementation of the nucleosome-nucleosome interaction potential. In principle, the simulations could be used to reproduce the entire force-extension curves that we measured. However, previous work showed that the current implementation of the simulations could not reach equilibrium conditions [32], which is observed experimentally for these transitions. Moreover, the harmonic nucleosome-nucleosome interaction potential used in our rbMC model is probably underestimating the energy for larger deviations from the equilibrium geometry. The very small periodic modulation in stacking energy for short NRLs is therefore likely an underestimation. A more confined stacking interaction increases the energy for non-optimal stacking configurations and would differentiate stronger between viable and non-viable states. The experimental curves on short NRLs indeed featured a more abrupt inhibition of stacking.

Note that the fitted rupture energies cannot be compared directly with the energies that were used as inputs for the rbMC simulations. The rupture energies that we fitted represent the difference in free energy between stacked and unstacked nucleosomes, including all contributions, such as deformation of the nucleosomal DNA, histone-DNA interactions, DNA-DNA repulsion and deformation of the linker DNA. In the rbMC simulations on the other hand, the energies were assigned to specific contacts, and deformation energies of the linker DNA were implicitly modeled from the standard deviation of the step parameters and computed afterward. Nevertheless, both the fitted unstacking and partial unwrapping energies were in roughly in range with the input parameters of the rbMC simulations.

Nucleosome-nucleosome stacking is thought to be primarily, but not exclusively, mediated by interactions between the H4 histone tail and its adjacent nucleosome. To improve our rbMC model, we therefore propose that stacking can better be described by multiple freely-jointed chains that represent the histone tails, rather than a harmonic nucleosome-nucleosome interaction potential. The H4-deletion experiments on NRL 197 fibers suggest however that additional histone interactions play a role in nucleosome stacking. The large and positively charged H3 tail is likely to interact with the negatively charged linker DNA for example [34, 35]. This may severely affect the mechanical properties of the DNA and would invalidate the use of standard step parameters for rbMC simulations of linker DNA. Similar effects can be anticipated for linker histones and other additional DNA organizing proteins, that we did not include in this study. The rather large impact of the linker DNA on the orientation of adjacent nucleosomes, *e.g.* 36 degrees of twist per base pair, probably makes that our naive interaction potential can still reproduce major experimental features.

To date, structural data on nucleosome-nucleosome stacking predominantly show short-range interactions that feature close packing of interacting nucleosomes [36–39] similar to the packing in crystal structures [40]. These data were generally obtained using nucleosome core particles rather than chromatin fibers, which lifts geometrical constraints of the linker DNA. Experimental data on fibers have only reported close-packing in tetra-nucleosomes [5, 6], showing that such constraints may compromise stacking. It should also be taken into account that the buffer conditions that are required for successfully obtaining high-resolution structural data may drive the fiber into particular states. Here, we measured in buffer conditions that mimic *in vivo* conditions. Interestingly, recent single-molecule FRET experiments on NRL 167 chromatin fibers reported a very dynamic arrangement, that was interpreted as rapid changes in the register of tetranucleosomal fiber sub-structures [41], highlighting the transient nature of nucleosome stacking interactions. This more dynamic and on average less tightly packed structure is fully consistent with our current findings.

The ability to account for minimal differences in composition between fibers,

as well as the ability to resolve different mechanical aspects of fiber folding, like stiffness, rupture energy, and cooperativity of unfolding, gives single-molecule force spectroscopy a competitive edge with respect to bulk techniques like sedimentation velocity measurements [31, 42, 43], native gel electrophoresis [44] and Small Angle X-ray Scattering [6, 45]. Furthermore, although only a single dimension of the fiber is measured, the nanometer accuracy and the direct manipulation of individual fibers yields quantitative, structural insight. We show here that sufficiently refined coarse-grain modeling provides a powerful combination that helps to resolve the dynamic structure of chromatin higher-order folding.

It is remarkable that our statistical mechanics model, with only three free fitting parameters, could capture all the different fiber topologies and compositions. Previously we have only validated the model for two NRLs [10, 11, 30]. Despite our best efforts to maintain the integrity of the fibers, we observed some variation in fiber fit parameters between fibers reconstituted from DNA with the same sequence. A large fraction of these variations was captured by taking into account the number of nucleosomes. Nevertheless, there was some correlation between $k$ and $\Delta G_1$ within each NRL, as depicted in Supplementary Figure 5.S2. Interestingly, the correlation for NRL 167, NRL 176 and NRL 177 was relatively small: $R = 0.46$. 197 NRL fibers had a larger correlation coefficient: $R = 0.82$. Ideally, $k$ and $\Delta G_1$ would be uncorrelated since the statistical mechanics model should separate mechanical properties from fiber composition. However, when fibers are not fully saturated, defects within a fiber would affect fiber folding. Defects at different locations could induce systematic variations that cannot be captured in our statistical physics model, and lift the independence of the fit parameters. It is noteworthy that the fibers that were attributed to fold in a 2-start topology, which corresponds to cooperative (un-)stacking [30], featured the lowest correlation in fit parameters. Such cooperative stacking would reduce both compositional defects, as nucleosomes are better stabilized, and provide additional robustness against stretching and unstacking, having two stacks of nucleosomes in series. Fibers that fold in 1-start topologies would therefore be more susceptible for stacking defects, leading to convoluted fitting parameters. Thus, the larger correlation between the stiffness and the first rupture energy may indirectly point to 1-start folding.

By comparing the force-extension curves of equal NRL fibers with different DNA sequence, we showed that mechanical properties of the linker DNA have a defining role in chromatin fiber folding. This may have important consequences *in vivo*. A-tracts, for instance, are predominantly found in the linker DNA regions of chromatin [46]. Their higher stiffness and intrinsic curvature drive nucleosomes into flanking positions [47]. As a secondary effect, these A-tracts may drive nucleosomes into orientations that can prohibit or enhance stacking, depending on the precise position and length of the A-tract and linker length, and therefore also control higher-order folding. More global effects can also

be envisioned: the most abundant linker lengths *in vivo* are $10n + 5$ [19, 21], which we show here to fold into 0-start, open chromatin fibers for short linker lengths. Increasing the linker length promotes 1-start topologies. Nucleosome positioning may thus modulate the accessibility of DNA locally and/or globally. It can result in direct effects, by wrapping the DNA in a nucleosome, or indirect effects, by controlling nucleosome stacking and the higher-order structure of the chromatin fiber.

The debate on the role, abundance, and existence of the 30-nm chromatin fiber *in vivo*, has not settled. Current experimental data, based on sophisticated EM and chromatin capture techniques, tend towards dismissing such structures [2]. In this study we show, that especially for short NRLs, a nucleosome shift of a single base pair can make the difference between stacked and non-stacked nucleosomes. The recent study by Baldi *et al.* [19] suggests that regular NRLs are more prominent than previously thought. However, one would need single base pair positioning accuracy for multiple nucleosomes in individual chromatin fragments to determine if this affects fiber formation and current technology does not allow for such accuracy. Pending such data, it is more likely to assume that irregular linker lengths result in irregular nucleosome stacking patterns along the fiber. Single-molecule force spectroscopy experiments on natively assembled chromatin [48] showed that despite such anticipated irregularity nucleosome stacking is prominent. It is therefore more compelling to discuss chromatin folding in terms of nucleosome stacking pairs that are defined by, amongst other features, the linker length, rather than imposing long-range regularity to understand chromatin higher-order structure [49].

The sequence-dependence of nucleosome stacking, both by the positioning of nucleosome pairs and by modulation of the bendability and twistability of the linker DNA, brings up an intriguing possibility. Next to post-translational histone modifications, and remodeler-induced nucleosomal rearrangements, the DNA sequence itself may partially encode its (lack of) higher-order structure. Therefore, our genome may have evolved to mechanically favor particular higher-order organizations. Future studies will need to verify this hypothesis. Here we show that such features require single base pair accuracy, as a shift of a nucleosome with one base pair can make a major difference in chromatin higher-order structure.

## 5.4 Materials and Methods

### 5.4.1 rbMC simulations

Chromatin fibers were simulated using a rigid base pair model, expanded to include (un)wrapping of nucleosomal DNA and (un)stacking of nucleosomes in 1-start and 2-start chromatin fibers [32]. Chromatin fibers under force, as shown in Figure 5.1, were simulated following the approach by De Jong *et al.*

[32]. Fibers with 16 repeats of NRL 167 and NRL 197 were simulated without nucleosome stacking (0-start), stacked in a 1-start structure, or stacked in a 2-start structure. Rupture energy was set at 25 $k_BT$, and interaction energy at 2.5 $k_BT$ per contact point. More details can be found in supplementary information.

### 5.4.2 DNA substrates

The DNA substrates, containing an array of 601 nucleosome positioning sequences [27], were constructed as described by *et al.* [28]. An insert containing a single nucleosome positioning sequence was cloned between the XbaI and EcoRI site in a pUC18 plasmid. The plasmid was split in two fractions: an acceptor and a donor plasmid. The acceptor plasmid was digested with BstXI and DraIII and purified. The donor plasmid was digested with BstXI and BglI, which produced a backbone and an insert containing the desired sequence. We introduced phosphatase to prevent self-ligation. The insert was purified and ligated into the acceptor plasmid by T4 ligase. The ligated plasmid then contained two identical inserts. The plasmid was subsequently transformed into *Escherichia coli* DM1-blue competent cells, amplified and harvested. Repeating these steps doubled the number of inserts in the plasmid, which was done three times to create arrays of 16 inserts. Sequences are listed in supplementary information.

### 5.4.3 Chromatin reconstitution

Plasmids were linearized with BsaI and BseYI, creating linear arrays with $\sim$ 1000 base pairs of DNA handles on each side. A digoxigenin was introduced at the BsaI site and a biotin at the BseYI site with a Klenow reaction [50]. Chromatin was reconstituted following [52]. DNA and human WT histone octamers (Epicypher, Durham, USA) were added together in 2M sodium buffer (2M NaCl / 1x TE buffer), which was slowly dialyzed to 200 mM. The level of saturation of the chromatin fibers depended on the ratio between 601 nucleosome positioning sequences and histone octamers: excess histones yielded additional tetrasomes and nucleosomes on the DNA handles. A too low stoichiometry resulted in under saturated fibers. More details and a step-by-step protocol for DNA labeling and chromatin reconstitution can be found in [51]. Drosophila WT and gH4 histones were purified and histone octamers were purified as described [52].

### 5.4.4 Force spectroscopy on single chromatin fibers

All force spectroscopy experiments were done on a home-built microscope. Details can be found in supplementary information. The force-extension curves were measured by increasing the force to 9 pN, releasing the force to 0.5 pN,

and increasing the force to 55 pN, and releasing again. This trajectory was chosen to minimize the sticking of the beads to the cover slip. Usually, a small force was sufficient to unstick a bead. A force of 0.5 pN was enough to keep the bead suspended above the cover slip whilst still fully compacting the fiber. The magnets were moved at a constant speed of 0.25 mm per second, which increased and decreased the force exponentially.

### 5.4.5 Statistical mechanics model

To infer mechanical properties of individual chromatin fibers the force-extension curves were fitted to a statistical mechanics model developed by Meng *et al.* [30]. The model and fitting procedure are described in supplementary information.

## 5.5 Acknowledgments

The authors would like to thank Chenyi Wu for his help in cloning novel 601 arrays and Klaas Hermans and Nikolay Korolev for valuable discussions. 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.

## 5.6 Supplementary Material

### 5.6.1 rbMC simulations

Chromatin fibers were simulated using a rigid base pair model, expanded to include (un)wrapping of nucleosomal DNA and (un)stacking of nucleosomes in 1-start and 2-start chromatin fibers [32]. For each structure, 1000 Monte Carlo iterations were calculated while the fibers were exposed to 3 pN of force. To speed up the simulations, the contour length was limited to 1800 base pairs, resulting in DNA handles that were shorter than in the experiments. Chromatin fibers in the absence of force, as shown in Figures 5.2-5.4, were simulated with 8 601 repeats. 10,000 Monte Carlo iterations were calculated, where each step corresponded to sequential replacement and evaluation of every base pair in the DNA tether. Energy penalties were included for steric clashes between nucleosomes and excluded volume effects. Every 200 iterations, an independent structure was stored for further analysis. To represent the typical conformation of a set of nucleosomes, step parameters of all 10,000 independent states were averaged. The standard deviations were calculated from the energies of all stored states. The total energy was calculated as the sum of the base pair step energies in the 6 degrees of freedom plus the wrapping and stacking energies. The energies depicted in Figure 5.4 were calculated relative to the 0-start fiber.

### 5.6.2 DNA substrates

#### NRL 167x16 — NRL 177x16 and NRL 192x16 — NRL 202x16 series

All plasmids in this series share the 601 nucleosome positioning sequence [27] (underlined) and a basis linker sequence of 20 base pairs (red), optionally extended in single base pair steps with the sequence in lowercase:

[ 5' CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGCAAGCTCTAGCACCGCTTA
AACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGC
ACGTGTCAGATATATACATCCTGTctgactggaccctatacgcggtattgaatcgacacCTTGC
CACCCCGGGCTGTGA 3' ]$_{16}$

#### Linker sequence comparison

For comparing the effect of the linker DNA sequence on fiber folding the following substrates were used:

##### 16 repeats of NRL 167 with linker DNA sequence 1

[ 5' CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGCAAGCTCTAGCACCGCTTA
AACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGC
ACGTGTCAGATATATACATCCTGTCTTGCCACCCCGGGCTGTGA 3' ]$_{16}$

##### 15 and 30 repeats of NRL 167 with linker DNA sequence 2

[ 5' CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGCAAGCTCTAGCACCGCTTA
AACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGC
ACGTGTCAGATATATACATCCTGTGCATGTACTCGGGGGCCGCC 3' ]$_{15/30}$

##### 16 repeats of NRL 197 with linker DNA sequence 3

[ 5' CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGCAAGCTCTAGCACCGCTTA
AACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGC
ACGTGTCAGATATATACATCCTGTtggaccctatacgcggtattgaatcgacacCTTGCCACCC
CGGGCTGTGA 3' ]$_{15/30}$

##### 15 and 25 repeats of NRL 197 with linker DNA sequence 4

[ 5' CTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGCAAGCTCTAGCACCGCTTA
AACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGC
ACGTGTCAGATATATACATCCTGTGCATGTATTGAACAGCGACCTTCTCGGGATGGACCCTATA
CGCGGCCGCC 3' ]$_{15/25}$

### 5.6.3 Magnetic tweezers setup

A multiplexed magnetic tweezers setup equipped with a NIKON CFI Plan Fluor objective MRH01401 (NA = 1.3, 40×, Oil, NIKON Corporation, Tokyo, Japan) was used to stretch our chromatin fibers. The field-of-view (FOV) was captured on a 25 Mpix Condor camera (cmv5012-F30-S[D]-M[F]-P8, CMOS Vision GmbH, Schaffhausen, Switzerland) using an infinity-corrected tube lens ITL200 (Thorlabs, Newton, USA). The flow cell was illuminated with a 100 mW, 645 nm collimated LED (LED-1115-ELC-645-29-2, IMM Photonics GmbH, Unterschleißheim, Germany). The large FOV allowed for hundreds of fibers to be measured simultaneously. The camera was connected to a PCIe-1433 (National Instruments, Austin, USA) plugged in a T7610 PC (Dell, Round Rock, USA) equipped with a ten-core Intel Xeon 2.8 GHz processor (E5-2680 v2, Intel, Santa Clara, USA) and 32GB DDR3 memory. The setup measured the full FOV at 30 fps. Samples were measured in custom-built flow cells, mounted on a multi-axis piezo scanner P-517.3CL (Physik Instrumente GmbH & Co. KG, Karlsruhe, Germany). Forces were exerted by a pair of N50 5 mm cube magnets (W-05-N50-G, Supermagnete, Webcraft GmbH, Gottmadingen, Germany).

Force $F$ was calculated using:

$$F(h) = F_{\max}\left(0.7\exp\left(-\frac{h}{c_1}\right) + 0.3\exp\left(-\frac{h}{c_2}\right)\right), \tag{5.1}$$

where $h$ was the magnet height, and $c_1 = 1.4$ and $c_2 = 0.8$ were decay lengths describing the double-exponential decay of the magnetic field and $F_{\max} = 85$ pN for 2.8 µm paramagnetic beads (Dynabeads M270 Streptavidin, Thermo Fisher Scientific). Forces were calibrated using equipartition theorem, as previously described [50, 53, 54].

### 5.6.4 Statistical mechanics model

To infer mechanical properties of individual chromatin fibers the force-extension curves were fitted to a statistical mechanics model developed by Meng *et al.* The model reduced chromatin fibers to their fundamental units: individual nucleosomes that could either be in a stacked conformation (I), partially unwrapped conformation (II), singly wrapped conformation (III), and fully unwrapped conformation (IV), as shown in Figure 5.4a. Each conformation $i$ had a corresponding extension and free energy which defines the thermodynamic equilibrium.

The probability $P_i(F)$ that a nucleosome is in conformation $i$ was defined as:

$$P_i(F) = \frac{1}{Z}\sum_i \exp\left(\frac{-g_i(F, z_i) + W}{k_{\mathrm{B}}T}\right), \tag{5.2}$$

where the partition function was defined as $Z = \sum_i \exp(-(g_i(F, z_i) + W)/k_\mathrm{B}T)$ [55], and $g_i(F, z_i)$ represents the free energy of conformation $i$. $W$ is the work done by the bead corresponding to:

$$W = \int_0^z F(\tilde{z})\mathrm{d}\tilde{z} = Fz - \int_0^F z(\tilde{F})\mathrm{d}\tilde{F}. \tag{5.3}$$

The nucleosomes are strongly compacted in the stacked conformation (conformation I). The extension $z_\mathrm{fiber}$ under force $F$ was described by a Hookean spring:

$$z_\mathrm{fiber}(F) = \frac{F}{k} + z_0, \tag{5.4}$$

where $k$ described the stiffness of the chromatin fiber and $z_0$ the extension of the folded fiber at $F = 0$. $z_0$ was fixed at a nucleosome line density of 1.5 nm. The free energy $g_\mathrm{fiber}$ under $F$ was defined as:

$$g_\mathrm{fiber}(F) = Fz_\mathrm{fiber}(F) - \int_0^F z_\mathrm{fiber}(\tilde{F})\mathrm{d}\tilde{F} = Fz_\mathrm{fiber}(F) - \frac{F^2}{2k}. \tag{5.5}$$

It required rupture energy $\Delta G_1$ to break the stacking interactions, unfold the fiber, and unwrap the outer turn of DNA from the histone core, yielding the partially unwrapped conformation (conformation II).

In the partially unwrapped conformation, the extension per nucleosome was dominated by the stretching of the bare DNA that was wrapped around the histone core, and could therefore be described by the WLC:

$$z_\mathrm{WLC}(F, L) = L\left(1 - \sqrt{\frac{k_\mathrm{B}T}{FP}} + \frac{F}{S}\right), \tag{5.6}$$

where $L$ is the contour length, $P$ the persistence length, and $S$ the stretch modulus of the DNA ($P = 50$ nm, $S = 1000$ pN [30, 56, 57]). In this partially unwrapped conformation, we set $L_1 =$ NRL - 132 base pairs.

The free energy of a stretched DNA molecule yields:

$$\begin{aligned} g_\mathrm{WLC}(F, L) = Fz_\mathrm{DNA}(F) - \int z_\mathrm{WLC}(F)\mathrm{d}F = \\ Fz_\mathrm{DNA}(F) - L\left(F - \sqrt{\frac{Fk_\mathrm{B}T}{P}} + \frac{F^2}{2S}\right). \end{aligned} \tag{5.7}$$

Therefore, the free energy of a nucleosome in the partially unwrapped conformation follows:

$$g_1(F, L_1) = g_\mathrm{WLC} + \Delta G_1. \tag{5.8}$$

The extension of a singly wrapped nucleosome was calculated with Equation 5.6 using $L_2 =$ NRL – 79 base pairs and free energy:

$$g_2(F, L_2) = g_\mathrm{WLC} + \Delta G_1 + \Delta G_2, \tag{5.9}$$

where $\Delta G_2$ was the interaction energy of the contact points broken in this transition.

The extension of the fully unwrapped conformation is equal to the extension of a DNA fragment with a length equal to the NRL [32, 58]. The free energy of this conformation was calculated by:

$$g_3(F, L_3) = g_{\mathrm{WLC}} + \Delta G_1 + \Delta G_2 + \Delta G_3, \tag{5.10}$$

with $\Delta G_3$ representing the wrapping energy of the inner wrap of DNA. The transition between the singly wrapped conformation and the fully unwrapped conformation was not in thermodynamic equilibrium as apparent from the as a stepwise unfolding transition. Therefore, the change in free energy of this transition could not be captured by the statistical mechanics model.

The total extension $Z_{\mathrm{tot}}(F)$ and total free energy $G_{\mathrm{tot}}(F)$ were calculated by a summation of all the extensions and free energies in all conformations:

$$Z_{\mathrm{tot}}(F) = \sum\nolimits_j n_j z_j(F) \tag{5.11}$$

$$G_{\mathrm{tot}}(F) = \sum\nolimits_j n_j g_j(F). \tag{5.12}$$

The force spectroscopy data were fitted with the Boltzmann distribution of conformations to obtain the mechanical properties, the total extension of the tether $\langle z_{\mathrm{tot}}(F)\rangle$, and the energy changes between the conformations:

$$\langle z_{\mathrm{tot}}(F)\rangle = \frac{1}{Z}\sum\nolimits_j z_j(F)\exp\left(-\frac{g_j(F)+W}{k_{\mathrm{B}}T}\right). \tag{5.13}$$

For a chromatin fiber with more than two nucleosomes, the nucleosomes can be distributed over a large, but finite number of states $s = n_I, n_{II}, n_{III}, n_{IV}$. Multiple indistinguishable states with an equal number of nucleosomes in each of these conformations were accounted for by a degeneracy factor $D(s)$:

$$D(s) = \prod_{i<j}\binom{n_i + n_j}{n_i} \tag{5.14}$$

where $i$ and $j$ were pairs of conformations in each state. The degeneracy factor was included in the Boltzmann distributions of states:

$$\langle z_{\mathrm{tot}}(F)\rangle = \frac{\sum_s Z_{\mathrm{tot}}(F)D(s)\exp(-(G_{\mathrm{tot}}(F)+W)/k_{\mathrm{B}}T)}{\sum_s D(s)\exp(-(G_{\mathrm{tot}}(F)+W)/k_{\mathrm{B}}T)}. \tag{5.15}$$

Note that the unwrapping lengths were adjusted since its introduction in 2015. Careful quantification of the step size by Kaczmarczyk *et al.* [59] revealed that 79 base pairs unfold in the transition from conformation III to conformation IV, indicating the unfolding of a complete turn of DNA. Accordingly, to

maintain the same extension in the transition between conformations partially unwrapped and singly wrapped conformation, the partially unwrapped conformation was attributed to constrains 132 base pairs [40]. This resolves the nature of the mysterious *extended conformation*. Several histone-DNA contact points must be broken in this second unfolding transition, which should account for the corresponding of $\Delta G_2$. In summary, the statistical mechanics model did not change, but the attribution of the number of wrapped base pairs in the two intermediate nucleosome conformations and the interpretation of the difference in free energy between them was different, and the names of these conformations were accordingly changed to partially unwrapped and single wrapped nucleosomes.

### 5.6.5 Fitting procedure

Prior to fitting the statistical mechanics model, thermal drift was accounted for by subtraction of a linear drift that was fitted from the low force parts ($F < 0.1$ pN) at the beginning and the end of each measurement. Drift typically did not exceed 1 nm/s. By aligning the force-extension curve at $F > 40$ pN with a WLC corresponding the contour length of the DNA substrate, the offset in $z$ was determined. Next, the number of 25 nm steps at $F > 10$ pN, corresponding to the last unwrapping transition, was counted to find the total number of nucleosomal particles reconstituted on the DNA template. The number of tetrasomes was subsequently fit to the statistical physics model in the force range $0.5 < F < 2.5$ pN. In this force range, nucleosomes remain stably folded and/or stacked, so excess extension should be attributed to tetrasomes [30].

Finally, fiber unfolding was fit in the range $0.5 < F < 10$ pN with three free parameters: stiffness $k$, rupture energy $\Delta G_1$, and interaction energy $\Delta G_2$. The degeneracy $D(s)$ was fixed to either Equation 5.14 (degenerate transitions indicating a 1-start fiber) or 1 (cooperative transitions indicating a 2-start fiber), depending on the best fit and was imposed on all fibers of a particular NRL.

The fitted parameters were plotted in a decorated boxplot. The center bar represents the median of the data (Q2), the edges of the box the medians of the split populations (Q1 and Q3); referred to as the interquartile range (IQR). The whiskers of the box plot described the spread of the data: Q1 – 1.5 × IQR to Q3 + 1.5 × IQR. Boxplots were decorated with a scatter plot of all data points, offset with random scatter in the $x$-direction. The p-values were calculated with t-test and one-way ANOVA where appropriate.

### 5.6.6 Fitted parameters of figures

| NRL (bp) | $N_8$ | $N_4$ | $k$ (pN/nm) | $\Delta G_1$ ($k_BT$) | $\Delta G_2$ ($k_BT$) |
|---|---|---|---|---|---|
| 167 | 15 | 0 | $1.1 \pm 0.1$ | $23 \pm 1$ | $9 \pm 1$ |
| 168 | 20 | 6 | $0.4 \pm 0.1$ | $17 \pm 1$ | $6 \pm 1$ |
| 169 | 18 | 3 | $0.4 \pm 0.1$ | $17 \pm 1$ | $6 \pm 1$ |
| 175 | 17 | 2 | $0.4 \pm 0.1$ | $20 \pm 1$ | $7 \pm 1$ |
| 176 | 17 | 3 | $0.8 \pm 0.1$ | $26 \pm 1$ | $7 \pm 1$ |
| 177 | 17 | 1 | $0.8 \pm 0.1$ | $27 \pm 1$ | $8 \pm 1$ |
| 192 | 19 | 4 | $0.2 \pm 0.1$ | $17 \pm 1$ | $8 \pm 1$ |
| 195 | 17 | 3 | $0.3 \pm 0.1$ | $22 \pm 1$ | $8 \pm 1$ |
| 197 | 18 | 3 | $0.3 \pm 0.1$ | $24 \pm 1$ | $8 \pm 1$ |
| 198 | 18 | 4 | $0.3 \pm 0.1$ | $26 \pm 1$ | $6 \pm 1$ |
| 200 | 17 | 2 | $0.2 \pm 0.1$ | $20 \pm 1$ | $7 \pm 1$ |
| 202 | 18 | 4 | $0.2 \pm 0.1$ | $17 \pm 1$ | $6 \pm 1$ |

**Table 5.S1**
Fitted parameters for chromatin fibers in Figures 5.5c-f. The ±-sign indicates the standard error of the fit.

| NRL (bp) | $k$ (pN/nm) | $\Delta G_1$ ($k_BT$) | $\Delta G_2$ ($k_BT$) |
|---|---|---|---|
| 167 | $1.3 \pm 0.5$ | $24 \pm 3$ | $7 \pm 1$ |
| 168 | $0.5 \pm 0.1$ | $16 \pm 2$ | $7 \pm 1$ |
| 169 | $0.4 \pm 0.1$ | $17 \pm 2$ | $7 \pm 1$ |
| 170 | $0.5 \pm 0.2$ | $19 \pm 2$ | $7 \pm 1$ |
| 171 | $0.5 \pm 0.1$ | $19 \pm 3$ | $8 \pm 1$ |
| 172 | $0.6 \pm 0.2$ | $19 \pm 2$ | $7 \pm 1$ |
| 173 | $0.4 \pm 0.1$ | $18 \pm 2$ | $8 \pm 1$ |
| 174 | $0.6 \pm 0.2$ | $20 \pm 2$ | $7 \pm 1$ |
| 175 | $0.5 \pm 0.2$ | $20 \pm 2$ | $7 \pm 1$ |
| 176 | $0.9 \pm 0.5$ | $24 \pm 3$ | $7 \pm 1$ |
| 177 | $1.0 \pm 0.5$ | $25 \pm 3$ | $7 \pm 1$ |
| 192 | $0.2 \pm 0.1$ | $17 \pm 2$ | $7 \pm 1$ |
| 193 | $0.2 \pm 0.1$ | $20 \pm 2$ | $7 \pm 1$ |
| 194 | $0.3 \pm 0.1$ | $21 \pm 3$ | $7 \pm 1$ |
| 195 | $0.3 \pm 0.1$ | $21 \pm 2$ | $7 \pm 1$ |
| 196 | $0.3 \pm 0.1$ | $22 \pm 2$ | $7 \pm 1$ |
| 197 | $0.3 \pm 0.1$ | $23 \pm 2$ | $7 \pm 1$ |
| 198 | $0.3 \pm 0.1$ | $24 \pm 3$ | $7 \pm 1$ |
| 199 | $0.3 \pm 0.1$ | $23 \pm 3$ | $7 \pm 1$ |
| 200 | $0.3 \pm 0.1$ | $21 \pm 2$ | $7 \pm 1$ |
| 201 | $0.3 \pm 0.1$ | $21 \pm 3$ | $6 \pm 1$ |
| 202 | $0.2 \pm 0.1$ | $21 \pm 3$ | $7 \pm 1$ |

**Table 5.S2**
Population parameters for Figure 5.6. The ±-sign indicates the standard deviation of the populations ($n > 35$).

| NRL (bp) | $N_8$ | $N_4$ | $k$ (pN/nm) | $\Delta G_1$ ($k_BT$) | $\Delta G_2$ ($k_BT$) |
|---|---|---|---|---|---|
| 167 WT | 16 | 2 | $0.9 \pm 0.1$ | $24 \pm 1$ | $5 \pm 1$ |
| 167 gH4 | 15 | 1 | $0.3 \pm 0.1$ | $15 \pm 1$ | $5 \pm 1$ |
| 197 WT | 16 | 6 | $0.3 \pm 0.1$ | $21 \pm 1$ | $6 \pm 1$ |
| 197 gH4 | 17 | 6 | $0.2 \pm 0.1$ | $20 \pm 1$ | $6 \pm 1$ |

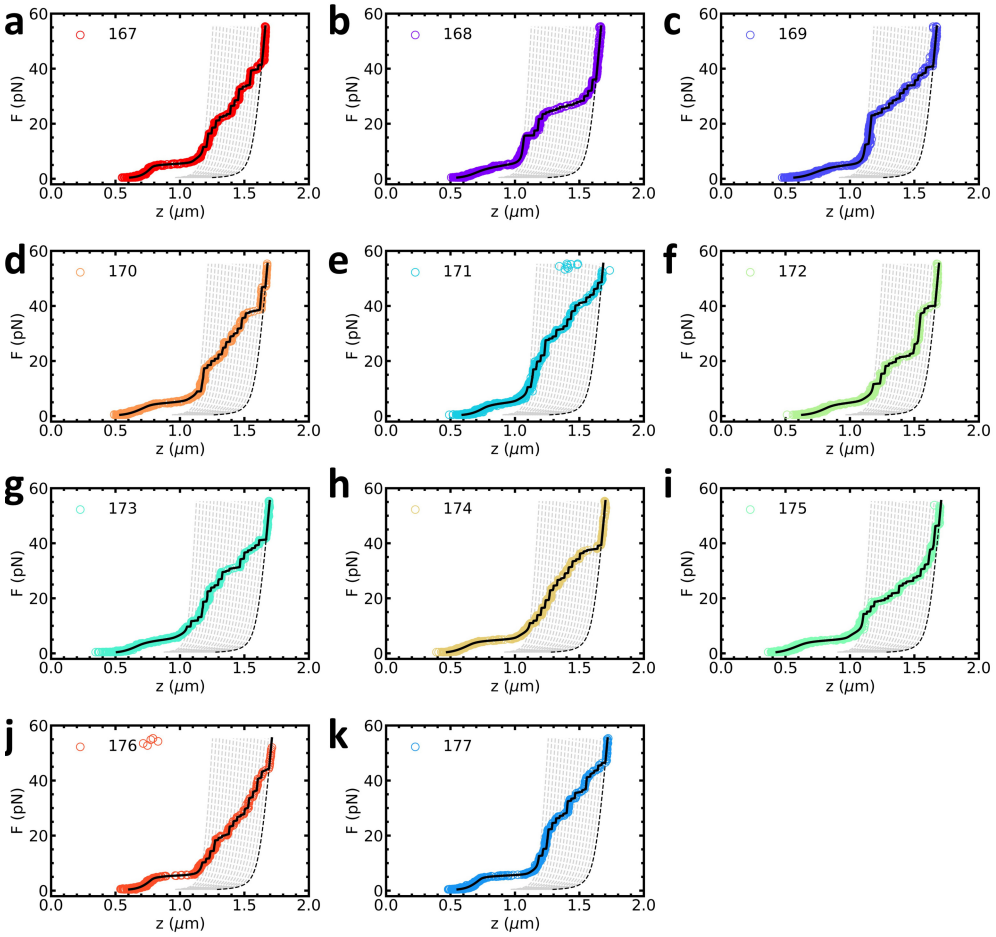
**Table 5.S3**
Fitted parameters for chromatin fibers in Figure 5.7a. The ±-sign indicates the standard error of the fit.

| NRL (bp) | histones | $k$ (pN/nm) | $\Delta G_1$ ($k_BT$) | $\Delta G_2$ ($k_BT$) |
|---|---|---|---|---|
| 167 | *Human* WT | $1.3 \pm 0.5$ | $24 \pm 3$ | $7 \pm 1$ |
| 167 | *D. melanogaster* WT | $0.8 \pm 0.4$ | $22 \pm 2$ | $6 \pm 1$ |
| 167 | *D. melanogaster* gH4 | $0.5 \pm 0.2$ | $17 \pm 2$ | $6 \pm 1$ |
| 197 | *Human* WT | $0.3 \pm 0.1$ | $23 \pm 2$ | $7 \pm 1$ |
| 197 | *D. melanogaster* WT | $0.2 \pm 0.1$ | $20 \pm 2$ | $6 \pm 1$ |
| 197 | *D. melanogaster* gH4 | $0.3 \pm 0.1$ | $20 \pm 2$ | $6 \pm 1$ |

**Table 5.S4**
Population parameters for Figures 5.7c-e. The ±-sign indicates the standard deviation of the populations ($n > 42$).

| NRL (bp) | $N_8$ | $N_4$ | $k$ (pN/nm) | $\Delta G_1$ ($k_BT$) | $\Delta G_2$ ($k_BT$) |
|---|---|---|---|---|---|
| 167 | 15 | 0 | $1.1 \pm 0.1$ | $23 \pm 1$ | $9 \pm 1$ |
| 168 | 20 | 6 | $0.4 \pm 0.1$ | $17 \pm 1$ | $6 \pm 1$ |
| 169 | 18 | 3 | $0.4 \pm 0.1$ | $17 \pm 1$ | $6 \pm 1$ |
| 170 | 17 | 0 | $0.5 \pm 0.1$ | $19 \pm 1$ | $8 \pm 1$ |
| 171 | 19 | 6 | $0.5 \pm 0.1$ | $19 \pm 1$ | $8 \pm 1$ |
| 172 | 16 | 2 | $0.4 \pm 0.1$ | $19 \pm 1$ | $8 \pm 1$ |
| 173 | 21 | 6 | $0.4 \pm 0.1$ | $20 \pm 1$ | $8 \pm 1$ |
| 174 | 19 | 1 | $0.5 \pm 0.1$ | $20 \pm 1$ | $8 \pm 1$ |
| 175 | 20 | 1 | $0.4 \pm 0.1$ | $20 \pm 1$ | $9 \pm 1$ |
| 176 | 17 | 3 | $0.8 \pm 0.1$ | $26 \pm 1$ | $7 \pm 1$ |
| 177 | 17 | 1 | $0.8 \pm 0.1$ | $27 \pm 1$ | $8 \pm 1$ |
| 192 | 17 | 0 | $0.2 \pm 0.1$ | $16 \pm 1$ | $8 \pm 1$ |
| 193 | 18 | 4 | $0.2 \pm 0.1$ | $19 \pm 1$ | $8 \pm 1$ |
| 194 | 12 | 0 | $0.3 \pm 0.1$ | $21 \pm 1$ | $9 \pm 1$ |
| 195 | 16 | 5 | $0.3 \pm 0.1$ | $21 \pm 1$ | $7 \pm 1$ |
| 196 | 19 | 7 | $0.4 \pm 0.1$ | $22 \pm 1$ | $7 \pm 1$ |
| 197 | 18 | 0 | $0.3 \pm 0.1$ | $23 \pm 1$ | $8 \pm 1$ |
| 198 | 17 | 6 | $0.5 \pm 0.1$ | $23 \pm 1$ | $7 \pm 1$ |
| 199 | 16 | 3 | $0.3 \pm 0.1$ | $24 \pm 1$ | $8 \pm 1$ |
| 200 | 16 | 1 | $0.3 \pm 0.1$ | $22 \pm 1$ | $9 \pm 1$ |
| 201 | 15 | 4 | $0.2 \pm 0.1$ | $21 \pm 1$ | $7 \pm 1$ |
| 202 | 12 | 0 | $0.2 \pm 0.1$ | $20 \pm 1$ | $8 \pm 1$ |

**Table 5.S5**
Fitted parameters for chromatin fibers in Figure 5.S1. The ±-sign indicates the standard error of the fit.


a
167
b
168
c
169
d
170
e
171
f
172
g
173
h
174
i
175
j
176
k
177
F (pN)
z (μm)
0
20
40
60
0.0
0.5
1.0
1.5
2.0

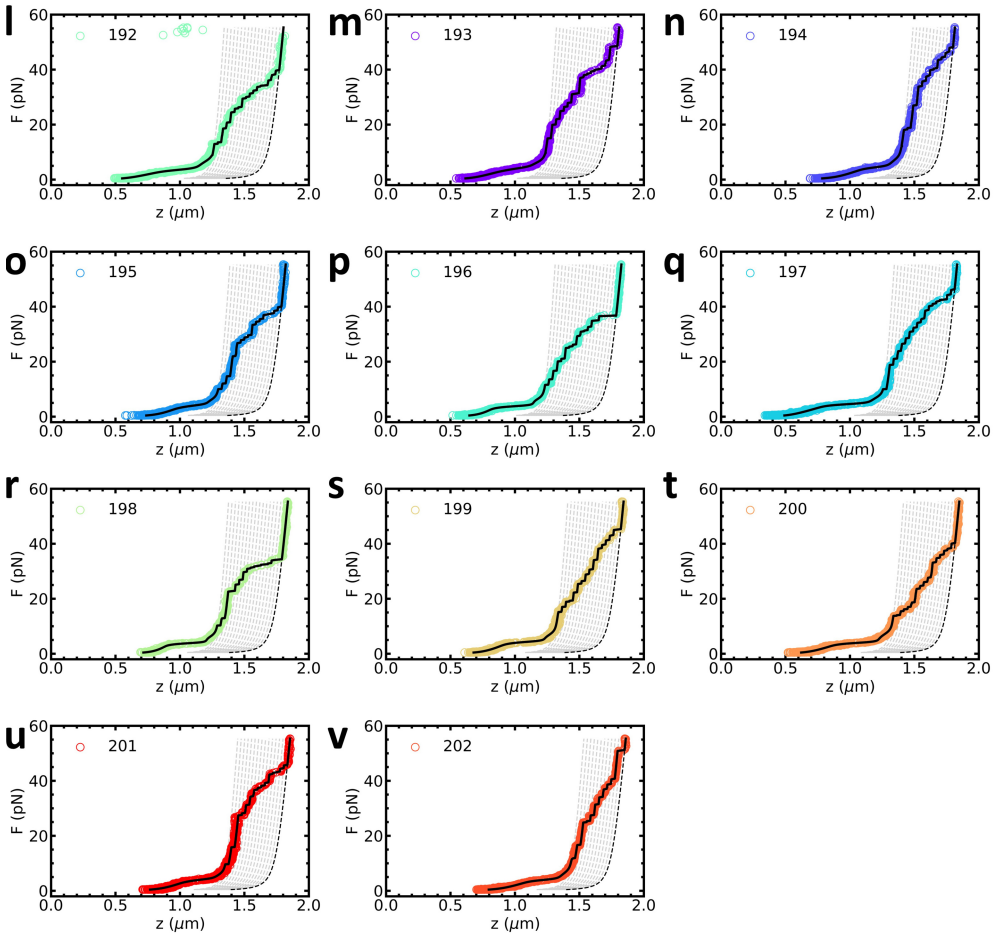



**Figure 5.S1**

**Typical force-extension curves show unique chromatin fiber unfolding features for different NRLs.** In general, all fibers unfolded at low forces ($F < 5$ pN). The fibers featured a small second transition at slightly higher forces ($F \approx 7$ pN) and stepwise unwrapping at high forces ($F > 10$ pN). Unwrapping intermediate states are indicated by the grey dashed lines. When all nucleosomes were completely unwrapped, the data followed the WLC, indicated by the black dashed line. Curves were depicted in range NRL 167 – 177 (a-k) and NRL 192 – 202 (l-v). Only the pulling trace is shown. Fitted parameters of all curves can be found in Table 5.S5 of the supplementary material. Although some tethers ruptured at high forces (*e.g.* NRL 171, 176 and 192), the curve could still be fitted and used for statistics.

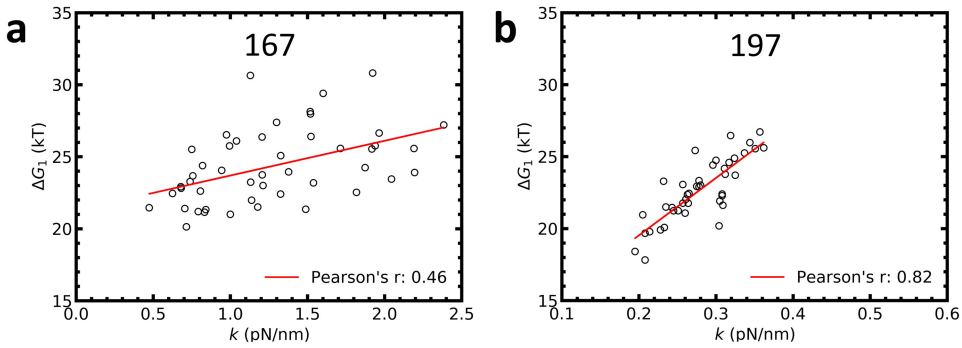



**Figure 5.S2**
**The correlation between stiffness $k$ and rupture energy $\Delta G_1$ hints at local defects in the fiber.** a) Chromatin fibers with NRL 167 featured a relatively low correlation between $k$ and $\Delta G_1$ ($R \approx 0.5$). b) The correlation was stronger for NRL 197 ($R \approx 0.8$). Such an increased correlation could point to defects in the composition of or nucleosome stacking in the fiber, which partially invalidate the statistical mechanics model, which assumes that all nucleosomes feature the same stages of unfolding.

# Bibliography V


[1] Michael G. Poirier et al. "Spontaneous Access to DNA Target Sites in Folded Chromatin Fibers". In: *Journal of Molecular Biology* 379.4 (2008), pp. 772–786. ISSN: 00222836. DOI: 10.1016/j.jmb.2008.04.025. arXiv: NIHMS150003.

[2] Beat Fierz and Michael G Poirier. "Biophysics of Chromatin Dynamics". In: *Annual review of biophysics* 48 (2019), pp. 321–345.

[3] Alan P Wolffe. "Gene regulation: insulating chromatin". In: *Current Biology* 4.1 (1994), pp. 85–87.

[4] Karolin Luger, Mekonnen L Dechassa, and David J Tremethick. "New insights into nucleosome and chromatin structure: an ordered state or a disordered affair?" In: *Nature reviews Molecular cell biology* 13.7 (2012), p. 436.

[5] 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.

[6] 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.

[7] Viviana I Risca et al. "Variable chromatin structure revealed by in situ spatially correlated DNA cleavage mapping". In: *Nature* 541.7636 (2017), p. 237.

[8] Tsung-Han S Hsieh et al. "Mapping nucleosome resolution chromosome folding in yeast by micro-C". In: *Cell* 162.1 (2015), pp. 108–119.

[9] 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.

[10] 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.

[11] 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.

[12] Kazuhiro Maeshima, Satoru Ide, and Michael Babokhov. "Dynamic chromatin organization without the 30-nm fiber". In: *Current opinion in cell biology* 58 (2019), pp. 95–104.

[13] Horng D Ou et al. "ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells". In: *Science* 357.6349 (2017), eaag0025.

[14] Kazuhiro Maeshima et al. "Liquid-like behavior of chromatin". In: *Current Opinion in Genetics and Development* 37 (2016), pp. 36–45. ISSN: 18790380. DOI: 10.1016/j.gde.2015.11.006.

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

[16] Kristin Brogaard et al. "A map of nucleosome positions in yeast at base-pair resolution". In: *Nature* 486.7404 (2012), p. 496.

[17] Eran Segal et al. "A genomic code for nucleosome positioning". In: *Nature* 442.7104 (2006), p. 772.

[18] Cizhong Jiang and B Franklin Pugh. "Nucleosome positioning and gene regulation: advances through genomics". In: *Nature Reviews Genetics* 10.3 (2009), p. 161.

[19] Sandro Baldi et al. *Genome-wide measurement of local nucleosome array regularity and spacing by nanopore sequencing*. Tech. rep. Nature Publishing Group, 2018.

[20] Sean C Clark et al. "Differential nucleosome spacing in neurons and glia". In: *Neuroscience Letters* 714 (2020), p. 134559.

[21] J Widom. "A relationship between the helical twist of DNA and the ordered positioning of nucleosomes in all eukaryotic cells." In: *Proceedings of the National Academy of Sciences of the United States of America* 89.February (1992), pp. 1095–1099. ISSN: 0027-8424. DOI: 10.1073/pnas.89.3.1095.

[22] Ognjen Perišić, Rosana Collepardo-Guevara, and Tamar Schlick. "Modeling studies of chromatin fiber structure as a function of DNA linker length". In: *Journal of molecular biology* 403.5 (2010), pp. 777–802.

[23] Elena F Koslover et al. "Local geometry and elasticity in compact chromatin structure". In: *Biophysical journal* 99.12 (2010), pp. 3941–3950.

[24] Mikhail V Bass et al. "Nucleosome spacing periodically modulates nucleosome chain folding and DNA topology in circular nucleosome arrays". In: *Journal of Biological Chemistry* 294.11 (2019), pp. 4233–4246.

[25] Davood Norouzi and Victor B Zhurkin. "Dynamics of chromatin fibers: Comparison of Monte Carlo simulations with force spectroscopy". In: *Biophysical journal* 115.9 (2018), pp. 1644–1655.

[26] Davood Norouzi and Victor B. Zhurkin. "Topological polymorphism of the two-start chromatin fiber". In: *Biophysical Journal* 108.10 (2015), pp. 2591–2600. ISSN: 15420086. DOI: 10.1016/j.bpj.2015.04.015.

[27] P T T Lowary and J Widom. "New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning." In: *Journal of molecular biology* 276.1 (1998), pp. 19–42. ISSN: 0022-2836. DOI: 10.1006/jmbi.1997.1494. eprint: NIHMS150003.

[28] Chenyi Wu et al. “The construction of customized nucleosomal arrays”. In: *Analytical biochemistry* 496 (2016), pp. 71–75.

[29] Babatunde Ekundayo, Timothy J Richmond, and Thomas Schalch. “Capturing structural heterogeneity in chromatin fibers”. In: *Journal of molecular biology* 429.20 (2017), pp. 3031–3042.

[30] 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.

[31] Sarah J Correll, Michaela H Schubert, and Sergei A Grigoryev. “Short nucleosome repeats impose rotational modulations on chromatin fibre folding”. In: *The EMBO Journal* 31.10 (2012), pp. 2416–2426. ISSN: 1460-2075. DOI: 10.1038/emboj.2012.80.

[32] Babette E de Jong et al. “Rigid Basepair Monte Carlo Simulations of One-Start and Two-Start Chromatin Fiber Unfolding by Force”. In: *Biophysical journal* 115.10 (2018), pp. 1848–1859.

[33] WJA Koopmans et al. “spFRET using alternating excitation and FCS reveals progressive DNA unwrapping in nucleosomes”. In: *Biophysical journal* 97.1 (2009), pp. 195–204.

[34] Zhenhai Li and Hidetoshi Kono. “Distinct roles of histone H3 and H2A tails in nucleosome stability”. In: *Scientific reports* 6 (2016), p. 31437.

[35] 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.

[36] Jonas J Funke et al. “Exploring nucleosome unwrapping using DNA origami”. In: *Nano letters* 16.12 (2016), pp. 7891–7898.

[37] Jonas J Funke et al. “Uncovering the forces between nucleosomes using DNA origami”. In: *Science Advances* 2.11 (2016), e1600974.

[38] Mikhail Eltsov et al. “Nucleosome conformational variability in solution and in interphase nuclei evidenced by cryo-electron microscopy of vitreous sections”. In: *Nucleic acids research* 46.17 (2018), pp. 9189–9200.

[39] S Mangenot et al. “Interactions between isolated nucleosome core particles: A tail-bridging effect?” In: *The European Physical Journal E* 7.3 (2002), pp. 221–231.

[40] 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.

[41] Beat Fierz. “Dynamic chromatin regulation from a single molecule perspective”. In: *ACS chemical biology* 11.3 (2015), pp. 609–620.

[42] Sergei A Grigoryev et al. “Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions”. In: *Proceedings of the National Academy of Sciences* 106.32 (2009), pp. 13317–13322.

[43] CL Woodcock et al. "A chromatin folding model that incorporates linker variability generates fibers resembling the native structures". In: *Proceedings of the National Academy of Sciences* 90.19 (1993), pp. 9021–9025.

[44] Benedetta Dorigo et al. "Nucleosome arrays reveal the two-start organization of the chromatin fiber." In: *Science (New York, N.Y.)* 306.5701 (Nov. 2004), pp. 1571–3. ISSN: 1095-9203. DOI: 10.1126/science.1103124.

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

[46] Feng Cui and Victor B Zhurkin. "Distinctive sequence patterns in metazoan and yeast nucleosomes: implications for linker histone binding to AT-rich and methylated DNA". In: *Nucleic acids research* 37.9 (2009), pp. 2818–2829.

[47] Eran Segal and Jonathan Widom. "Poly (dA: dT) tracts: major determinants of nucleosome organization". In: *Current opinion in structural biology* 19.1 (2009), pp. 65–71.

[48] Nicolaas Hermans et al. "Toehold-enhanced LNA probes for selective pull down and single-molecule analysis of native chromatin". In: *Scientific reports* 7.1 (2017), p. 16721.

[49] Shujun Cai et al. "The in situ structures of mono-, di-, and trinucleosomes in human heterochromatin". In: *Molecular biology of the cell* 29.20 (2018), pp. 2450–2457.

[50] Thomas B Brouwer et al. "Unraveling DNA Organization with Single-Molecule Force Spectroscopy Using Magnetic Tweezers". In: *Bacterial Chromatin*. Springer, 2018, pp. 317–349.

[51] Artur Kaczmarczyk et al. "Probing Chromatin Structure with Magnetic Tweezers". In: *Nanoscale Imaging*. Springer, 2018, pp. 297–323.

[52] Henrike Klinker et al. "Rapid purification of recombinant histones". In: *PloS one* 9.8 (2014).

[53] Aartjan J.W. W Te Velthuis et al. "Quantitative guidelines for force calibration through spectral analysis of magnetic tweezers data". In: *Biophysical Journal* 99.4 (2010), pp. 1292–1302. ISSN: 00063495. DOI: 10.1016/j.bpj.2010.06.008.

[54] ID Vilfan et al. "Magnetic tweezers for single-molecule experiments". In: *Handbook of Single-Molecule Biophysics*. Springer, 2009, pp. 371–395.

[55] Ignacio Tinoco Jr and Carlos Bustamante. "The effect of force on thermodynamics and kinetics of single molecule reactions". In: *Biophysical chemistry* 101 (2002), pp. 513–533.

[56] Carlos Bustamante, Zev Bryant, and Steven B. Smith. "Ten years of tension: single-molecule DNA mechanics". In: *Nature* 421.6921 (2003), pp. 423–427. ISSN: 0028-0836. DOI: 10.1038/nature01405.

[57] Terence Strick et al. "Twisting and stretching single DNA molecules". In: *Progress in biophysics and molecular biology* 74.1-2 (2000), pp. 115–140.

[58] Nikolay Korolev, Lars Nordenskiöld, and Alexander P. Lyubartsev. "Multiscale coarse-grained modelling of chromatin components: DNA and the nucleosome". In: *Advances in Colloid and Interface Science* (2016), pp. 1–13. ISSN: 00018686. DOI: 10.1016/j.cis.2016.02.002.

[59] A Kaczmarczyk. "Nucleosome stacking in chromatin fibers probed with single-molecule force-and torque-spectroscopy". PhD thesis. Leiden University, 2019.

# Chapter 6

# Mechanical and Structural Properties of Archaeal Hypernucleosomes

Many archaeal species express homologs of eukaryotic histones H3, H4, H2A and H2B, which organize the genome and play a key role in gene regulation. The structure and function of archaeal histone-DNA complexes remain however largely unclear. Recent X-ray crystallography studies of *Methanothermus fervidus* histone HMfB bound to DNA fragments show formation of hypernucleosomes consisting of DNA wrapped around an 'endless' histone-protein core. Such a structure is in line with analysis of digested nucleoprotein complexes suggesting that these complexes consist of an integer number of dimeric units of archaeal histones *in vivo*. However, if and how such a hypernucleosome structure assembles on a long DNA substrate and which interactions provide for its stability, remains unclear.

Here, we describe micromanipulation studies of complexes of the *M. fervidus* histones HMfA and HMfB with single DNA tethers. Our experiments show hypernucleosome assembly which results from cooperative binding of histones to DNA, facilitated by stacking interactions between neighboring histone dimers. Furthermore, rotational force spectroscopy demonstrates that the HMfB-DNA complex has a left-handed chirality, but that torque can drive it in a right-handed conformation. The structure of the hypernucleosome thus depends on stacking interactions, torque, and force. *In vivo*, such modulation of the archaeal hypernucleosome structure may play an important role in transcription regulation as a responsive to environmental changes.

## Significance statement

Archaea are highly relevant as they are one of the main pillars of life and form a link between the last universal common ancestor in life on earth (LUCA) and Eukaryotes becomes more unequivocal. In eukaryotes, histones play an important role in genome compaction and gene regulation. Archaea also express histones, but they lack a key feature of the eukaryotic histone: the N-terminal tail, important in eukaryotic transcription regulation. With no N-terminal tail, Archaea must have an alternative way of modulating DNA accessibility, although the archaeal transcription regulation mechanisms are poorly understood. We investigated how archaeal histones compact DNA. Its mechanisms of compaction are suggestive of primordial bacterial-like transcription regulatory mechanisms preceding eukaryotic regulatory mechanisms.

This chapter is based on: Henneman B., Brouwer T. B., van Emmerik C., van der Valk R. A., Timmer M., Kirolos N., van Ingen H., van Noort S. J. T., Dame R. T.: Mechanical and Structural Properties of Archaeal Hypernucleosomes, *submitted*.

## 6.1 Introduction

Dynamic genome organization is a prerequisite for a compact yet active genome throughout all domains of life. Current views on evolution agree that eukaryotes are part of the branch of Archaea, which are single-cellular organisms that share many cellular mechanisms with Eukaryotes [1–4]. Eukaryotes use histones, which wrap DNA into nucleosomes to compact and functionally organize their genomes. These histones have N-terminal tails that can be post-translationally modified, which changes their physico-chemical properties, and is key in defining the functional chromatin state [5]. Most archaeal species express rudimentary homologs of eukaryotic histones with the characteristic histone fold [6–8], but lack the N-terminal tail [9–13]. In addition to histones, Archaea express other small architectural proteins, nucleoid-associated proteins (NAPs). These NAPs are also involved in genome organization and may complement or compete with histones to regulate genes [14, 15].

The prototypical archaeal histones are HMfA and HMfB from *Methanothermus fervidus* [16]. HMfA and HMfB are homologs of eukaryotic histones in terms of sequence and structure [13]. HMfA and HMfB share 84% sequence identity. These proteins are expressed at varying ratios as a function of growth phase [17]. Although estimates of absolute expression levels of individual histones are lacking, HMfA is prevalent during exponential growth (1.5× the amount of HMfB), and present in equal amounts as HMfB during stationary growth [17]. This might imply that these proteins have evolved different DNA binding properties – and correspondingly distinct functions – in the cell. HMfA and HMfB are capable of forming homo- and heterodimers, which subsequently can form larger structures [17–20]. The binding of HMf dimers to DNA results in bends in DNA at low protein concentrations, and has been reported to yield beads-on-a-string nucleosome-like structures at higher protein to DNA ratios [16]. HMf proteins bind preferentially to intrinsically curved DNA sequences *in vitro* [21]. There is currently no evidence of the existence of specific DNA sequence signatures that enhance histone binding *in vivo*, although studies in *Methanothermobacter thermautotrophicus* and *Thermococcus kodakarensis* indicate that repetitive motifs of AT and GC base pair steps favor histone positioning [22]. These motifs facilitate the distortion needed for DNA wrapping. Systematic evolution of ligands by exponential enrichment (SELEX) experiments have yielded two high-affinity DNA binding sequences for HMfB *in vitro* [23]. Binding of HMf proteins to this sequence yields nucleosome-like structures, in which 60 base pairs of DNA is proposed to be wrapped around a tetrameric protein core composed of two interacting HMf dimers [24]. Here, we will refer to the DNA compaction mode of HMf proteins as wrapping, even when the DNA is wrapped less than one turn.

The histone tetramer model is supported by micrococcal nuclease (MNase) digestion studies of chromatin from *Haloferax volcanii*, which yield undigested

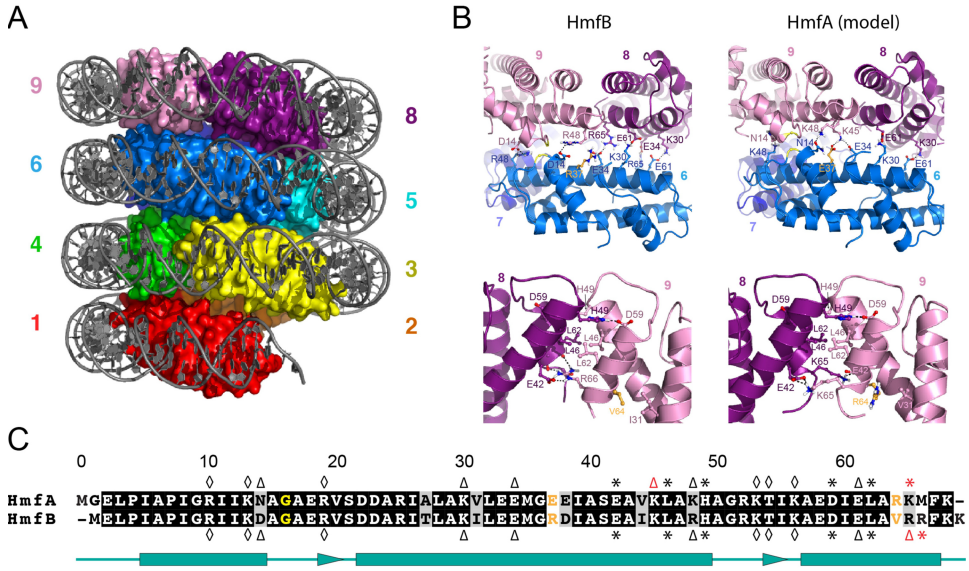



**Figure 6.1**
**The hypernucleosome and *Methanothermus fervidus* histones HMfB and HMfA.** a) The HMfB hypernucleosome structure as described by Mattiroli *et al.* (PDB: 5T5K [25]). Figure taken from Henneman *et al.* 2018 with permission [26]. B) Hypernucleosome interfaces of HMfB (left; water-refined and starting from the 5T5K structure) and HMfA (right; water-refined and starting from the homology model). Top panels show the stacking interface, bottom panels show the tetramer interface. Interacting residues are labeled; hydrogen bonds are shown by dashed lines. Important residue differences between HMfA and HMfB are colored in orange. The closely packed residues G16 in the stacking interface are shown in yellow spheres (Cα only). C) Sequence alignment of HMfA and HMfB. Conserved residues are shown in black boxes, conservative changes in grey and non-conservative changes in white boxes. Residues are numbered according to HMfB sequence. The two sequences are 84% identical and have 94% similarity. Residues directly contacting the DNA are indicated with diamonds (◊), residues forming the tetramerization interface are marked with asterisks (∗) and stacking residues are indicated with triangles (△) for HMfA (top) and HMfB (bottom). Differences in interacting residues between HMfA and HMfB are indicated with red symbols. Important residue differences between HMfA and HMfB are colored in orange and G16 residues are shown in yellow, matching panel B.

DNA fragments 60 base pairs in size [22, 27]. However, the histone tetramer model has been challenged by the results of similar studies in *T. kodakarensis*, which yield DNA fragments ranging from 30 – 450 base pairs in 30 base pairs increments [28]. The sizes of these DNA fragments were interpreted as reflecting the existence of histone multimers of varying sizes bound along the genome. A notable recent crystallography study of HMfB on 90 base pair DNA fragments supports the formation of a multimeric histone filament established by interactions between adjacent histone dimers and of wrapping of DNA fragments around an endless protein core, creating a quasi-continuous superhelix (Figure 6.1A) [25]. This structure is referred to as the hypernucleosome [26]. The hypernucleosome is left-handed like the eukaryotic nucleosome, although it has been reported that archaeal histones can accommodate both left- and right-handed wrapping configurations [20, 29, 30]. It was proposed that hypernucleosomes are not just held together by interactions between adjacent dimers, but also by stacking interactions between the layers of dimers within the hypernucleosome (Figures 6.1B and 6.1C) [26]. The disparate findings from *H. volcanii* and *T. kodakarensis*, as well as the crystal structure of HMfB-DNA complexes, indicate that molecular insights into the role of histones in archaeal chromatin organization are important for understanding archaeal genome compaction.

In this study, we provide direct evidence of hypernucleosome formation and its capability of transitioning in handedness. HMfA and HMfB homodimers exhibit subtle differences in amino acid composition, resulting in different stacking energies in the hypernucleosomes. This results in different stabilities of the hypernucleosome, which may be the key mechanism of transcription regulation in Archaea.

## 6.2 Results

### 6.2.1 Both HMfA and HMfB cooperatively compact DNA

To determine the degree of DNA compaction by HMfA and HMfB, we performed Tethered Particle Motion experiments (TPM) [34]. Here, the Root Mean Squared displacement (RMS) of a bead connected to a surface-attached DNA tether provides a quantitative readout of DNA conformation (Supplementary Figure 6.S1A). We investigated the binding of HMfA and HMfB to a 685 base pairs DNA molecule containing a random, naturally occurring sequence (*see* Materials and Methods). We observed an abrupt and drastic reduction in RMS of the DNA tether over a narrow concentration range during titration of both proteins, which indicates that DNA compaction by both HMfA and HMfB is highly cooperative. Such cooperativity is due to either direct protein-protein interactions or facilitated protein-DNA binding through structural effects of adjacently bound HMfA or HMfB proteins [20, 35] (Figure 6.2).

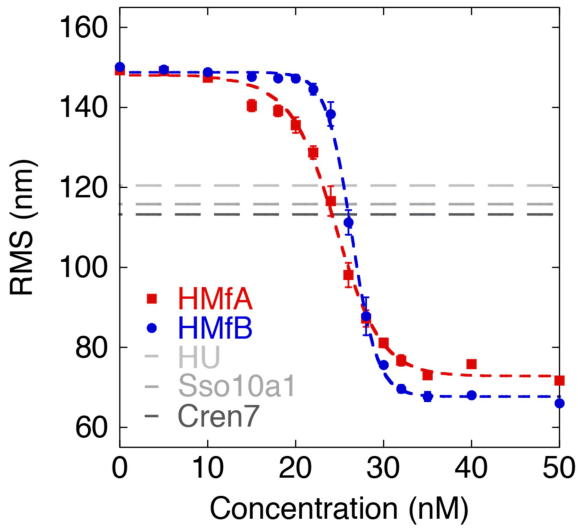



**Figure 6.2**
***Methanothermus fervidus* histone proteins HMfA and HMfB cooperatively compact DNA in Tethered Particle Motion experiments.** Root mean square displacement (RMS) excursion of the bead for HMfA and HMfB on a DNA substrate is shown as a function of protein concentration. Dashed lines are to guide the eye. Error bars indicating the standard error of the data points ($N \geq 100$) are small and mostly hidden behind the data points. In grey, saturation levels of the bacterial NAP HU and archaeal NAPs Sso10a1 and Cren7 – which are known to organize and compact DNA – are indicated, as reported by Driessen *et al.* 2014, Driessen *et al.* 2016 and Driessen *et al.* 2013 [31–33].

As a dimer, HMf proteins bend DNA to an angle that is similar to those induced by other DNA-bending proteins, such as the bacterial NAP HU and the archaeal NAPs Cren7 and Sso10a1 [16, 31–33, 36, 37]. These proteins reduce the RMS of a 685 base pairs DNA substrate from 150 to 110-120 nm at saturating protein concentrations [31, 32]. A much stronger reduction in RMS down to 70 nm was observed for HMfB and $\sim$ 75 nm for HMfA. The extent of this change cannot be explained by DNA bending induced by individual HMf dimers. The high degree of DNA compaction together with the pronounced binding cooperativity suggests adjacent packing of HMfB (and HMfA) proteins that together wrap DNA into multiple turns. By determining the maximum deflection of the bead in TPM, we found an end-to-end distance of $24 \pm 4.7$ nm for the HMfB-DNA complex at saturating protein concentrations for this specific DNA substrate (Supplementary Figure 6.S1). This number agrees with the theoretical end-to-end distance of the hypernucleosomal structure observed in the crystallography studies [25]. Both HMf proteins bind to DNA in a cooperative manner, but HMfB shows a more pronounced cooperativity than HMfA. Differences in the stacking interface or tetramerization interface of

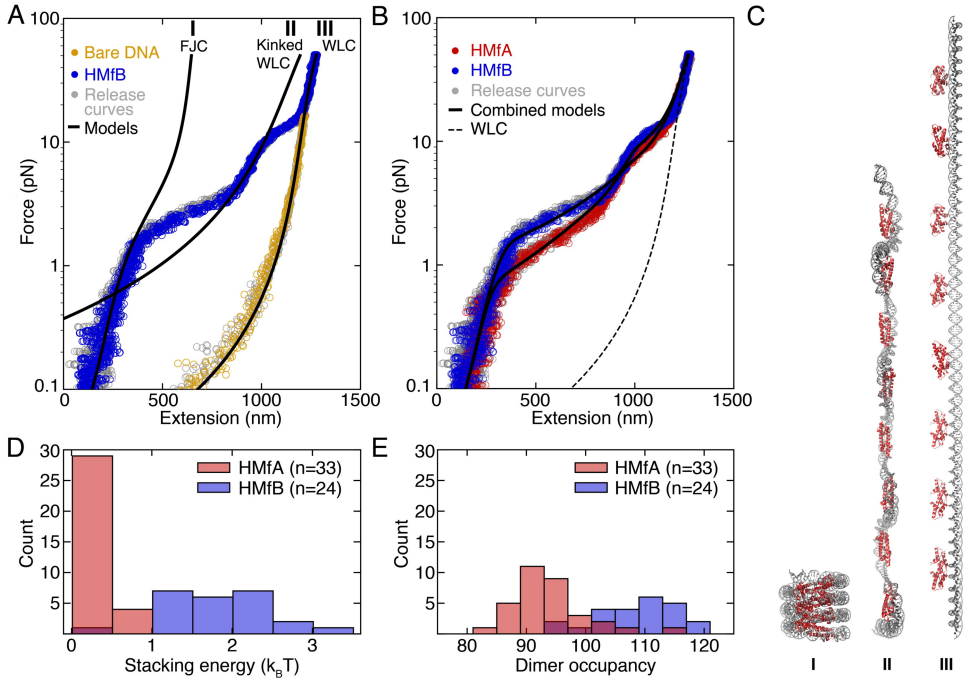



**Figure 6.3**
**Force spectroscopy experiments on the hypernucleosome reveal stronger stacking in HMfB tethers than in HMfA tethers.** A) Force spectroscopy on a HMfB-DNA complex (blue dots) at 100 nM of HMfB reveals three levels of compaction. A bare DNA molecule is shown in yellow dots. We fitted different parts of the curve to a freely-jointed chain (FJC) (I), kinked worm-like chain (WLC) (II), and WLC (III). The release curves overlapped with the stretch curves, which indicates that HMfB-DNA the stretch-release cycle was in equilibrium. B) Comparison between HMfA hypernucleosomes (red dots) and HMfB hypernucleosomes (blue dots). The WLC model for bare DNA was added as a reference (dashed black line). The free energy for dimer stacking and DNA wrapping was extracted by fitting the unfolding model (black lines). C) Structural models of three states of the HMfB-DNA complex corresponding to the model fit in A, illustrating a 7-fold compaction of the hypernucleosome compared to bare DNA. D) Histogram of stacking energies of HMfA and HMfB. E) Histogram of HMfA and HMfB dimers occupancy.

HMfA and HMfB may underlie the difference in DNA compaction (Figures 6.1B and 6.1C).

### 6.2.2 HMfB organizes long DNA tethers into endless hypernucleosomes

The mechanical stability of the hypernucleosome was further investigated by force spectroscopy. HMfB hypernucleosomes were reconstituted *in vitro* on a torsionally unconstrained DNA construct, at 100 nM HMfB, well above the $K_d$. Hence, we expect full coverage of the DNA. Force spectroscopy

revealed characteristic reversible force-extension curves (Figure 6.3A, 6.3B). Binding of HMfB resulted in a large compaction of the DNA. The absence of hysteresis indicates that the manipulation of the HMfB-DNA complex occurs in thermodynamic equilibrium: stretch-release cycles could be repeated multiple times without qualitative changes in the data. The force-extension curve of the HMfB-DNA complex showed three regimes of extension (Figure 6.3A). In regime I, at forces below 1 pN, the force-extension curve featured a small extension and a high stiffness. In regime II, between 2 and 10 pN, the tether is longer and appears to be less stiff than in regime I. In regime III, at forces above 20 pN, the extension of the complex follows that of bare DNA, which is stiffer than the complex in regime II. Most of the force-extension curve of the HMfB-coated DNA could be captured in these three regimes, and we interpret the intermediate parts as transitions between these regimes.

We modeled the two most extended structures, in force regime II and III, with kinked and regular worm-like chains (WLCs). In regime III the complex followed the force-extension curve of bare DNA with a persistence length of 50 nm. For regime II, we fitted the same contour length and a persistence length of approximately 4 nm, suggesting a much more flexible tether. The most compacted structure, regime I, could not be modeled by a WLC and shows a linear extension between 0.1 and 2 pN. We used the linear part of a freely-jointed chain (FJC) to describe this regime of the force-extension curve.

The mechanical properties of the complex suggest three different structures of the tether. Stacking of HMfB dimers into a hypernucleosome would create a relatively stiff structure that can only be stretched to a limited extent without breaking the protein-protein stacking interactions. Based on the crystal structure of the HMfB hypernucleosome [25], we expect an approximate height of 4 nm per dimer for the most condensed structure (Figure 6.3C, I). When the stacking interactions are broken due to excessive force, individual dimers may remain bound to the DNA. Such a structure would have a much larger extension per HMfB dimer than the stacked hypernucleosome. If all DNA-protein contacts remain intact, the DNA follows a highly curved trajectory (Figure 6.3C, II). Such curved DNA results in a decrease of the apparent persistence length [38, 39], which is consistent with the observed force-extension relation in regime II. Increasing the force further breaks protein-DNA interactions to yield an unperturbed DNA trajectory (Figure 6.3C, III), which would have similar mechanical properties as bare DNA. Indeed, at forces larger than 20 pN the force-extension curve overlaps with that of DNA. Thus, the experimental force-extension curves suggest a two-step transition from a hypernucleosome into a fully stretched DNA tether.

For a more quantitative analysis we developed a statistical mechanics model that includes the transitions between the three structures, as described in materials and methods. Release curves fully overlapped with the stretch curves. The absence of hysteresis in all curves indicates that all transitions

are in thermodynamic equilibrium, allowing an equilibrium model to describe the transitions between the states. In analogy with the forced unfolding of eukaryotic chromatin [40], we modeled the entire force-extension curve as a linear combination of the extension of individual dimer-DNA complexes in either of the three states supplemented by a small fraction of bare DNA (*see* Materials and Methods for model details). For every dimer that binds to the DNA tether, the extension changes and the free energy is reduced by a binding and/or stacking energy. For each dimer in state II, we imposed a deflection angle of 10 degrees resulting in a reduction of the persistence length as described by Kulić and Schiessel [38].

The force-induced transitions between all three states of the HMfB-coated DNA tether were adequately captured by the model (Figure 6.3B). The first transition (I to II) at 2 pN indicates a stacking energy of $1.8 \pm 0.7\ k_{\mathrm{B}}T$, which indicates that the stacked hypernucleosome is easily disrupted by force (Figure 6.3D). Such unstacking yielded a lengthening of the tether of $\sim 500$ nm. Unwrapping of the DNA from the dimers at 10 pN takes more energy, $8.7 \pm 0.9\ k_{\mathrm{B}}T$, and increases the extension of the tether by $\sim 200$ nm. These transitions appear as gradual changes in extension, likely because of the non-cooperative behavior of the large number of dimers, the small changes in extension per dimer and the fast kinetics. For multiple tethers, we fitted an average of $108 \pm 7$ dimers per tether. A footprint of 30 base pairs per HMfB dimer would allow 121 dimers to bind this DNA molecule. It therefore appears that the DNA molecule is not fully saturated, implicating defects in stacking of the hypernucleosome (Figure 6.3E). Interestingly, the fitted number of dimers did not change in multiple stretch and release curves of the same hypernucleosome, suggesting that the dimers remain bound to and positioned on the DNA, even at high forces.

TPM experiments showed that HMfB compacts DNA slightly more than HMfA (Figure 6.2). In force spectroscopy, the force-extension curves for HMfA-DNA complexes were similar to those for HMfB-DNA complexes (Figure 6.3B), although the transitions between stacked and unstacked states occurred at lower forces, *i.e.* at 0.5 rather than 2 pN. Correspondingly, we found a smaller stacking energy and dimer occupancy for the HMfA-DNA complex compared to the HMfB-DNA complex, but a similar wrapping energy (Figure 6.3D, 6.3E). The fitted stacking energy, wrapping energy, and dimer occupancy is summarised in Table 6.1. The high similarity in stiffness and extension of the complexes at forces below the unstacking transition (I to II) and the similar wrapping energy after unstacking, highlights the universality of the hypernucleosome structure, whereas the difference in stacking energy and occupancy clearly distinguishes the mechanical properties of the protein-DNA filaments formed by the two homologs. This may be biologically relevant as differences in mechanical stability may modulate the balance between gene compaction and accessibility *in vivo*.

| | stacking energy (mean ± SD) | wrapping energy (mean ± SD) | dimer occupancy (mean ± SD) |
|---|---|---|---|
| **HMfA** | 0.3 ± 0.2 $k_BT$ | 9.0 ± 2.0 $k_BT$ | 94 ± 7 dimers |
| **HMfB** | 1.8 ± 0.7 $k_BT$ | 8.7 ± 0.9 $k_BT$ | 108 ± 7 dimers |

**Table 6.1**
Mean stacking energy, wrapping energy and dimer occupancy of HMfA and HMfB.

### 6.2.3 Stacking interactions mediate hypernucleosome formation

The basis for stacking of HMfA and HMfB homodimers lies in their primary structure. We analyzed both histones in terms of interacting residues in the hypernucleosome structure(Figure 6.1C). Using water refinement with HADDOCK [41] on the HMfB hypernucleosome crystal structure and modeling the HMfA hypernucleosome structure based on that same crystal structure, we were able to identify the possible stacking interactions within the hypernucleosome. For HMfB, we determined interactions between R48-D14, K30-E61 and E34-R65. For HMfA, we found homologous interactions between K31-E62 and E35-K62. Thus HMfB has 1 stacking interaction more than HMfA. The common stacking interactions of both proteins have a similar number and type of hydrogen bonds. When we compare the theoretical total free energy of all interface residues (stacking interfaces and dimer-dimer interfaces including hydrophobic interactions), we surprisingly find that HMfA contacts are energetically more stable than HMfB contacts. Reduced repulsive forces between the HMfA dimers, because of the charge swap at position 37 and the lack of positive charge at position 69, also favor a more stable HMfA complex. However, at HMfB position 64, there is a valine residue that features a hydrophobic interaction with I31, whereas HMfA has an arginine at this position (Figure 6.1C). The large size of this residue disrupts the hydrophobic interaction. This may make the C-terminal α-helix of HMfA more flexible, leading to a less compact and less stable hypernucleosome. Earlier studies support this interpretation, as a V64R mutation in HMfB leads to reduced electrophoretic mobility of the histone-DNA complex and a slightly impaired nucleosome formation [42–44].

### 6.2.4 The hypernucleosome forms a left-handed structure on DNA

Wrapping DNA into a hypernucleosome imposes a distinct chirality. To reveal handedness and torsional stiffness of the hypernucleosome, we used rotational force spectroscopy on a torsionally constrained HMfB-DNA complexes. As opposed to the experiments described in the previous section, the hypernucleosome was unable to release torsional stress by swiveling around the attachment

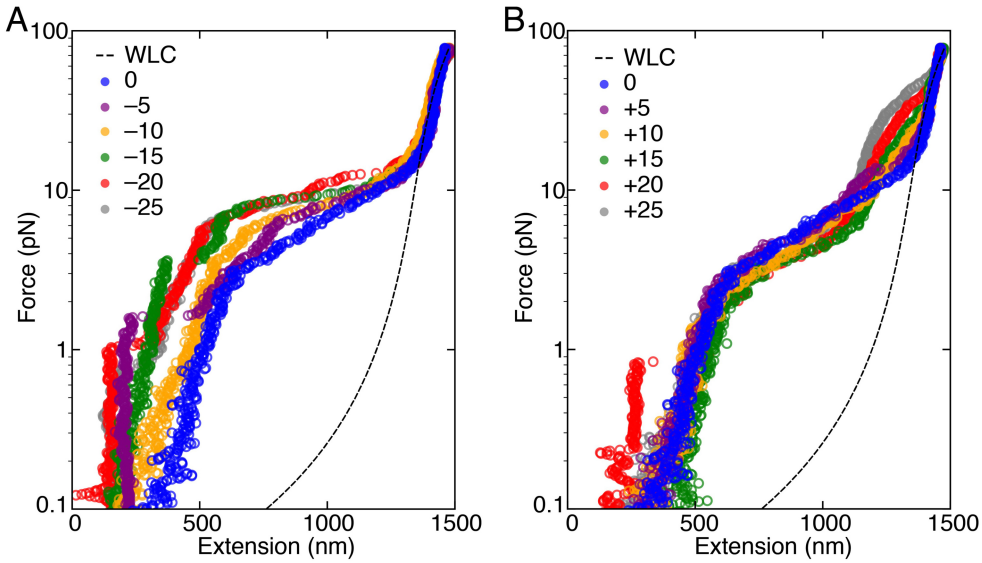



**Figure 6.4**
**Stretching a torsionally constrained HMfB-DNA complex reveals left-handedness of the hypernucleosome, and shows increased affinity of HMfB dimers for positively twisted DNA.** A) Negative twist as a result of 5-25 rotations increases the unstacking force of the hypernucleosome. In addition, excess negative twist causes the HMfB-DNA complex to buckle, resulting in a dramatic decrease of its extension. Both observations point to the left-handedness of the hypernucleosome. The unwrapping transition appears largely unaffected, resulting in a continuous decondensation of the tether around 10 pN. B) Positive twist as a result of 5-25 rotations recovers the unstacking plateau. Strikingly, positive twist increases the rupture force of the unwrapping transition of HMfB dimers. An increasing force is necessary to reverse the bending effect of HMfB dimers on DNA. At very low forces the beads appear to stick to the cover slip for certain curves (-5, -20, and +20 rotations, especially).

point in this experiment. The resulting torque could stabilize the HMfB-DNA hypernucleosome when stretched.

We monitored the extension of the hypernucleosome at a constant force, when applying negative (Figure 6.4A) and positive (Figure 6.4B) twist. The force plateau that represents the unstacking transition shifted up to 10 pN for negatively twisted complexes. As more force is required to unstack the hypernucleosome, the energy for the unstacking transition is increased. Large amounts of negative twist compacted the HMfB-DNA complex by buckling the hypernucleosome at forces below 1 pN. The second transition, from wrapped to unbent DNA, appeared unaffected by negative twist. Positive twist, on the other hand, only marginally extended the range of the unstacking transition to 2-10 pN. Such behavior is characteristic of a left-handed structure, and in agreement with the HMfB-DNA co-crystal structure [25]. Strikingly, the bent

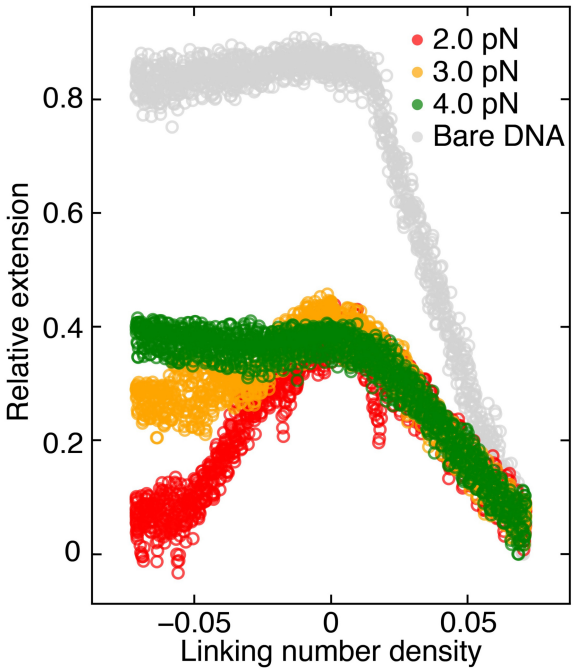



**Figure 6.5**
**Torque stabilizes the hypernucleosome and reveals a right-handed chirality of unstacked HMfB-DNA complexes.** A torsionally constrained HMfB hypernucleosome was compacted by twisting the molecule under constant force. At 2.0 pN, both positive and negative twist reduced the extension, presumably by buckling. At higher forces, negative twist caused the HMfB dimers to unwrap from the DNA, which induced stretching of the HMfB-DNA complex. These measurements suggest that the unstacked HMfB-DNA complex is right-handed. As a reference, we included a twist-extension curve of bare DNA at 1.0 pN. Note that the slope for the HMfB-DNA complex is much smaller than that of bare DNA.

DNA was stabilized by positive twist, but not by negative twist. Thus, it appears that unstacked HMfB dimers can better accommodate positive twist than bare DNA. In accordance, the affinity for HMfB dimers increased when overtwisting the DNA.

### 6.2.5 Unstacked HMfB-DNA complexes impose a right-handed DNA structure

The unstacking transition of a torsionally constrained hypernucleosome features an increased and more gradual force plateau compared to a torsionally unconstrained hypernucleosome (Supplementary Figure 6.S2). As a result, the unstacking transition merges with the transition that we attribute to unwrapping DNA from the HMfB dimers (regime II to III in Figure 6.3A). Since HMfB dimers appeared to remain bound to the DNA, the unwrapping of one dimer

would create additional torsional stress, which impedes unwrapping of the next dimer, resulting in an anti-cooperative transition that manifests itself by an increased force range in which unfolding takes place.

When maintaining a fixed force, and monitoring the extension as a function of twist, we were able to determine the handedness of the unstacked HMfB-DNA complex (Figure 6.5). At $F = 2.0$ pN, both negative and positive twist induced a reduction in extension due to buckling of the hypernucleosome structure. Note that the presence of HMfB-induced bends favored plectonemes over DNA melting, which would occur in bare DNA for negative twist at this force. An asymmetry in twist response in HMfB-decorated tethers occurred at $F = 3.0$ pN. At this force, negative twist helped to unwrap the HMfB dimers, increasing the length of the tether. Similar to bare DNA, positive twist induced supercoils in the tether. This implies that the unstacked HMfB-DNA complex formed a right-handed superhelix and that the HMfB dimer DNA complexes were stabilized by positive torque. The force required to completely unwrap the DNA also depended on the applied twist. In this intermediate force regime, HMfB dimers appeared to have a higher wrapping affinity for positively twisted DNA than for negatively twisted or relaxed DNA. This observation was supported by the release curves (Supplementary Figure 6.S3), which showed that positively twisted complexes more readily restack into hypernucleosomes than relaxed or negatively twisted tethers.

The slope of the curve beyond the buckling transitions, as shown in Figure 6.5, reveals more details of the twisted complex. Bare plectonemic DNA reduces its extension by 40-100 nm per rotation, depending on force, as the size of the loop at the tip of the plectoneme decreases with force [45]. For unstacked HMfB-DNA complexes we measured a reduction of the extension of 16 nm per rotation, much lower than bare DNA, which was independent of force. This is consistent with a fixed HMfB-induced bend at the tip of the plectoneme that defines the tip curvature, and thus the geometry of the plectoneme.

## 6.3 Discussion

The hypernucleosome structure may be a key feature of archaeal chromosome compaction. Here we characterized the structural and mechanical properties of single DNA tethers containing such hypernucleosomes. We found that hypernucleosomes composed of stacked *M. fervidus* histone homodimers HMfA and HMfB compact DNA much stronger than similar DNA-bending proteins. Both HMfA-DNA and HMfB-DNA complexes featured a two-step unfolding mechanism that could be captured quantitatively in the transitions between a hypernucleosome, an array of wrapped dimers that kink the DNA trajectory, and a stretched DNA molecule to which dimers remain bound. HMfB featured a higher stacking energy than HMfA, resulting in a larger occupancy as well as

a higher stability of the hypernucleosome against force.

The stiffness of hypernucleosomes is larger than that of stacked eukaryotic nucleosomes [40]. This may be explained by the different nature of the stacking interactions. Interactions between eukaryotic nucleosomes are known to be mediated by flexible histone tails. Our data suggest that the archaeal hypernucleosome on the other hand is mediated by stacking interactions between the globular histones that form the hypernucleosome. Stacking energies of HMfA and HMfB dimers ($\sim 2\ k_{\mathrm{B}}T$) are much lower than the 17 $k_{\mathrm{B}}T$ that we reported for eukaryotic nucleosome stacking [46]. The unwrapping of DNA from the HMfA and HMfB dimers on the other hand starts at similar forces as the unwrapping of eukaryotic nucleosomes. However, in hypernucleosomes this happened reversibly, whereas nucleosome unwrapping is generally not in equilibrium.

The equilibrium conditions made it possible to capture the entire force-extension curve in a statistical physics model. This worked better for the unstacking transition than for the unwrapping transition. It should be noted though that modeling of the wrapped dimers as kinks in the DNA trajectory might be somewhat naive, as the high occupancy of the DNA yields a dimer-dimer distances that are much smaller than the typical deflection length for which the Kulić and Schiessel model is valid [38]. Moreover, this model assumes a flat one-dimensional kink, while the super-helical structure of the unfolded hypernucleosome may be more complex. A deflection angle (10°), which matched the curves best, is rather small for the 30 base pair footprint of a HMfA or HMfB dimer. Nevertheless, the resulting force dependent reduction of the end-to-end distance matched quantitatively with the force-extension curve. We resolved a clear transition from a flexible (regime II) to a stiffer WLC (regime III), using the same number of dimers as in the unstacking transition at low force, indicating that this unfolding intermediate can be effectively modeled as an array of wrapped dimers.

In the crystal structure, the hypernucleosome compacts DNA in a left-handed manner, which resembles DNA wrapping by eukaryotic histones. In accordance, we found that negative supercoiling stabilized this most compact regime of the HMfB-DNA complex. However, after breaking the stacking by force, a right-handed HMfB-DNA structure was observed. Such a right-handed HMfB-DNA complex has been described previously as a functional form, responding to environmental cues *in vitro* [17, 30]. Depending on conditions, force and torque, a right-handed HMfB-DNA complex may have a biological function in Archaea.

The archaeal genome is also organized by other NAPs. The components of the organizational machinery however differ per phylum, class and order. Histones are found in almost every branch of the archaeal domain. Currently, evidence in support of hypernucleosome formation has only been reported for *M. fervidus* using X-ray crystallography, and for *T. kodakarensis* showing *in*

*vivo* effects on transcription and MNase digestion results [25, 28]. It is likely though that histones from other species also assemble into hypernucleosomes [26]. Initiation of the hypernucleosome is probably stochastic, and initial complex formation at multiple sites may lead to frustration when two expanding hypernucleosomes meet. Therefore, long stretches of hypernucleosome will probably be rare in the presence of other factors that regulate chromatin structure [47]. Ultimately, long straight hypernucleosomes might clash with physical boundaries of the cell. Such large and highly ordered structures should be visible with microscopic techniques, but empirical evidence for such large filaments is currently lacking [28].

*M. fervidus* histones form homo- and heterodimers, and, based on sequence similarities, histones from other species may be able to do so as well. The majority of archaeal chromosomes contain genes coding for two or more different histones. This may provide Archaea with a means to regulate gene expression via changing the length and stability of the hypernucleosome in response to growth phase or environmental cues [17]. For Archaea, transcription regulation may follow a variety of mechanisms [48]. For instance, heterodimers may bind to DNA and interact with other heterodimers via one multimerization site, while other sites remain unbound. Homodimers may form hypernucleosomes, which can be interrupted by other homodimers incapable of certain interactions. Furthermore, sequence specificity in certain histone variants may direct histones to specific genomic targets. The emergence of the N-terminal tail has diverted the eukaryotic histone from the archaeal histone, eventually resulting in alternative abilities to regulate genes than in Archaea [26]. The current methods of biological analysis at the single-molecule level allow for detailed, mechanistic measurements that have the potential to shed light on the evolutionary path of genome compaction and transcription regulation.

## 6.4 Materials and Methods

### 6.4.1 DNA substrate preparation

For the TPM DNA substrate, a random sequence of 50% GC content (`CGGC GCAAATTCGTGACCAGTTGCATCAGCTGCGTGAGCTGTTTATCGCAGCATCGTAACAGGATAGT GAAGAAGACT`) was cloned into pBR322, as described previously [32], resulting in plasmid pRD121. We used PCR to generate and amplify a 685 base pairs linear substrate (sequence in supplementary material) containing the cloned sequence, using digoxygenin- and biotin-labeled oligonucleotides [34].

For magnetic tweezers a 3646 base pairs DNA fragment based on plasmid pFW01 (pBPCYC1(wt)/3 derivative [49], sequence in supplementary material) was digested with BsaI and BseYI (New England Biolabs). The fragment was isolated using agarose gel electrophoresis and purified using a Promega Wizard SV Gel & PCR cleanup kit. It was subsequently labeled with digoxigenin and

biotin at either end by a Klenow reaction.

For the torsionally constrained construct, pUC18-based plasmid pTB01, (containing 12 repeats of the Widom 601 sequence separated by 25 base pairs of linker DNA [50], sequence in supplementary material), was digested with BsaI and BseYI. The 4092 base pairs digestion product was isolated using agarose gel electrophoresis and purified using a Promega Wizard SV Gel & PCR cleanup kit. Subsequently, a 650 base pairs handle containing multiple digoxigenin- or biotin-modified bases was ligated on each end. Details of both protocols are provided elsewhere [51]. The 4092 base pairs substrate, based on plasmid pTB01, qualitatively behaved the same as the 3646 base pairs pFW01-based substrate. We did not observe different features of the hypernucleosome on this nucleosome reconstitution template.

### 6.4.2 Tethered Particle Motion

The tethered particle motion experiments were carried out as described previously [52]. The measurement buffer contained 50 mM Tris-HCl, pH 7 and 75 mM KCl. We used an anisotropic ratio cut-off of 1.3 and a standard deviation cut-off of 8% to select single-tethered beads. The end-to-end distance of the DNA substrate was determined by taking the average of the extremities (2.5% largest values) of the deflection of 25 beads in the $xy$-plane at 50 nM of HMfB and subtracting the radius of the bead (Supplementary Figure 6.S1). The end-to-end distance was obtained by triangular calculation.

### 6.4.3 Force spectroscopy and rotational spectroscopy

HMfA and HMfB were diluted to 100 nM in measurement buffer. 1 pM DNA substrate was introduced in the flow cell and incubated with histones, following the TPM protocol, but using paramagnetic M270 beads (Dynabeads; Thermo Fisher Scientific, USA) instead of polystyrene beads. Details of the force spectroscopy experiments and the multiplexed magnetic tweezers setup have been described elsewhere [51]. In each measurement, the force was increased exponentially from 0.01 to 40 pN in 70 seconds, and decreased at the same speed.

### 6.4.4 Quantitative modeling of stretching the hypernucleosome

To interpret the force-extension behavior of the hypernucleosomes, we developed a statistical mechanics model. HMfA and HMfB dimers bound to the DNA were distributed over three conformations, representing different levels of compaction. In the lowest force regime, HMfA and HMfB homodimers assembled into a hypernucleosome. The extension of the hyper-nucleosome was modeled by a

freely-jointed chain (FJC) [53, 54]:

$$z_{\mathrm{FJC}}(F) = L_{\mathrm{dimer}} \left( \coth\left(\frac{Fb}{k_{\mathrm{B}}T}\right) - \frac{k_{\mathrm{B}}T}{Fb} \right), \tag{6.1}$$

where $z_{\mathrm{FJC}}(F)$ is the extension per dimer, $b$ the Kuhn length that characterizes the tether flexibility, $F$ the force, $k_{\mathrm{B}}$ Boltzmann's constant, $T$ the temperature, and $L_{\mathrm{dimer}}$ the height of a single dimer (when fully wrapped), which was fixed at 4 nm. The FJC was chosen for both its initial linear increase of extension with force, as observed experimentally, and its asymptote to contour length at high force. For low forces ($Fb \ll k_{\mathrm{B}}T$), the FJC describes a Hookean spring with linear extension, and its stiffness $k$ of the stacked fiber is inversely proportional to the effective Kuhn length [53]:

$$k = \frac{3k_{\mathrm{B}}T}{bL_{\mathrm{dimer}}}. \tag{6.2}$$

Note that only the linear part of the FJC, *i.e.* up to an extension up to 2 nm per dimer, mattered, as the dimers unstacked at larger forces. The second force regime is characterized as a beads-on-a-string structure. In this structure, DNA wraps onto non-interacting HMfA or HMfB homodimers resulting in moderate compaction of the fiber. The beads-on-a-string structure was modeled as an extensible worm-like chain (WLC) [40, 55]:

$$z_{\mathrm{WLC}}(F) = L \left( 1 - \frac{1}{2}\sqrt{\frac{k_{\mathrm{B}}T}{FP}} + \frac{F}{S} \right), \tag{6.3}$$

in which $z_{\mathrm{WLC}}(F)$ is the extension of an HMf dimer-DNA complex, L is the contour length of the DNA substrate, S the stretch modulus of DNA and P the persistence length. Since the footprint of an HMfB homodimer is 30 base pairs [56], each transition of a dimer from the hypernucleosome into a beads-on-a-string conformation reduces the contour length of the FJC by $L_{\mathrm{dimer}}$ and increases the contour length of the WLC by 30 base pairs. Moreover, each HMf dimer that bends the DNA, which reduces the persistence length of DNA to an apparent persistence length $P_{\mathrm{app}}$ [38]:

$$P_{\mathrm{app}} = \frac{P}{\left(1 + PN^2 8 \left[1 - \cos\left(\frac{\alpha}{4}\right)\right]/2\right)^2}, \tag{6.4}$$

in which $N$ is the number of dimers in the beads-on-a-string conformation and $\alpha$ is the DNA deflection angle induced by the dimer. In addition, we added a small free energy contribution per kink $g_{\mathrm{kink}}$ that is proportional to the square root of force and did not exceed 0.01 $k_{\mathrm{B}}T$ [57]. The model matched the data best using a deflection angle of 10 degrees, which corresponds roughly to the structure depicted in Figure 6.3C.

At high forces, the HMfA and HMfB homodimers appeared to remain bound, but did not bend the DNA. Each transition therefore reduced the number of kinks in the DNA, resulting in an increase of the apparent persistence length to that of bare DNA, for which we used $P = 50$ nm and $S = 1000$ pN [40, 54, 58].

A large number of fiber states $j$ can be defined, reflecting different distributions of each of the three dimer conformations $i$, representing the hypernucleosome, the beads-on-a-string, or the straight conformation. The total extension $z_j(F)$ of the tether is a linear combination of the extension of each dimer conformation $z_i(F)$ multiplied by the number of dimers $n_i$ in conformation $i$:

$$z_j(F) = \sum\nolimits_i n_i z_i(F). \tag{6.5}$$

The free energy per dimer in the hyper-nucleosome is given by:

$$\begin{aligned} g_{\mathrm{FJC}} &= \int z_{\mathrm{FJC}}(F)\mathrm{d}F + g_{\mathrm{stack}} + g_{\mathrm{wrap}} = \\ L_{\mathrm{dimer}}\frac{k_{\mathrm{B}}T}{b}&\left(\ln\left(\sinh\left(\frac{Fb}{k_{\mathrm{B}}T}\right)\right) - \ln\left(\frac{fB}{k_{\mathrm{B}}T}\right)\right) + g_{\mathrm{stack}} + g_{\mathrm{wrap}}, \end{aligned} \tag{6.6}$$

using a protein-protein interaction energy $g_{\mathrm{stack}}$ and a protein-DNA interaction energy $g_{\mathrm{wrap}}$.

The free energy per dimer in the wrapped and the straight conformation is given by:

$$g_{\mathrm{WLC}} = \int z_{\mathrm{WLC}}(F)\mathrm{d}F = L\left(F - \sqrt{\frac{Fk_{\mathrm{B}}T}{P}} + \frac{F^2}{2S}\right), \tag{6.7}$$

complemented by $g_{\mathrm{wrap}}$ for the wrapped conformation. In addition, the work $W$ done by the bead corresponds to:

$$W = \int F(z_j)\mathrm{d}z. \tag{6.8}$$

The total free energy $G_j(F)$ of a particular tether is the work by the bead minus the sum of the free energy contributions of each dimer with an additional free energy contribution for the number of kinks $N \cdot g_{\mathrm{kink}}$ [57]:

$$G_j(F) = W - \sum\nolimits_i n_i g_i(F) + N \cdot g_{\mathrm{kink}}. \tag{6.9}$$

Finally, the force dependent extension of the HMf-DNA complex was calculated as the Boltzmann-weighted mean extension:

$$\langle z_{\mathrm{tot}}(F)\rangle = \frac{\sum_j z_j(F)\exp\left(-G_j(F)/k_{\mathrm{B}}T\right)}{Z} - z_0 \tag{6.10}$$

with partition function $Z = \sum_j \exp\left(-G_j(F)/k_{\mathrm{B}}T\right)$.

Prior to fitting the entire curve, the $z$-offset $z_0$ was determined by fitting the WLC for bare DNA at forces over 30 pN. Subsequently, the number of dimers on the DNA was determined based on the maximum compaction of state I and state II. The stiffness $k$ of the hypernucleosome was fixed at 1.2 pN/nm, yielding a good agreement with the data. Since the hypernucleosome unstacked at a relatively low force, fitting this value was problematic. Equation 6.10 was fitted to the force-extension curve in the force range between 0.5 and 55 pN to extract the number of dimers, the stacking energy $g_{\mathrm{stack}}$ and wrapping energy $g_{\mathrm{wrap}}$ from each curve.

## 6.5 Acknowledgments

We thank Kathleen Sandman and John Reeve for providing wild type HMfA and HMfB. We thank Helmut Schiessel for fruitful discussions and guidance during the development of the statistical mechanics model. Research in this article is supported by grants from the Netherlands Organization for Scientific Research [VICI 016.160.613] (RTD), [VICI 680.47.616] (JvN), and [VIDI 723.013.010] (HvI), the FOM Foundation for Fundamental Research on Matter program 'Crowd management: The physics of genome processing in complex environments' (RTD) and the Human Frontiers Science program [RGP0014/2014] (RTD).

## 6.6 Supplementary Material

### 6.6.1 Sequence of the TPM substrate

5' TTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGG
AATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT
TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAG
GGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGAC
ATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAAGAATTCCGGCGCAAATTC
GTGACCAGTTGCATCAGCTGCGTGAGCTGTTTATCGCAGCATCGTAACAGGATAGTGAAGAAGA
CTAAGCTTTAATGCGGTAGTTTATCACAGTTAAATTGCTAACGCAGTCAGGCACCGTGTATGAA
ATCTAACAATGCGCTCATCGTCATCCTCGGCACCGTCACCCTGGATGCTGTAGGCATAGGCTTG
GTTATGCCGGTACTGCCGGGCCTCTTGCGGGATATCGTCCATTCCGACAGCATCGCCAGTCACT
ATGGCGTGCTGCTAGCGCTATATGCGTTGATGCAATTTCTATGCGCACCCGTTCTCGGAGCACT
GTCCGACCGCTTTGGCCGCCGCCCAGTCCTGCTCGCTTCGCTACTTGG 3'

### 6.6.2 Sequence of the torsionally free substrate

5' ACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGG
GCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGG
AAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCAT
CGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGA
GTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCA
GAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGT
CATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAG
TGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCA
GAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACC
GCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACT
TTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGG
CGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGG
TTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCG
CGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCT
ATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCT
CTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAA
GCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAG
AGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAA
ATACCGCATCAGGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGG
GCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAA
CGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGCCAAGCTTGCATGCCTGCA
GGTCGACGAGGTGATATTTCCAATTTGGGAAATTTCCCAAATCAGTAATGTAGCCTCTACGGGT
GTCTCTGTCAGCCCCGTGGTCGCCAGCACAGAATGTATCGTACCCCTGAAGGTAGTTTTTTACC
GCCGTGGCACACGATAAAGGTGCACCTTGTGATAATAAGGTGGAAAAATATATATGAAAAAGTG
AAATTGATTGTGGCTGCACTAGGACATCATTATTTCTTACTTGGCTATTTACACGTACTTACGC
TGGCTGTATATCATTTAAGGGGCGGAGGACGAAGAGGACGGACCCGAGATCATCCGGTCCAAGA

AACGGGTCATCCGGTCCTTAGCATTGTCTAGACTATCTAGGGCAGGACGGACATCCACGTGGAA
AGTAGGCATTCCGTTTTCGTCGTCGGGCCCTCCGTAGAAATCCAAGACGTATCTAACTTCCTTG
AAGGTTGGAGGTTGTTGTTCCGCTTTGCGCTCGCCTCGGAGTACAATCCAGTCGTGCCTGTCGA
ATGGTAGTTCTTGGCTAAAATGGGACGGAAACAGTAGGCCGCACAGGTGCATCCAGCGAGCACG
AGGGCTCAATACGCCCGGTTTCCCCATGAATTTCAGCAACTTAGGCTGCACGTGGCTTTCATCT
GTGTGCGGTTTTTCCCATTCGAGCACTTCCTGCCAGCACCCTTCATTTAGAAAGTTGTGGACCT
GCACCATCGACTCCACTGCATCTTCGGCGACTTCGCCGCTACCGCCAATCTTGCCCTTTCTAAC
CATAGCATTGTACATCTGTTGTGGAGAAGGATACTCCCAGAACTCGTTACTGTCTGGACTCTTG
GGGATGCTGGAGATGGTCCGATCAACGGGCAAGTCCATCTTTTGGCCAGGCTGTTTGGATGCTG
CCAACTCCGGCATATTGTTCAGCGGGTTTATTCTATCGTTATCTCCCTGCATAACGGGGCACTC
AGAGGATGGTGGCGACGACGACGACGACTCGTGCATGACTGGGCACCCTGACATGGATGATACT
GCTGCCCCACCAATATCTTTGCCCGTAGTTTTTTGATCTGCCCAAAACCAACCCATTTTTTGTG
AATTCATAGTATCGATACCCTCCTAAGTTTTGTAATCTATAAAGTTAGCAATTTAACTAAGTGT
AAAAACTTAGCTAGCTTACTAAAAAGATAATTACTGAAAAGCTAGCTTGGCCGGATCCGATCGA
CAAAGGAAAAGGGGCCTGTTTACTCACAAGCTTTTTTCAAGTAGGTAATTAAGTCGTTTCTGTC
TTTTTCCTTCTTCAACCCACCAAACGCGTACTTGGTACCAGGAATATATTTCTTTGGGTTAGTC
AAGTACTCTGACATGTTATTTTCGTCCCACAACACGTTTTTCTTGATATTGGCATCTGTGTACG
AATACCCTTCAGCTTGACCAGAGTGTCTGCCAAAGATACCATGCAAGTTTGGACCAACCTTATG
TGGGCCACCCTTTTCCACGGTGTGGCATTGTAGACATCTAGTCTTGAAAAGTGTAGCACCTTTC
TTAGCAGAACCGGCCTTGAATTCAGCCATGGATATATCTCCTTCTTAAAGTTAAACAAAATTAT
TTCTAGATTAGTTAGTTATGGATCCCCGGGTACCGAGCTCGAATTCGTAATCATGGTCATAGCT
GTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAG
TGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCG
CTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGG
CGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGG
CTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATA
ACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT
GCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAG
AGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGC
GCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGT
GGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTG
G 3'

### 6.6.3 Sequence of the torsionally constrained substrate

5' TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCAC
AGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGC
GGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGC
GGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCCATTCGCCATTCAG
GCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAA
GGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTA
AAACGACGGCCAGTGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGAATATCCCGCCCTGGA
GAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTA

CGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGA
TATATACATCCTGTGCATGTAACTAGAGAATATCCCGCCCTGGAGAATCCCGGTGCCGAGGCCG
CTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTT
AACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGTGCATGT
AACTAGAGAATATCCCGCCCTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGC
TCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTC
CCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGTGCATGTAACTAGAGAATATCCCGCCC
TGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCA
CGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGT
CAGATATATACATCCTGTGCATGTAACTAGAGAATATCCCGCCCTGGAGAATCCCGGTGCCGAG
GCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCG
TTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGTGC
ATGTAACTAGAGAATATCCCGCCCTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGA
CAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATT
ACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGTGCATGTAACTAGAGAATATCCC
GCCCTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAA
CGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCAC
GTGTCAGATATATACATCCTGTGCATGTAACTAGAGAATATCCCGCCCTGGAGAATCCCGGTGC
CGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCC
CGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCT
GTGCATGTAACTAGAGAATATCCCGCCCTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCG
TAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGG
GATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGTGCATGTAACTAGAGAATA
TCCCGCCCTGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCT
TAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAG
GCACGTGTCAGATATATACATCCTGTGCATGTAACTAGAGAATATCCCGCCCTGGAGAATCCCG
GTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGT
CCCCCGCGTTTTAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACA
TCCTGTGCATGTAACTAGAGAATATCCCGCCCTGGAGAATCCCGGTGCCGAGGCCGCTCAATTG
GTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTTAACCGCCA
AGGGGATTACTCCCTAGTCTCCAGGCACGTGTCAGATATATACATCCTGTGCATGTAACTAGTT
ATCGCATGCAAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTC
ACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGA
GCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCA
GCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCT
TCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAA
AGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGG
CCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCC
CCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAA
GATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTAC
CGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGG
TATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGC
CCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATC

GCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAG
TTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGC
TGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGG
TAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGAT
CCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGG
TCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATC
AATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCT
ATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTA
CGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACC
GGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCA
ACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAG
TTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGG
TATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGC
AAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTAT
CACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTC
TGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCT
TGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTG
GAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTA
ACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCA
AAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCA
TACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACAT
ATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCA
CCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGC
CCTTTCGTC 3'

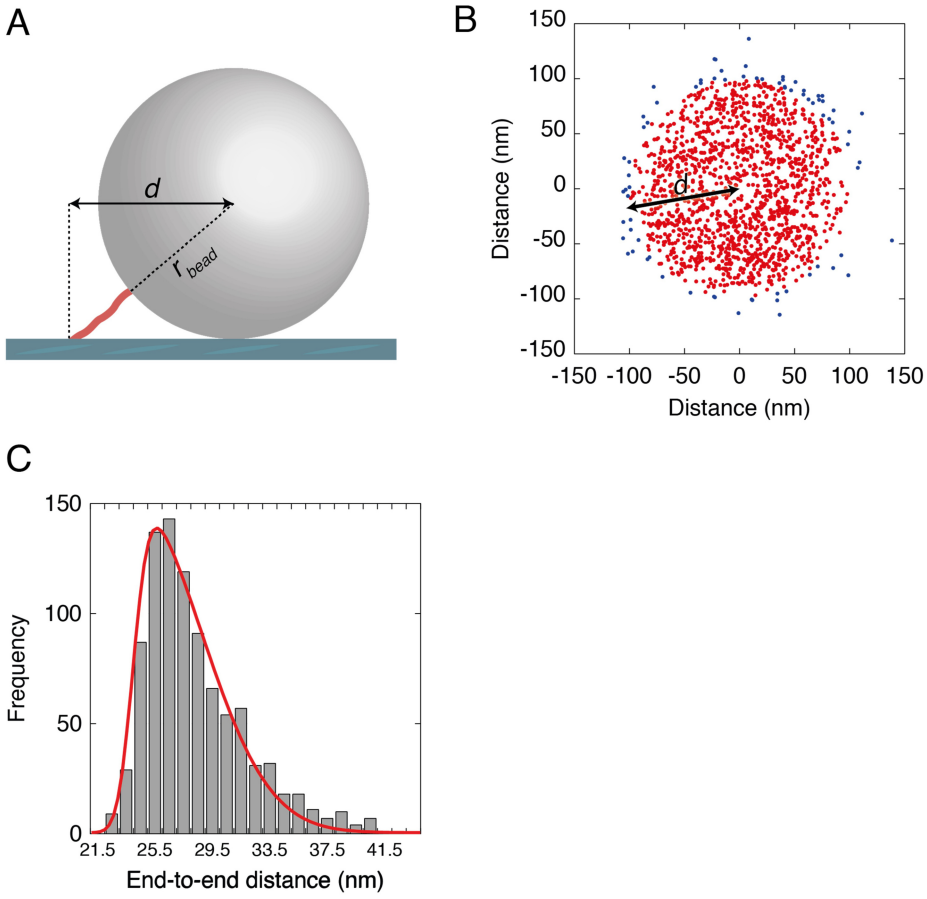



**Figure 6.S1**

**End-to-end distance of the HMfB-DNA complex using TPM.** A) Schematic drawing of the parameters used in the calculation of the end-to-end distance. $d$ is the horizontal distance between the center of the bead and the place where the DNA is attached to the flow cell; $r_{\text{bead}}$ is the radius of the bead. B) Positions of the TPM bead over 60 seconds at 50 nM HMfB (red dots). The 2.5% positions at the extremities of the plot (blue dots) represent the end-to-end distance of the HMfB-DNA complex. C) Skewed Gaussian fit (red line) of the end-to-end distance of the 2.5% most distant positions with respect to the center of 25 beads.

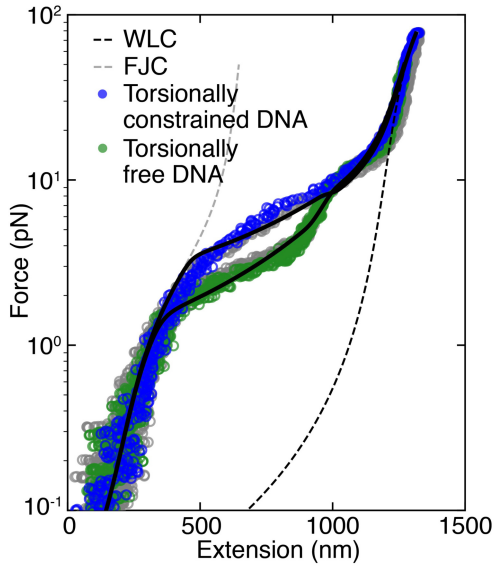



**Figure 6.S2**

**Force spectroscopy on a torsionally constrained and a torsionally free HMfB-DNA complex.** At low force ($F < 2$ pN) both complexes were identical in length, and the curve overlapped with the freely-jointed chain model (FJC). At high force ($F > 20$ pN) the curves overlapped with the worm-like chain model (WLC) of bare DNA. In between, the hypernucleosome unstacked and subsequently unwrapped. Due to the chiral structure of the hypernucleosome, unstacking induces torque, which stabilized the structure. The torsionally constrained hypernucleosome unstacked at $\sim$ 10 pN, compared to $\sim$ 1.5 pN for the torsionally unconstrained hypernucleosome. The unwrapping transition appeared unaffected. Fitting the torsionally constrained hypernucleosome yielded a higher stacking energy ($4.2 \pm 0.1$ $k_BT$ vs $1.6 \pm 0.1$ $k_BT$ per dimer), reflecting an increased energy cost for unstacking the hypernucleosome. The wrapping energy was largely unaffected ($9.9 \pm 0.1$ $k_BT$ vs $10.0 \pm 0.1$ $k_BT$ per dimer). The fit, however deviated somewhat from the data in the transition region. Both molecules appeared to contain 103 HMfB dimers, from which 95 were stacked in a hypernucleosome. The force-extension curves of both molecules converged to a WLC with a contour length of 3646 base pairs at forces over 30 pN.

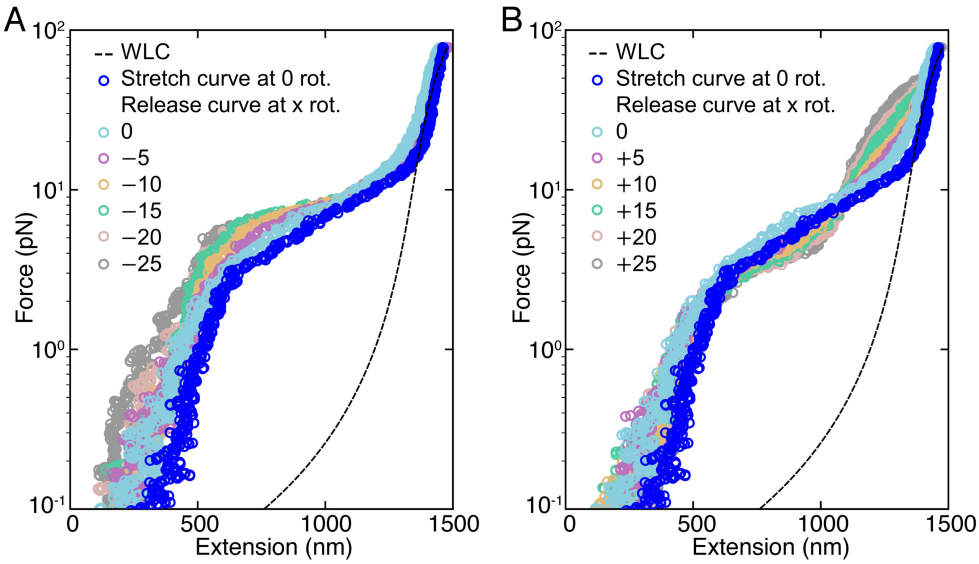



**Figure 6.S3**

**Release curves of torsionally constrained HMfB-DNA complexes.** A) Force spectroscopy on negatively twisted HMfB-DNA complexes. The release curves suggest that overtwisted DNA has a higher affinity to wrap DNA onto the HMfB dimers compared to relaxed DNA, shown by the reduced extension of the release curve. As a reference, the stretch curve for 0 twist was added in bright blue. Upon negative twist, the hysteresis reduced. B) Force spectroscopy experiments on positively twisted HMfB-DNA complexes. At forces between 10 and 40 pN the extension was reduced with increasing twist, which was also observed in the pulling curve.

# Bibliography VI


[1] Tom A Williams et al. "An archaeal origin of eukaryotes supports only two primary domains of life". In: *Nature* 504.7479 (2013), p. 231.

[2] Katarzyna Zaremba-Niedzwiedzka et al. "Asgard archaea illuminate the origin of eukaryotic cellular complexity". In: *Nature* 541.7637 (2017), p. 353.

[3] Laura Eme et al. "Archaea and the origin of eukaryotes". In: *Nature Reviews Microbiology* 15.12 (2017), p. 711.

[4] Cymon J Cox et al. "The archaebacterial origin of eukaryotes". In: *Proceedings of the National Academy of Sciences* 105.51 (2008), pp. 20356–20361.

[5] Gregory D Bowman and Michael G Poirier. "Post-Translational Modi fi cations of Histones That In fl uence Nucleosome Dynamics". In: *Chemical Reviews* 115 (2015), pp. 2274–2295. ISSN: 0009-2665. DOI: 10.1021/cr500350x.

[6] Kathleen Sandman and John N Reeve. "Archaeal chromatin proteins: different structures but common function?" In: *Current opinion in microbiology* 8.6 (2005), pp. 656–661.

[7] Eveline Peeters et al. "The interplay between nucleoid organization and transcription in archaeal genomes". In: *Nature Reviews Microbiology* 13.6 (2015), p. 333.

[8] Martijn S Luijsterburg et al. "The major architects of chromatin: architectural proteins in bacteria, archaea and eukaryotes". In: *Critical reviews in biochemistry and molecular biology* 43.6 (2008), pp. 393–418.

[9] Erin A Becker et al. "Phylogenetically driven sequencing of extremely halophilic archaea reveals strategies for static and dynamic osmo-response". In: *PLoS genetics* 10.11 (2014), e1004784.

[10] Carol J Bult et al. "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii". In: *Science* 273.5278 (1996), pp. 1058–1073.

[11] Trevor J Darcy, Kathleen Sandman, and John N Reeve. "Methanobacterium formicicum, a mesophilic methanogen, contains three HFo histones." In: *Journal of bacteriology* 177.3 (1995), pp. 858–860.

[12] R Tabassum, KM Sandman, and JN Reeve. "HMt, a histone-related protein from Methanobacterium thermoautotrophicum delta H." In: *Journal of bacteriology* 174.24 (1992), pp. 7890–7895.

[13] Bram Henneman and Remus T. Dame. "Archaeal histones: dynamic and versatile genome architects". In: *AIMS Microbiology* 1.1 (2015), pp. 72–81. ISSN: 2471-1888. DOI: 10.3934/microbiol.2015.1.72.

[14] Toshiaki Fukui et al. “Complete genome sequence of the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1 and comparison with Pyrococcus genomes”. In: *Genome research* 15.3 (2005), pp. 352–363.

[15] Iain Anderson et al. “Complete genome sequence of Methanothermus fervidus type strain (V24S T)”. In: *Standards in genomic sciences* 3.3 (2010), p. 315.

[16] K Sandman et al. “HMf, a DNA-binding protein isolated from the hyperthermophilic archaeon Methanothermus fervidus, is most closely related to histones.” In: *Proceedings of the National Academy of Sciences* 87.15 (1990), pp. 5788–5791.

[17] Kathleen Sandman et al. “Growth-phase-dependent synthesis of histones in the archaeon Methanothermus fervidus”. In: *Proceedings of the National Academy of Sciences* 91.26 (1994), pp. 12624–12628.

[18] David R Musgrave, Kathleen M Sandman, and John N Reeve. “DNA binding by the archaeal histone HMf results in positive supercoiling”. In: *Proceedings of the National Academy of Sciences* 88.23 (1991), pp. 10397–10401.

[19] Rowan A Grayling, Kathleen Sandman, and John N Reeve. “DNA stability and DNA binding proteins”. In: *Advances in protein chemistry*. Vol. 48. Elsevier, 1996, pp. 437–467.

[20] Frédéric Marc et al. “Archaeal histone tetramerization determines DNA affinity and the direction of DNA supercoiling”. In: *Journal of Biological Chemistry* 277.34 (2002), pp. 30879–30886.

[21] MT Howard et al. “HMf, a histone-related protein from the hyperthermophilic archaeon Methanothermus fervidus, binds preferentially to DNA containing phased tracts of adenines.” In: *Journal of bacteriology* 174.23 (1992), pp. 7864–7867.

[22] Narasimharao Nalabothula et al. “Archaeal nucleosome positioning in vivo and in vitro is directed by primary sequence motifs”. In: *BMC genomics* 14.1 (2013), p. 391.

[23] Kathryn A Bailey et al. “Archaeal histone selection of nucleosome positioning sequences and the procaryotic origin of histone-dependent genome evolution”. In: *Journal of molecular biology* 303.1 (2000), pp. 25–34.

[24] Kathryn A Bailey et al. “Both DNA and histone fold sequences contribute to archaeal nucleosome stability”. In: *Journal of Biological Chemistry* 277.11 (2002), pp. 9293–9301.

[25] Francesca Mattiroli et al. “Structure of histone-based chromatin in Archaea”. In: *Science* 357.6351 (2017), pp. 609–612. ISSN: 0036-8075. DOI: 10.1126/science.aaj1849.

[26] Bram Henneman et al. “Structure and function of archaeal histones”. In: *PLoS genetics* 14.9 (2018), e1007582.

[27] Ron Ammar et al. “Chromatin is an ancient innovation conserved between Archaea and Eukarya”. In: *elife* 1 (2012), e00078.

[28] Hugo Maruyama et al. “An alternative beads-on-a-string chromatin architecture in Thermococcus kodakarensis”. In: *EMBO reports* 14.8 (2013), pp. 711–717.

[29] Artem K Efremov et al. “Transcriptional repressor TrmBL2 from Thermococcus kodakarensis forms filamentous nucleoprotein structures and competes with histones for DNA binding in a salt-and DNA supercoiling-dependent manner”. In: *Journal of Biological Chemistry* 290.25 (2015), pp. 15770–15784.

[30] David Musgrave, Patrick Forterre, and Alexei Slesarev. “Negative constrained DNA supercoiling in archaeal nucleosomes”. In: *Molecular microbiology* 35.2 (2000), pp. 341–349.

[31] Rosalie PC Driessen et al. “Effect of temperature on the intrinsic flexibility of DNA and its interaction with architectural proteins”. In: *Biochemistry* 53.41 (2014), pp. 6430–6438.

[32] Rosalie PC Driessen et al. “Diverse architectural properties of Sso10a proteins: Evidence for a role in chromatin compaction and organization”. In: *Scientific reports* 6 (2016), p. 29422.

[33] Rosalie PC Driessen et al. “Crenarchaeal chromatin proteins Cren7 and Sul7 compact DNA by inducing rigid bends”. In: *Nucleic acids research* 41.1 (2012), pp. 196–205.

[34] Ramon A van der Valk, Niels Laurens, and Remus T Dame. “Tethered particle motion analysis of the DNA binding properties of architectural proteins”. In: *The Bacterial Nucleoid*. Springer, 2017, pp. 127–143.

[35] Miroslav Tomschik et al. “The archaeal histone-fold protein HMf organizes DNA into bona fide chromatin fibers”. In: *Structure* 9.12 (2001), pp. 1201–1211.

[36] Yingang Feng, Hongwei Yao, and Jinfeng Wang. “Crystal structure of the crenarchaeal conserved chromatin protein Cren7 and double-stranded DNA complex”. In: *Protein Science* 19.6 (2010), pp. 1253–1257.

[37] Liqing Chen et al. “The hyperthermophile protein Sso10a is a dimer of winged helix DNA-binding domains linked by an antiparallel coiled coil rod”. In: *Journal of molecular biology* 341.1 (2004), pp. 73–91.

[38] 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].

[39] J. van Noort et al. “Dual architectural roles of HU: Formation of flexible hinges and rigid filaments”. In: *Proceedings of the National Academy of Sciences* 101.18 (2004), pp. 6969–6974. ISSN: 0027-8424. DOI: 10.1073/pnas.0308230101.

[40] 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.

[41] GCP Van Zundert et al. “The HADDOCK2. 2 web server: user-friendly integrative modeling of biomolecular complexes”. In: *Journal of molecular biology* 428.4 (2016), pp. 720–725.

[42] Klaas Decanniere et al. “Crystal structures of recombinant histones HMfA and HMfB from the hyperthermophilic archaeon Methanothermus fervidus”. In: *Journal of molecular biology* 303.1 (2000), pp. 35–47.

[43] Divya J Soares, Kathleen Sandman, and John N Reeve. "Mutational analysis of archaeal histone-DNA interactions". In: *Journal of molecular biology* 297.1 (2000), pp. 39–47.

[44] Kathleen Sandman and John N Reeve. "Chromosome packaging by archaeal histones". In: *Adv. Appl. Microbiol* 50 (2001), pp. 75–99.

[45] Jan Lipfert, Sven Klijnhout, and Nynke H Dekker. "Torsional sensing of small-molecule binding using magnetic tweezers". In: *Nucleic acids research* 38.20 (2010), pp. 7122–7132.

[46] 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.

[47] Travis J Sanders et al. "TFS and Spt4/5 accelerate transcription through archaeal histone-based chromatin". In: *Molecular microbiology* 111.3 (2019), pp. 784–797.

[48] Hiromi Nishida and Taku Oshima. "Archaeal histone distribution is associated with archaeal genome base composition". In: *The Journal of general and applied microbiology* 63.1 (2017), pp. 28–35.

[49] W Brent R Pollock et al. "Bacterial expression of a mitochondrial cytochrome c. Trimethylation of Lys72 in yeast iso-1-cytochrome c and the alkaline conformational transition". In: *Biochemistry* 37.17 (1998), pp. 6124–6131.

[50] P T T Lowary and J Widom. "New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning." In: *Journal of molecular biology* 276.1 (1998), pp. 19–42. ISSN: 0022-2836. DOI: 10.1006/jmbi.1997.1494. eprint: NIHMS150003.

[51] Thomas B Brouwer et al. "Unraveling DNA Organization with Single-Molecule Force Spectroscopy Using Magnetic Tweezers". In: *Bacterial Chromatin*. Springer, 2018, pp. 317–349.

[52] Bram Henneman et al. "Quantitation of DNA-binding affinity using Tethered Particle Motion". In: *Bacterial Chromatin*. Springer, 2018, pp. 257–275.

[53] TR Strick et al. "Physical approaches to the study of DNA". In: *Journal of statistical physics* 93.3-4 (1998), pp. 647–672.

[54] Terence Strick et al. "Twisting and stretching single DNA molecules". In: *Progress in biophysics and molecular biology* 74.1-2 (2000), pp. 115–140.

[55] John F. Marko and Eric D. Siggia. "Stretching DNA". In: *Macromolecules* 28.26 (Dec. 1995), pp. 8759–8770. ISSN: 15205835. DOI: 10.1021/ma00130a008.

[56] Suzette L Pereira et al. "Archaeal nucleosomes". In: *Proceedings of the National Academy of Sciences* 94.23 (1997), pp. 12633–12637.

[57] IM Kulić and Helmut Schiessel. "DNA spools under tension". In: *Physical Review Letters* 92.22 (2004), p. 228101.

[58] Carlos Bustamante, Zev Bryant, and Steven B. Smith. "Ten years of tension: single-molecule DNA mechanics". In: *Nature* 421.6921 (2003), pp. 423–427. ISSN: 0028-0836. DOI: 10.1038/nature01405.

# 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.

# Samenvatting

De genetische informatie van alles wat leeft ligt verborgen in het DNA. Elke menselijke cel bevat bijvoorbeeld ongeveer 2 meter DNA wanneer je het volledig uit zou strekken. Op een cellulaire schaal is dit een enorme lengte, welke duizenden keren moet worden gecondenseerd om fysiek in de celkern te passen. Tegelijkertijd moet de genetische informatie in het DNA selectief toegankelijk blijven, om in cellulaire ontwikkeling of onderhoud van de cel te voorzien. Cellen manipuleren de mate van DNA-condensatie door epigenetische factoren, welke grotendeels bepalen welke genen actief worden uitgelezen en welke transcriptioneel stil blijven.

Tientallen jaren van onderzoek naar dit systeem hebben de wetenschappelijke wereld geen consensus gebracht over hoe de cel de variërende mate van DNA-compressie kan realiseren *in vivo*. Eukaryoten, zoals mensen, comprimeren het DNA door het op te vouwen rond histon-octameren, een complex van eiwitten in de vorm van klosjes. Samen met het opgevouwen DNA wordt deze structuur een nucleosoom genoemd. Meerdere nucleosomen komen samen om hogere-orde structuren te vormen die uiteindelijk een compacte chromatinevezel vormen. Verschillende structurele modellen voor deze vezels zijn beschreven in de literatuur, waarvan de 1-start solenoïde vezel en de 2-start zigzag vezel verreweg het populairst zijn. Het wetenschappelijk debat over de aanwezigheid van dergelijke georganiseerde structuren *in vivo* blijft echter voortduren, omdat ondubbelzinnige experimentele resultaten helaas uitblijven. In onze groep hebben we eerder aangetoond dat de lengte van het linker DNA, het DNA wat zich tussen aangrenzende nucleosomen bevindt, de mechanische eigenschappen van de chromatinevezels bepaalt. Dit impliceert het bestaan van meer dan één hogere-orde chromatinestructuur, hoewel het volledig beeld deels ontbreekt. Meer inzicht in de structuur en het ontstaan van de chromatinevezel is wenselijk, omdat het disfunctioneren van compressiemechanismen resulteert in een verscheidenheid aan medische aandoeningen. Enkele medische aandoeningen die in verband zijn gebracht met afwijkingen in de DNA-(re)organisatie zijn bijvoorbeeld diabetes type 1, het fragiele X-syndroom of verschillende vormen van kanker.

Het biofysisch onderzoek naar structuur en dynamica van verschillende eiwitten zoals chromatine is in een stroomversnelling terecht gekomen door de on-

twikkeling van krachtspectroscopische technieken die aan individuele moleculen kunnen meten. Moderne microscopen kunnen deze moleculen manipuleren met krachten op de pico-Newton-schaal, waarbij de onderzoeker kan meekijken en de meting kan aansturen zonder vertraging. Magnetische pincetten, bijvoorbeeld, kunnen dit soort krachten op een enkele chromatinevezel uitoefenen. Daarnaast heeft deze microscoop het unieke vermogen om de vezel om zijn as te draaien zodat torsie in de vezel opbouwt, om zo nog meer te weten te komen over de mechanische eigenschappen van de vezel. Torsie speelt een belangrijke rol in DNA-compressie, wanneer bijvoorbeeld genoom-organiserende eiwitten het DNA opwinden of loslaten, buigen, of lussen vormen.

In Hoofdstuk 2 introduceren we een op maat gemaakte microscoop voor onze experimenten: een magnetisch pincet dat in staat is om aan honderden moleculen tegelijkertijd te meten. We beschrijven verder hoe we chromatinevezels bouwen en hoe we deze bevestigen aan een magnetisch bolletje. Dit bolletje van polystyreen, ter grootte van een micrometer, is gevuld met ijzer nano-deeltjes. Vervolgens kunnen we het DNA blootstellen aan kracht, in zogenaamde strek-experimenten, door middel van manipulatie van een magnetisch veld. Ten slotte lichten we toe hoe mechanische parameters van nucleo-eiwitfilamenten kunnen worden geëxtraheerd door de experimentele data te fitten met een statistisch mechanisch model. Dergelijke experimenten resulteren in een gedetailleerde beschrijving van de mechanische eigenschappen van chromatinevezels, welke inzicht kunnen verschaffen over de hogere-orde vezelstructuur.

Veel krachtspectroscopietechnieken die individuele moleculen meten maken gebruik van meettechnieken die de driedimensionale positie van objecten ter grootte van een micrometer met nanometer precisie kunnen bepalen. Deze zogenaamde tracking-methoden meten bijvoorbeeld de positie van een (magnetisch) bolletje verbonden met het (bio)molecuul waarin de onderzoeker geïnteresseerd is. Door de positie van het bolletje als een functie van de kracht te meten kan men de mechanische eigenschappen van het molecuul afleiden. De driedimensionale positie van deze magnetische bolletjes kan worden bepaald door analyse van het diffractie-patroon, wat gevormd wordt door interferentie van de lichtbron van de microscoop met licht wat kaatst van de het bolletje. Bestaande tracking-methoden zijn vaak bewerkelijk, daardoor traag, en daarom niet optimaal voor microscopen die honderden metingen parallel uitvoeren. In Hoofdstuk 3 introduceren we 3D phasor tracking, een snel en robuust tracking algoritme met nanometrische precisie en een bereik van minimaal 10 µm in de $z$-richting. Het algoritme is gebaseerd op een 2D-kruiscorrelatie tussen het diffractiepatroon en enkele computer-gegenereerde referentiebeelden waarbij gebruik wordt gemaakt van zogenaamde snelle fouriertransformaties. Ons algoritme verwerkt tot 10,000 diffractiepatronen per seconde. We hebben deze methode geïmplementeerd in onze magnetische pincetopstelling en waren in staat om de 3D-positie van meer dan 100 bolletjes zonder vertraging te meten met behulp van een generieke CPU. 3D phasor tracking kan biofysische meettechnieken verbeteren waar

tracking met een hoge doorvoer vereist is. De ontwikkeling van deze efficiënte meetmethode was essentieel voor het werk dat verder aan bod komt in dit proefschrift.

In Hoofdstuk 4 bestuderen we de mechanische eigenschappen van chromatinevezels met Monte Carlo simulaties. Om dit te verwezenlijken bouwen we een bestaand rigide basenpaar DNA-model verder uit. We implementeren het (af-)wikkelen van nucleosomaal DNA en (ont-)stapelen van nucleosomen in 1-start en 2-start chromatinevezels. Ons model stelt ons in staat om de verschillende stadia van ontvouwing van chromatinefibers onder invloed van kracht te reproduceren. De resultaten zijn grotendeels consistent met een 2-startvouwing voor chromatinevezels met kort linker DNA en een 1-startvouwing voor chromatinevezels met langere linkers. Omdat de simulaties de driedimensionale structuur van de chromatinevezels met basepaarnauwkeurigheid reproduceren op basis van realistische fysische eigenschappen van het DNA, ondersteunen ze de interpretatie van krachtsspectroscopiedata. Hierdoor kunnen we de mechanische eigenschappen die we uit krachtspectroscopieexperimenten halen koppelen aan gedetailleerde modellen van de structuur van chromatine.

We maken gebruik van het magnetische pincet om in Hoofdstuk 5 diepgaand onderzoek te doen naar de ontvouwing van chromatinevezels. In dat hoofdstuk analyseren we meer dan 1000 metingen om het effect van het systematisch verlengen van het linker DNA op de mechanische eigenschappen van chromatinevezels in kaart te brengen. Het verlengen van het linker DNA doen we in stappen van één basepaar. We verrijken de experimentele data met rigide-basenpaar Monte Carlo simulaties. Zowel de simulaties als de experimentele data onthullen een periodieke modulatie van de vouwenergie met linkerlengte, gedreven door de beperkte flexibiliteit van het linker DNA. De meest stabiele stapelinteracties tussen de nucleosomen worden waargenomen voor linkerlengten in veelvouden van 10 basenparen, wat overeenkomt met de periodiciteit van de B-DNA-helix. We analyseren nucleosoomstapeling in zowel 1- als 2-starttopologieën en laten zien dat de lengte van het linker DNA bepaalt hoe de chromatinevezel vouwt. Korte linkers vormen een 2-startvezel, terwijl lange linkers leiden tot een 1-startconformatie. Hiernaast laten we in Hoofstuk 5 ook zien dat de sequentie van het linker DNA ook de nucleosoomstapeling moduleert. Dat de balans tussen de verschillende vezelstructuren afhangt van zowel de linkerlengte als de sequentie van het linker DNA, impliceert een belangrijke rol voor beide factoren in de epigenetische regulatie in de cel.

Waar de eerste hoofdstukken in dit proefschrift de vouwing van eukaryoot chromatine beschrijven, en hoe de fysieke organisatie van het genoom een sleutelrol speelt in genregulatie, is het laatste deel van dit proefschrift gewijd aan de genomische organisatie in archaea, onze evolutionaire voorganger. In recent werk, waarbij gebruik werd gemaakt van röntgen-kristallografie, is de structuur van het *Methanothermus fervidus* histon HMfB gebonden aan DNA-fragmenten van 80 basenparen voor het eerst in beeld gebracht. Deze publicatie laat zien

dat DNA zich 'eindeloos' rond een stapel van histon-eiwitten wikkelt, en zo een zogenaamd hypernucleosoom kan vormen. Het blijft echter onduidelijk of en hoe deze hypernucleosoomstructuur zich op een langer DNA-substraat assembleert, en welke interacties verantwoordelijk zijn voor de stabiliteit van de structuur. Elektrostatische interacties kunnen de superstructuur moduleren, wat duidt op een epigenetisch mechanisme. In Hoofdstuk 6 beschrijven we verscheidene experimenten die de vouwing van individuele DNA-moleculen onder invloed van *M. fervidus* histonen HMfA en HMfB ophelderen. De experimentele procedure, geïntroduceerd in Hoofdstuk 2 voor de studie naar eukaryoot chromatine, is met minimale aanpassingen ook toepasbaar op archaeale chromatinevezels. Het statistisch mechanisch model moet echter aangepast worden om de experimentele data te fitten. We hebben kunnen afleiden dat hypernucleosoom-assemblage het resultaat is van coöperatieve binding van archaeale histonen aan het DNA, gedreven door interacties tussen naburige histondimeren. Bovendien toont rotatiekrachtspectroscopie aan dat een HMfB-DNA-complex een linkshandige chiraliteit heeft, maar dat opbouw van torsie kan resulteren in de vorming van een rechtshandige conformatie. Samenvattend hangt de structuur van het hypernucleosoom dus af van stapelinteracties, torsie en kracht. Dit duidt op een belangrijke rol voor archaeale histonen als transcriptieregulator, die afhankelijk van omgevingsfactoren het DNA kunnen opvouwen of toegankelijk maken. Het bestuderen van deze archaïsche cellulaire processen kan inzicht verschaffen in de meer gecompliceerde mechanismen die aan de basis liggen van eukaryote DNA-vouwing.

Het werk in dit proefschrift draagt bij aan de fundamentele kennis van epigenetische mechanismen rond de structuur van chromatine. De krachtspectroscopieexperimenten aan chromatinevezels in Hoofdstuk 5 leveren overtuigend bewijs dat er meer dan één hogere-orde chromatinestructuur gevormd wordt *in vitro*. De verschillende topologieën van de chromatinevezels onder invloed van de lengte van het linker DNA zijn in zowel experimentele als gesimuleerde data gevonden. Dit zou een relatief ongeordende structuur van chromatine *in vivo* kunnen verklaren, waar de lengte en de sequentie van de linker varieert. In de toekomst kunnen deze experimenten worden uitgebreid om in kaart te brengen waar de omschakeling tussen de 2-start- en de 1-start-vezel plaatsvindt en welke factoren van invloed zijn op deze balans. Er is nog steeds geen experimenteel bewijs voor reguliere structuren *in vivo*. De lengtes van het linker DNA in de cel zijn gekwantiseerd, en omdat lange linkers relatief veel voor komen (Widom - PNAS, 1992), voorspellen we dat kleine stukjes van deze 1-start chromatinevezels op korte termijn zullen worden geïdentificeerd. Belangrijk is dat de modulatie van de mechanische vezeleigenschappen met de linkerlengte, zoals hier beschreven, het bestaan van 1-startstructuren verder onderbouwt.

Een opmerkelijke vinding beschreven in Hoofdstuk 5 is dat nucleosoomstapeling in 1-start vezels niet wordt gemedieerd door de histon H4-staart. Dit is ogenschijnlijk in tegenspraak met eerdere experimenten van onze groep

(Kaczmarczyk *et al.* - JBC, 2017) die aantoonden dat chromatinevezels met wederzijds verbonden H4-staarten de mechanische eigenschappen van deze vezels insloten. Dit kan duiden op een grotere verscheidenheid aan hogere-orde structuren van chromatinevezels dan aanvankelijk gedacht. We ontdekten ook dat de sequentie van het linker DNA de stabiliteit van de chromatinevezel in verrassend grote mate beïnvloedt. Zowel histon-staart gemedieerde nucleosoomstapeling als linker DNA-sequentie effecten moeten verder worden geëvalueerd door middel van simulatie en experiment om dit verder te verklaren. Ons Monte Carlo-model kan met enkele aanpassingen worden gebruikt om de verschillende stapelmechanismen binnen chromatinevezels te vergelijken en kan verder worden uitgebreid met sequentie-effecten.

De volgende fase van dit onderzoek is het karakteriseren van chromatine met dezelfde samenstelling en nucleosomale bezettingen als in de cel. Onze geparallelliseerde tracking-methode zal het onderzoek naar deze hoogst heterogene monsters vergemakkelijken. Dergelijke experimenten kunnen het mechanisme van DNA-organisatie *in vivo* verder ophelderen. In onze groep hebben we de eerste stappen in deze richting reeds gezet. We hebben nucleosoomstapelingseigenschappen van natief chromatine, geëxtraheerd uit gist, onderzocht. We maakten hier gebruik van een door onze groep ontwikkelde zuiveringstechniek om chromosomale loci in oorspronkelijke staat te isoleren (Hermans *et al.* – Scientific Reports, 2017). Dit werk opent een nieuw onderzoeksgebied naar fysisch begrip van epigenetische regulatie door de cel. Tot slot hopen we dat het onderzoek beschreven in dit proefschrift bijdraagt aan een beter begrip van epigenetische aandoeningen, wellicht buiten de biofysische discipline alleen, wat vervolgens de hoeksteen kan vormen waarop verbeterde behandelingen van dergelijke aandoeningen worden gefundeerd.

# List of Publications

1. **Brouwer T. B.**, Pham C. Kaczmarczyk A., de Voogd W. J., Botto M., Vizjak P., Müller-Planitz F., and van Noort S. J. T.: A Critical Role for Linker DNA in Higher-Order Folding of Chromatin fibers, *submitted*.

2. Henneman B.†, **Brouwer T. B.**†, van Emmerik C., van der Valk R. A., Timmer M., Kirolos N., van Ingen H., van Noort S. J. T., Dame R. T.: Mechanical and Structural Properties of Archaeal Hypernucleosomes, *submitted*.

3. **Brouwer T. B.**, Hermans N., van Noort S. J. T.: Multiplexed Nanometric 3D Tracking of Microbeads Using an FFT-Phasor Algorithm, Biophysical Journal, 118.9, 2020: 2245-2257..

4. 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.

5. **Brouwer T. B.**, Kaczmarczyk A., Pham C., van Noort S. J. T.: Unravelling DNA organization with single-molecule Force Spectroscopy using Magnetic Tweezers, in Bacterial Chromatin, Humana Press, New York, NY, 2018: 317-349.

6. Kaczmarczyk A., **Brouwer T. B.**, Pham C., Dekker N. H., van Noort S. J. T.: Probing Chromatin Structure with Magnetic Tweezers, in Nanoscale Imaging, Humana Press, New York, NY, 2018: 297-323.

7. Kaczmarczyk A., Allahverdi A., **Brouwer, T. B.**, Nordenskiöld L., Dekker N. H., van Noort, S. J. T.: Single-molecule force spectroscopy on histone H4 tail-cross-linked chromatin reveals fiber folding, Journal of Biological Chemistry, 292(42), 2017: 17506-17513.

† These authors contributed equally.

8. Hermans N., Huisman J. J., **Brouwer T. B.**, Schächner C., van Heusden G. P. H., Griesenbeck J., van Noort, S. J. T.: Toehold-enhanced LNA probes for selective pull down and single-molecule analysis of native chromatin, Scientific reports, 7(1), 2017: 16721.

# Curriculum Vitae

I was born in Amsterdam, on the 25$^{\text{th}}$ of Augustus in 1987. My family moved from Amsterdam to Bergen (N-H) when I was very young. I went to the Petrus Canisius College, first in Bergen and then in Alkmaar. In high-school, I discovered that I had a passion for Science. I finished high-school in 2005 and graduated both in biology and physics.

In 2005 I started my Bachelor's degree in *Medical Natural Sciences* at the Vrije Universiteit in Amsterdam. After a brief period as a medical student (2006-2007), I realized I actually wanted to be a scientist and rejoined the *Medical Natural Sciences* program. I completed two internships: at the Department of Condensed Matter Physics of the Vrije Universiteit, I characterized a multimode fiber-top refractometer under the supervision of prof. dr. Davide Iannuzzi. During another project, I optimized the experimental protocol for the irradiation of living mice at the Department of Radiation Oncology of the Vrije Universiteit Medical Center under the supervision of dr. Stan Heukelom. I completed my Bachelor's degree in 2009.

I took a year off, traveled through Central America, and resumed my studies in 2010: a Master's degree in *Medical Natural Sciences*, with a specialization in the *Physics of Life*. Throughout my studies, I volunteered in several NGOs, to improve human rights (with Amnesty International) and fight climate change (with Greenpeace). Under the supervision of prof. dr. ir. Gijs Wuite and prof. dr. ir. Erwin Peterman I studied the speed and dynamics of molecular motors with fluorescence experiments, using a confocal microscope that I constructed myself. After completing my master thesis I traveled to Canada for 7 months to work as a visiting graduate student researcher in the lab of prof. dr. Nancy Force in Vancouver. I studied the extracellular matrix protein collagen II and the mechanics and dynamics of short DNA molecules ($\sim$ 1000 base pairs). The work experience as well as the personal interactions with the extremely motivated and inspiring scientists in both labs convinced me that I wanted to pursue a career in science, at least for a while. I received my Master's degree on the 30$^{\text{th}}$ of August in 2013.

In the PhD position in the group of prof. dr. ir. John van Noort at Leiden University, which I joined in 2014, I found exactly what I was looking for. A

chance to combine state-of-the-art technological development with biophysical experimentation. To study life itself, at the molecular level, is a unique experience that continues to be mind-boggling when one tries to imagine it.

As a PhD candidate, I supervised two bachelor students: Elger Vlieg and Willem Jan de Voogd, and one master student: Margherita Botto. Furthermore I assisted in teaching the bachelor course *LabVIEW*, which instructed students on the basics of visual programming. Several times I presented my research at (international) conferences, which I saw as opportunities to connect with other researchers in the field.

I instigated two collaborations between research groups. The first project was a collaboration between the group of prof. dr. Marileen Dogterom at the Technical University Delft and our own, on which I worked with dr. Stef van der Meulen. We designed a method to attach mobile biochemical linkers to magnetic beads, and successfully performed several proof-of-principle experiments. Unfortunately, this project is not included in this thesis. The second project was a collaboration between the group of dr. Remus Dame and the group of dr. Hugo van Ingen. I worked together with dr. Ramon van der Valk, dr. Bram Henneman, and dr. Clara Emmerik to investigate the structure of archaeal chromatin fibers. Our combined efforts are described in Chapter 6) of this thesis.

In 2019 I seized the opportunity to travel again. Together with my girlfriend Sandra I explored Latin America for 11 months until we were forced to return home because of the COVID-19 pandemic. In 2020 I started a position as a data scientist for NKC, Europe's largest association (and insurer) for campervans.