DISSERTATION

submitted to the

Combined Faculties for the
Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg
Germany

for the degree of
Doctor of Natural Sciences

put forward by

Dipl.-Phys. Yang Zhang

born in: Shanxi

Date of oral examination: 07.02.2013

Chromosomes in Interphase and Mitosis

Referees: Prof. Dr. Dieter W. Heermann
Prof. Dr. Michael Hausmann

Abstract

The three-dimensional organization of the chromatin fiber is driven by entropy. Therefore, the folding of the chromatin fiber is essentially a problem of statistical physics. In the present thesis, two questions in the context of chromatin folding which are still not fully understood are investigated: on the one hand the organization of chromatin in mitosis and on the other hand the changes of chromatin organization in the damage response to ionizing radiation.

In the first part we develop a model that explains the condensation of mitotic chromosomes by size-restricted dynamic looping of the chromatin fiber. Our results show also that chromatin loops can contribute to the experimentally determined bending rigidity of mitotic chromatids and generate the correct force-extension behaviour. In a next step. this folding model is then extended to describe sister chromatids by including dynamic binding and unbinding of sister fibers. We assess the interplay between cohesion and condensation and show that alignment and cohesion of sister chromatids requires detailed control of the abundance of tethering points between them.

In the second part we examine the damage response of interphase chromosomes. With an expression-dependent folding model and utilizing experimental data on the transcriptional activity of cells that were exposed to ionizing radiation, we first show that the overall organization of chromatin does not change after irradiation. By modeling actual fiber breaks in local environments we demonstrate that broken ends are passively transported to the surface of their domains and that this facilitates recognition of the break by diffusing proteins. Finally, we use a graph theoretical approach to analyze the structural changes of histone positions in localization microscopy images of cells that were exposed to ionizing radiation. We validate our previous results that no changes of the overall organization of chromatin is recognizable and demonstrate that highly packaged heterochromatic areas of the genome decondense upon irradiation.

Zusammenfassung

Entropie steuert die dreidimensionale Organisation der Chromatinfaser. Die Faltung der Chromatinfaser ist daher grundsätzlich eine Aufgabenstellung für die Statistische Physik. Die vorliegende Dissertation untersucht zwei Fragen im Zusammenhang mit der Chromatinfaltung, welche noch immer nicht vollständig geklärt sind: Einerseits die Organisation der Chromatinfaser in der Mitose, andererseits die Änderung der Organisation vom Chromosomen bei der Reaktion auf Strahlungsschäden.

Im ersten Teil haben wir ein Modell entwickelt, welches die Kondensation von Chromosomen in der Mitose durch dynamische und in ihrer Größe begrenzte Bildung von Schleifen erklärt. Unsere Ergebnisse zeigen, dass Schleifen der Chromatinfaser einen Beitrag zur experimentell ermittelten Biegesteifigkeit von Chromosomen leisten können und die korrekte Kraftauslenkungscharakteristik erzeugen. Im nächsten Schritt haben wir das Faltungsmodell durch die Einbeziehung von dynamischem Binden und Wiederablösen zwischen Schwesterfasern zur Beschreibung von Schwesterchromatiden erweitert. Wir untersuchen an dieser Stelle das Zusammenspiel zwischen Kohäsion und Kondensation und zeigen, dass die Ausrichtung und das Zusammenbinden von Schwesterchromatiden eine genaue Regulierung der Menge an Anbindungspunkte zwischen ihnen erfordert.

Im zweiten Teil untersuchen wir die Reaktion auf Schäden in Interphase-Chromosomen. Mit Hilfe eines expressionsabhängigen Faltungsmodells und unter Zuhilfenahme von experimentellen Daten zur Transkriptionsaktivität in Zellen, welche ionisierender Strahlung ausgesetzt waren, zeigen wir zunächst, dass sich die Gesamtorganisation des Chromatins nach einer Bestrahlung nicht verändert. Durch Modellierung von tatsächlichen Faserbrüchen in lokalen Umgebungen demonstrieren wir dann, dass gebrochene Enden passiv an die Oberfläche ihrer Domäne transportiert werden und sich dadurch ihre Erkennung durch diffundierende Proteine verbessert. Schließlich ziehen wir einen graphentheoretischen Ansatz zur Analyse der strukturellen Veränderung von Histon-Positionen in Lokalisationsmikroskopiebildern bestrahlter Zellen heran. Hierbei bestätigen wir, dass keine Änderung der Gesamtorganisation im Zellkern erkennbar ist, und weisen ferner nach, dass stark kompaktifizierte, heterochromatische Bereiche des Genoms eine Dekondensation nach der Bestrahlung vollziehen.

Publications Related to this Thesis

Parts of this thesis have already been published. Papers in preparation are also listed. (Information as of December 9th, 2011)

- **Y. Zhang**, D.W. Heermann Loops determine the mechanical properties of mitotic chromosomes. *PLoS ONE* **(2011)**, *6*, e29225

  DOI: 10.1371/journal.pone.0029225

- **Y. Zhang**, S. Isbaner, D.W. Heermann Mechanics of mitotic chromosomes studied with a polymer model. *Front. Phys.* **(2013)**, *1*

  DOI: 10.3389/fphy.2013.00016

- **Y. Zhang**, D.W. Heermann DNA double-strand breaks: linking gene expression to chromosome morphology and mobility. *Chromosoma* **(2013)**, *epub ahead of print*

  DOI: 10.1007/s00412-013-0432-y

- **Y. Zhang**, C. Schindler, D.W. Heermann A role for entropy in DNA double-strand break recognition and repair. *in preparation* **(2013)**

- **Y. Zhang***, G. Máté*, S. Hillebrandt, P. Müller, M. Hausmann, D.W. Heermann Measuring structural changes in chromatin induced by ionizing radiation. *in preparation* **(2013)**

* shared first authorship due to equal contribution

Contents

Acknowledgments

I would like to thank my advisor, Prof. Dieter W. Heermann for his enduring support throughout my time at the Institute for Theoretical Physics. Without his valuable advice, his knowledge about computational modeling and his commitment to present the work at numerous conferences and meetings, this work would not have been possible. I am grateful for the opportunity to do this PhD which also included a visit at the Chinese Academy of Sciences in Shanghai.

I owe many thanks to Prof. Michael Hausmann for kindly agreeing to act as a referee for this thesis.

I would like to thank Prof. Li Yixue and Prof. Liu Lei for their great hospitality during my visit at the Chinese Academy of Sciences in Shanghai.

I appreciate fruitful discussions with Prof. Roel van Driel, Prof. Jürgen Hesser and Dr. Carsten Herskind.

I am thankful to Prof. Ras Pandey for his advice and encouragement.

This work was funded by the German National Academic Foundation (Studienstiftung des Deutschen Volkes). I also appreciate support by the Institute for Theoretical Physics and travel funding by the Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGS MathComp).

The former and current members of our group have made the work atmosphere incredibly pleasant. For this I would like to thank very much, Gabriell Máté, Liu Lei, Xing Fei, Xiong Wei, Chu Min, Jörg Eisele, Andreas Hofmann, Hansjörg Jerabek, Miriam Fritsche, Manfred Bohn, Benoit Knecht, Christina Schindler and Sebastian Isbaner.

I would like to thank Gabriell Máté and Liu Lei for proofreading parts of this thesis.

I am most grateful for the unconditional support and love of my family.

My special thanks go to Karla Höning for her encouragement and her patience with me. I could not have done this without her.

Chapter 1

Aims and Scope of this Thesis

1.1 Introduction

Understanding the functioning of our genome is crucial for the comprehension of many diseases such as cancer or dementia and the development of treatments for them [1–3]. In the last few decades, significant progress has been made in this regard. It is now possible to obtain the entire human genome for an individual person. After many years of research, full genome sequencing for humans was first accomplished by the *Human Genome Project* in 2003 [4]. Only one decade later, commercial sequencing is becoming available for clinical usage while costs are dropping [5–7]. Although the knowledge of the linear sequence on our genetic material is without doubt very important, it is by itself not enough to understand the functioning of our genome. Another important aspect is its three-dimensional organization. The interplay between genomic function and its spatial organization determines the operation of cells.

The genetic material is DNA which adopts the famous double-helical form. In this form, human DNA of a cell has a length on the scale of meters but has to fit into the cell nucleus with dimensions on the scale of microns. Packaging of the genome is thus an essential aspect for genome organization. Folding of the DNA in eukaryotes is hierarchical and therefore a multi-scale problem. It is facilitated by various types of proteins. In a first step, double-helical DNA is folded into chromatin by wrapping around histone protein complexes called nucleosomes [8]. The further folding of this $10\,nm$ thick fiber possibly involves the stacking of nucleosomes, resulting in the $30\,nm$ fiber [9–11]. This thicker fiber has been observed in *in vitro* experiments, but its existence *in vivo* is not finally confirmed [12, 13].

The lifetime of a cell can be divided into different stages. Especially the higher-order chromatin organization can be vastly different dependent on the stage. In interphase, cells grow and fulfill their designated functions. One of the major tasks is gene expression, the production of specific proteins and functional RNAs. Gene expression involves RNA transcription and protein translation: The genetic information stored on the DNA is used to synthesize RNA that can later be processed into proteins. It has become very clear now that there is a tight connection between the regulation of transcriptional activity and the three-dimensional chromatin organization in eukaryotic cells [14]. Especially the

formation of chromatin loops is a key feature in gene regulation [15–18]. Furthermore, transcriptional activity is not homogeneous along the genome which contributes to the physical domain formation in interphase chromosomes [19–21].

The amount of packaging of the genome reaches its maximum during mitosis - the division of a cell into two daughter cells. In this stage, the interphase chromosomes condense into rod-like and rigid objects. In mitosis, chromosome density is so high that they become visible with conventional light microscopy. This is also the reason that the mitotic phase of chromosomes gave the first evidence for their existence [22]. The length contraction of DNA here is up to 50 times higher than during the interphase [23]. Moreover, in mitosis, the two sister chromatids are attached to each other to ensure reliable distribution of them to the daughter cells. Attachment is released only late in mitosis and immediately before chromatids are pulled to opposite poles of the cell [24].

At the same time, organisms are not isolated from their environment. Cells are constantly exposed to exogenous but also endogenous threats. Especially double-strand breaks of the DNA are extremely dangerous and can lead to chromosome rearrangements, apoptosis and the formation of tumorous tissue [25–27]. A large part of the cell resources are therefore spend on machinery for hazard control and to maintain genomic integrity.

1.2 Scope of this Thesis

In this thesis, we engage in answering two important questions in the context of chromatin organization in eukaryotes. We first target the aspect of mitotic chromosome condensation. We especially focus on the relationship between the internal organization and macroscopic mechanical properties that were observed in experiments as well as the phenomenon of sister chromatid cohesion. Our intention is to develop a model for the condensation of mitotic chromosomes and to answer the question of how tethering between sister chromatid fibers influences this condensation. In the second part we attempt to determine how entropy influences the damage response of interphase chromosomes. Our goal is to identify the structural changes that chromatin undergoes when facing severe damage by exposure to ionizing radiation. Additionally, we seek to find the role that entropy can play in the recognition of DNA double-strand breaks.

Considering the huge complexity of nuclear systems and chromosome organization, Biophysical Modeling can give valuable contributions to the understanding of these structures. Biophysics applies methods of statistical physics and thermodynamics, amongst others, to construct abstract models for highly complex systems. One factor that is of immense significance to all problems in statistical physics is *entropy*. Naturally, as physicists, we would certainly agree that entropy must also play an important role in such complex systems as a cell nucleus. And indeed, it is now becoming clearer and clearer that entropy can help to guide actions in cells. Especially, there is strong evidence that the organization of the genome is based on entropy [28–33]. Segregation processes in prokaryotic and eukaryotic cell nuclei are also driven by entropy [34–38]. And finally, entropy can be important at the binding of proteins to the DNA chain [39, 40]. All these insights reaffirm that it is physical principles which fundamentally govern the behaviour of chromosomes in the nucleus.

The two major cell stages, interphase and mitosis, host chromosomes with completely different structures. Entering mitosis, the interphase chromosome undergoes an elemental transformation into a rigid object. It is clear that many factors contribute to this transformation. Despite the seemingly huge differences between interphase and mitotic

chromosomes, our physicists intuition lets us anticipate that the same motifs must be at work in the folding of our genome in both stages of the cell cycle. The formation of loops is a general principle of chromatin organization that has been shown experimentally and by modeling for the interphase and their existence in mitosis is undoubted [17,41,42]. The design and development of a unified model for chromosome organization that covers the complete cell cycle is however still illusory or at least a very remote goal. Any attempt to find such a model must first be able to capture the structural characteristics in each of the stages.

The mechanical properties of mitotic chromosomes were studied extensively in an attempt to draw conclusions on their internal organization. Mitotic chromosomes were found to be highly elastic and could be stretched to many times of their native length with an characteristic force-extension curve that shows the emergence of an force plateau of at large elongations. Different kinds of chromosomes were analyzed and the stiffness of mitotic chromosomes were found to be very different depending on the species and the cells that they were extracted from [43]. In our model, we intend to explore how internal formation of chromatin loops can affect and contribute to the mechanical properties of single chromatids. Another important facet that could influence the mechanical properties is the fact that sister chromatids are permanently attached to each other until the anaphase of mitosis. The tethering of the two chromatin fibers may have severe impact on their condensation. The interplay between tethering and condensation is thus the major question that we pose in the first part: How is the chromatin fibre folded in mitosis and how are sister chromatids connected to each other to form such highly condensed objects?

Cells are continually threatened by damage sources including radiation, drugs and chemical agents. These sources can cause severe disturbances on the genome. Single-strand breaks - breaks of one strand in the double-helical DNA - are a frequent damage in cells [44]. But when two single-strand breaks occur in close proximity to each other, a double-strand break (DSB) can form [45,46]. DSBs belong to the most dangerous forms of DNA damage [47]. Cells must therefore quickly resolve them since they can otherwise lead to cell death or mutations [25–27,48]. The cell reaction to DSBs includes quick recognition and recruitment of repair proteins as well as local chromatin rearrangements [46,49]. But are all of these structural changes entirely actively mediated or are they automatically instated driven by entropy? Since the recent development of modeling techniques for interphase chromosomes have been very successful, this gives us the chance to analyze how chromatin organization changes after being exposed to ionizing radiation. The model that we employ for this takes gene expression data as model input. Altered transcriptional activity has been reported as an effect that ionizing radiation have on cells [50, 51]. We further explore the altered behaviour in the vicinity of actual breaks of the fiber and how such a damage could be recognized by the cell based on that. We finally attempt to verify our studies with experimental evidence and analyze super-resolution microscopy data on the structure of the genome in irradiated cells.

1.3 Structure of this Thesis

In chapter 2, a short introduction to the biological background is given, since this is the general framework for all biophysical modeling in this thesis. The introduction includes general aspects to the constitution of cells, the function of genes, the current state of research on the organization of chromosomes in different stages of the cell cycle and the damage response to double-strand breaks. Chapter 3 then introduces the physical and

mathematical methods and tools that were used for the modeling and analysis attempts in the different projects. After a short introduction to statistical physics and polymer physics, two essential simulations methods, namely Metropolis Monte Carlo Simulations and Molecular Dynamics Simulations are briefly described.

In chapter 4 we analyze what internal organization mitotic chromosomes must have to exhibit the macroscopic mechanical properties that were observed in experiments. With a coarse grained description of the chromatin fiber we develop a model for the mitotic folding of it. The key feature of our model is the dynamic and size-restricted formation of condensin-mediated loops within the fiber. This loop formation not only condenses the fiber but can also exert the entropic forces that determine the mechanical properties of mitotic chromosomes.

In chapter 5 we extend our model to sister chromatids that are subject to cohesion in mitosis. We use another implicit and dynamic mechanism to model cohesin tethering of sister fibers. We analyze how the abundance of tethering points impact on the ability of sister chromatids to align and to condense and examine what influence they have on the mechanical properties of sister chromatid systems.

The second aspect of the consequences of entropy-driven organization that we examine is the damage response of chromosomes to radiation. We employ an expression-dependent folding model for interphase chromosomes. As data input we use microarray data for transcriptional activity for cells that were also exposed to different doses of γ-irradiation. Additionally, we induce direct breaks of the chromatin fiber in different chromatin domains and analyze their dynamic behaviour. The results are presented in chapter 6.

After having analysed the global structure, in chapter 7 we then focus on the local changes to the environment caused by DNA double-strand breaks. We examine the role of entropy in DSB signaling and recruitment by assessing the search times that repair proteins take to attach to broken chromatin fibers in chromosome sub-domains with different densities.

In the final project presented in chapter 8, we analyze high-resolution images of HeLa cells that were exposed to ionizing radiation in an attempt to validate our prior modeling results. Using statistical physics methods and a graph theoretical approach we analyze how neighborhood properties of marked histones and specific antibodies change after exposed to radiation.

Finally, in chapter 9 we give a short summary of all our results and comment on possible future works and directions for further studies.

Chapter 2

Chromosome Function and Physical Organization

2.1 Cells - The Basis of Living Organisms

The only planet in the universe on which life as we know it can exist is Earth. We all have a clear conception of what life is. An universal and precise definition of life is however not as simple as one would think. According to Koshland [52], there are seven essential pillars of life. These characteristics include the existence of a metabolism, the capacity to grow and to reproduce, the ability to adapt to its environment, to react to stimuli and the capability to preserve its state [52, 53]. Finally, the organism has to be composed of the smallest entities that can be considered alive themselves - the cells [52, 54]. Cells are the building blocks of life. The simplest life forms consist of only one single cell. More complex organisms such as mammals contain trillions of cells with diverse functions [55].

All cells are enclosed by a lipid bilayer, the cell membrane, separating its interior from the environment. This membrane is selectively permeable to regulate the inward and outward flow of material [56]. All components inside of the cell are embedded in a fluid - the cytoplasm [57]. Generally, there are two different types of cells: *eukaryotic cells* and *prokaryotic cells*. Prokaryote organisms are often unicellular, the two main types of prokaryotes are *bacteria* and *archaea*. Their cells do not possess a cell nucleus. Eukaryotes on the other hand are multicellular organisms. Plants, animals and fungi are all eukaryotes. The cells of these different species all share the common concept of a cell nucleus that contains the hereditary material which is important for reproduction [58]. In Figure 2.1 an overview over the organization of cells is given. Eukaryotic cells can reproduce by mitosis which denotes the asexual division into daughter cells, or by meiosis which denotes a special kind of cell division in the process of fertilization [54, 55]. But how is genetic information of such complicated organisms stored and how are these information passed on to the next generation of cells?

2.2 DNA, Genes and Chromosomes

In the middle of the 19th century, the Augustinian monk Gregor Mendel conducted experiments with pea plants in an effort to understand the functionality of heredity. In the experiments he focused on characteristics with only two different possible values, e.g. whether the flowers were red or white. This discretization of the characteristics allowed Mendel to make precise and quantitative statements about the heredity of the plants [59]. From his pea experiments he derived the famous rules now known as *Mendel's Laws*.

Today, Mendel is seen as the father of classical genetics and the discrete characteristics are called *Mendelian characters* [57]. In modern genetic theory, they are closely connected to *genes*. Genes are structures that carry instructions for specific proteins, so called *enzymes*, to initialize certain bio-synthetic reactions eventually resulting in specific characteristics of the organism. In the example of red of white flowers, a gene is responsible for the production of red pigments. If this gene is defect, no red pigments are manufactured by the cells and the flower remains white. However there are certainly much more complex characteristics being the outcome of many subsequent reactions, each in turn mediated by genes. The set of of all genes of an organism is called the *genome*. The human genome consists of more than 40 000 genes [61].

Chromosomes are the carrier of genes [62]. The human chromosomal set consists of 23 different chromosomes, with most cells being diploid, which means that they contain two identical copies of each chromosome and thus there are 46 chromosomes in total in the cell nucleus. The material that genes and chromosomes are made of is *Deoxyribonucleic Acid*,

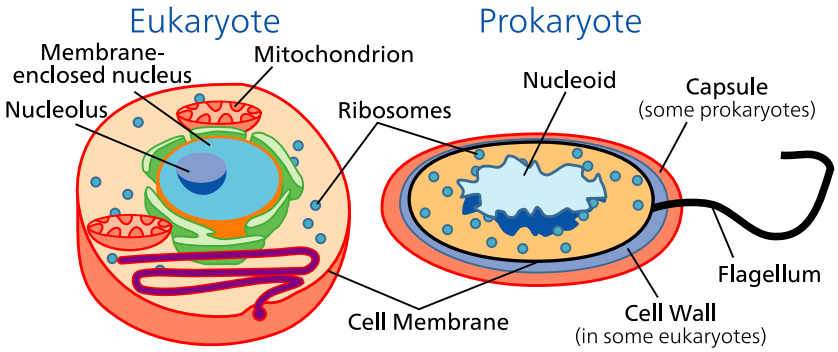

Figure 2.1: Illustration of eukaryote and prokaryotic cells. All cells have a cell membrane separating it from the outside. The components of the cell are embedded in the cytoplasm. Eukaryotic cells have a cell nucleus that contains the hereditary material while in prokaryotic cells it is just contained within the cell. The Figure is adapted from [60].

or simply *DNA*. It was first discovered as a component of the cell nucleus in the late 60s of the 19th century by Swiss doctor Friedrich Miescher at the *Universität Tübingen* [63]. However it was not until the middle of the 20th century that DNA was recognized as the genetic material. In 1953, Francis Crick and James Watson unraveled the basic structure of DNA using interpretations of X-ray crystallography experiments conducted by Rosalind Franklin and Maurice Wilkins [64, 65]. This is considered one of the most spectacular discoveries in cell biology.

The basic units for DNA are the deoxyribonucleotides. Each nucleotide is composed of a pentose (2-deoxyribose), a phosphate group and a nitrogen-containing base. Desoxyribonucleotides can have four different possible bases: *Adenine* (A), *Cytosine* (C), *Thymine* (T) and *Guanine* (G).

The deoxyribonucleotides are connected to each other via the phosphate groups, which form a phosphodiester linkage. The base groups of all the nucleotides face the same direction. This long strand is called single-stranded-DNA (ssDNA) [61]. However Crick and Watson discovered that in nature DNA is not organized in one strand, but rather in two anti-parallel strands with the bases of each strand directed towards each other. Hydrogen bonds between bases on opposing strands connect the two strands to each other. At this, A always binds to T via two hydrogen bonds, while a G base is always connected to a C base via three hydrogen bonds. Two bases connected in this way are called a basepair. The two strands are helical folded, forming the famous *double-helical* structure of DNA. Stability of the helix is guaranteed on the one hand by the hydrogen bonds, but more importantly by so-called *stacking interactions* between the bases at the center of the helix [58]. DNA forms long macromolecules and human chromosomes contain in the order of 10^8 base pairs.

2.3 Organization of Eukaryotic Chromosomes

2.3.1 Basic Organization of DNA

Histone Proteins and Nucleosomes

When unraveled, the DNA double helix of eukaryotic cells has contour lengths of up to two meters. The chromosomes however are packed in the cell nucleus on length scales

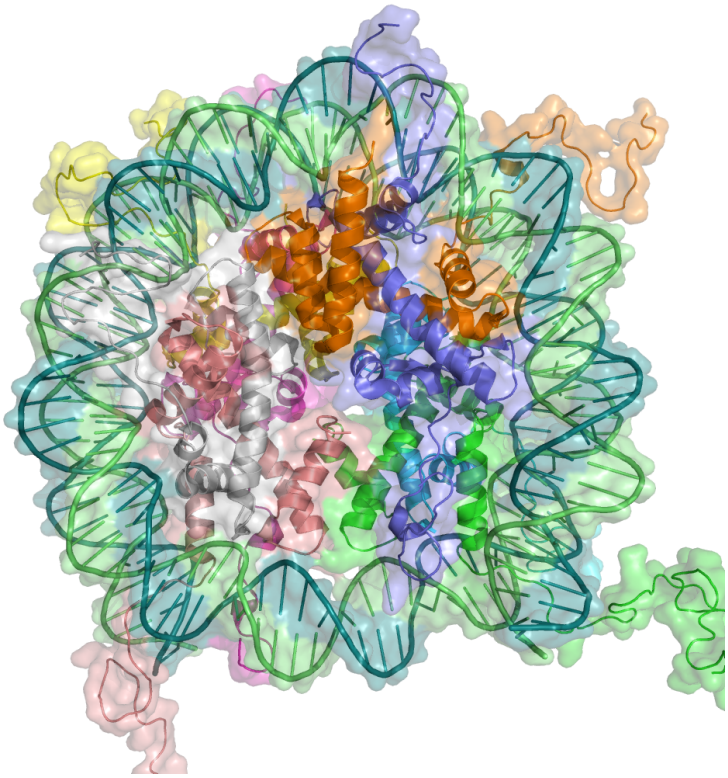

Figure 2.2: Structure of a Nucleosome. The DNA fiber is wrapped around histone octamers called nucleosomes. This first order packaging results in a beads-on-a-string type of fiber. Figure taken from [60].

of a few micrometers [54]. Therefore, chromosomes are highly compacted objects. The compaction in human cells can be in the range between 1000-fold compaction in interphase to over 15000-fold length-wise compaction during cell division [66–68]. This compaction is achieved by a hierarchical folding of the chromosome: the DNA is organized in different compaction levels and the Double Helix can be seen as the secondary structure in this hierarchy.

The folding of the double-helical DNA is facilitated by many different binding proteins. The combination of proteins and the DNA is referred to as *chromatin*. On the tertiary level, chromatin forms a *beads-on-a-string* structure with approx. 10 nm in diameter. Special proteins called *Histones* are responsible for this organization. Histones have an acidic character and are negative charged, facilitating interactions with the basic, positive charged DNA. In total, there are five different histones called H1, H2A, H2B, H3 and H4. Pairs of the latter four form spherical histone octamers around which approx. 150 bp of the Double Helix wraps. This histone octamer is referred to as a *nucleosome*. H1 is situated at the entrance of the DNA coils. Nucleosomes are lined up on the DNA fiber like beats on a string. However, this 10 *nm* fiber accounts for only a small fraction of the compaction of chromosomes in living cells [69].

The 30 nm fibre

Although it is obvious that chromatin has to be further compacted in order to fit into the cell nucleus, the way it does so still remains unclear. The idea of the 10 nm fibre coiling further into a thick filament with 30 nm in diameter was put forward very early [70, 71]. There are two competing models: in the *Crossed-Linker model*, subsequent nucleosomes are situated on opposite sides of the fibre (relative to the fibre axis) while in the *Solenoid model* they are situated directly next to each other. Either of the two models is able to describe the folding of the nucleosomes into the *30 nm fibre* [72]. However, even the existence of the 30 nm fibre in living cells is still under debate [12].

The folding behaviour above the level of the 30 nm fibre probably depends on the phase of the cell cycle. Chromatin structure in different phases are observed to be completely different. Metaphase chromosomes are condensed into the well-known, compact X-shaped organization with two *chromatides* that can be made visible via light microscopy. On the other hand, mammalian interphase chromosomes are in an dispersed state about 4 to 50 times larger than during mitosis [23]. Although growding the cell nucleus, only little intermingling between different chromosomes can be found, but rather distinct chromosome territories can be observed [73, 74].

2.3.2 The Cell Cycle

Most eukaryotic cells have the ability of self reproduction by dividing into two genetically identical daughter cells. During its lifetime, the cell and especially the cell nucleus go through different stages. This *cell cycle* can be divided into two main parts. In the *interphase*, *DNA transcription* takes place and the chromosomes are duplicated. *Mitosis* describes the process of cell division and thus the asexual transfer of genetic information. The duration of the cell cycle depends strongly on the cell type, ranging from less than an hour for frog embryos, hours for yeast cells and up to many months for human liver cells [56, 57]. Figure 2.3 gives an overview over the different stages of the cell cycle.

2.3.3 Chromosome Structure in Interphase

The interphase can be divided into 4 stages. During G_1-*Phase* the cell grows, develops its characteristic components and performs its specific tasks. Highly *differentiated* cells can pass into the G_0-*Phase* where the cell cycle is interrupted until outer signals initiate its continuation. The crucial *S-Phase* is the period in which DNA duplication is performed until all chromosomes have two identical *chromatids*. Moreover, many proteins needed for chromatin compaction such as histones, are produced during S-Phase. Finally, necessary preparations for cell division, e.g. checking for replication errors, are performed in the G_2-*Phase* [54, 58].

Gene Expression

The cell metabolism is the process by which cells fulfill their specific functions. Proteins are the main factors that regulate and maintain the cellular activity. Proteins can be build by the cell itself through protein synthesis. Other gene products that are important for cellular function are functional Ribonucleic Acids (RNAs). RNA is similar to DNA, but instead of the Guanine base group, there is an Uracil group. The production of gene products is called *gene expression*. The two main steps involved in this process are *transcription* and *translation*. While translation denotes the actual build procedure of

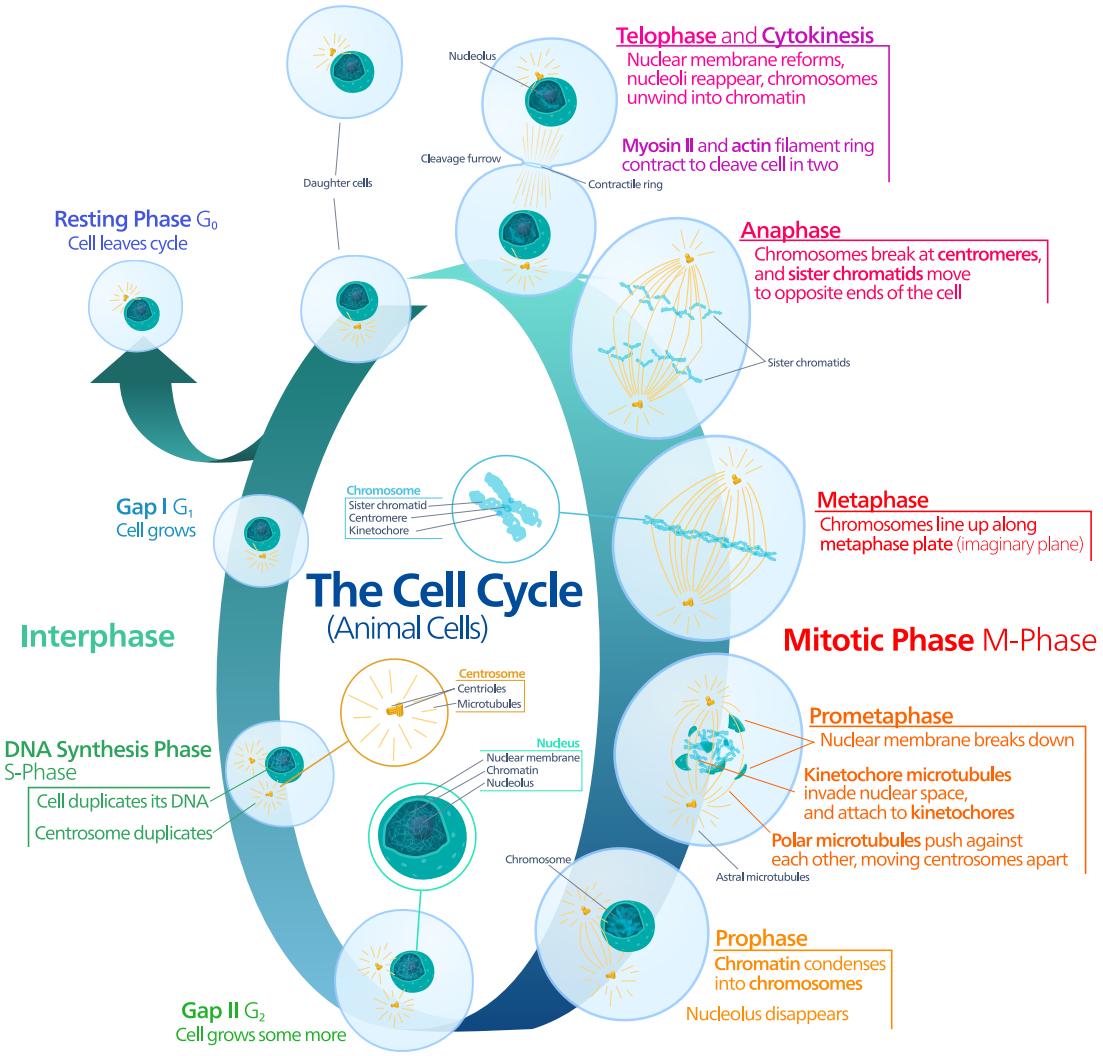

Figure 2.3: The Cell Cycle. Overview over the cell cycle in eucaryotic cells. The Figure is adapted from [60].

the protein which takes place outside of the nucleus, transcription denotes the procedure of creating the template for translation. In the transcription process, a *Messenger RNA (mRNA)* is synthesized from the genomic code. mRNA subsequently travels outside of the cell nucleus where in several steps a protein is synthesized based on the information of the mRNA [54, 56].

Transcription is initialized when the RNA polymerase enzyme attaches to a *promoter* region on the DNA chain. Well known examples for promoter regions are the *TATA Box* in eukaryotic cells. RNA polymerase then opens the double-helical DNA and reads the DNA sequence from the 3' end towards to 5' end. While reading the sequence, a RNA consisting of the complementary bases to the DNA sequence is synthesized. Each sequence of the DNA that can be transcribed into an RNA molecule is called a *transcription unit*. In eukaryotes, transcription units usually contain one single gene [54, 56, 61]. In Figure 2.4, the process of transcription of DNA by RNA polymerase is depicted schematically.

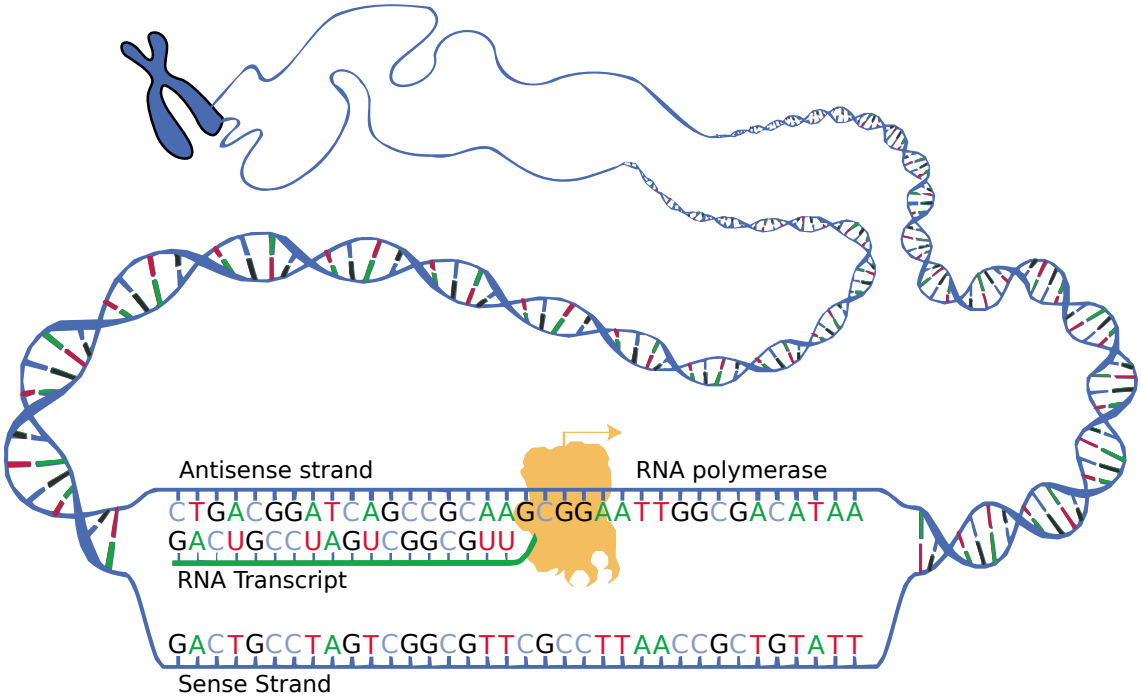

Figure 2.4: Transcription. The RNA polymerase opens the double helix and reads the DNA fiber from the 3' end to the 5' end. For each base on the DNA chain, the according base is produced for the RNA molecule. Note that on RNA there is no Guanine base and the corresponding base for Adenine is Uracil. Figure taken from [60].

Differing numbers of RNA molecules can be produced for each gene and many genes are not transcribed at all depending on the cell type. The amount of mRNA or other transcripts in the nucleus can be used as a measure of how active the expression of a gene is. Modern technologies have made it possible to measure and to evaluate the activity of a large number of predefined genes or even the whole genome. This so-called *expression profiling* yields the *expression levels* of genes and is an important tool to analyse cellular function [75]. Expression Profiling has been used to characterize malignant cells, cells that were treated with chemicals or ionizing radiation in numerous studies [50, 51, 76–79].

The most widespread method to determine expression levels is the usage of DNA microarrays [75]. The basic setup of a microarray are *DNA spots* which are attached to a solid surface in an array. The spots are chosen such that mRNA molecules can bind specifically to them. They can consist of DNA, *Complementary DNA (cDNA)* or *oligonucleotides* [80]. The specific binding of the target mRNA probes is enabled through base pairing of complementary bases on the spot and the probe, a process called *hybridization* [75]. In the process, mRNA probes are first isolated from the cell and then labeled with fluorescent tags. They are then incubated with a microarray so the mobile mRNA probes can hybridize to the specific spots on the microarray [75,81]. Finally the microarray is scanned with a microscope and the intensity data for each of the spots is computationally processed [81]. The results are relative values between the gene expression levels of the analyzed cell to control cells. Today, vast numbers of spots can be placed on a single chip, which makes it possible to profile the expression levels of a large number of genes at the same time.

Relationship between Gene Expression and Chromosome Structure

The cell possesses many different mechanisms to control the amount of each gene product that is to be synthesized. One important way for the regulation of gene expression is the regulation of the production of RNA molecules. In fact only a small fraction of the DNA is coded for gene expression. The non-coding DNA regions contain, among others, a large amount of regulatory units such as *insulators*, *enhancers* and *silencers* [54]. These genomic regions bind a large number of proteins that all belong to the class of *general transcription factors*. In the complex genome of eukaryotic cells, regulatory units for a certain gene are not necessarily close to the gene itself but can even be on other chromosomes [82, 83]. For the initiation of transcription, the proteins that are bound to these regulatory units are required to be present at the site of the transcription unit in order to aid RNA polymerase in numerous ways [54]. Therefore, genomically distant parts of the chromatin fiber can come close to each other causing the intermediate fiber segment to loop out, as has been demonstrated for various loci [84–86]. DNA is usually not present as open double stranded DNA in the nucleus but is rather packaged by nucleosome wrapping and higher-order folding. In regard of this, transcription initiation is most likely acting on a higher-order folded chromatin fiber such as the $30\,nm$ fiber [54].

Moreover, different transcription units can come physically together to be transcribed simultaneously in so-called *transcription factories*. Transcription factories are chromatin hubs where large assemblies of different functional units including transcription units, RNA polymerases and their mediators, enhancers and chromatin remodeling complexes. Especially genes that are regulated by the same regulatory units could co-localize to enhance the efficiency of transcription [87]. At this, it has been suggested that instead of the recruitment of proteins to the promoter site, transcription units could actively migrate to such functional sub-compartments which remain persistent after transcription terminated [88, 89]. Transcription factories could therefore be a major reason for the formation of sub-compartments of the chromatin fiber [90].

Generally, transcription requires relatively open forms of chromatin that is accessible for the various proteins required for the process. Additionally, in regard that distant fiber segments have to be able to loop in and out during the initiation and the termination of transcription, tight compaction would severely hinder these kind of movements. However, the organization of chromatin is by no means homogeneous along the fiber. Instead, chromatin can be organized in the open euchromatic form and the closed heterochromatic form. Heterochromatin is for example found at the centromere and the telomere region. Heterochromatic regions are silent and highly compacted. For transcription, chromatin needs to be in the open euchromatic form [91, 92]. Different studies have shown that the chromatin fiber is compartmentalized throughout interphase. It has been suggested that looping of the fiber leads to the formation of chromosome sub-domains [17, 31, 93]. The division of chromatin sub-domains into two types, an open and active one and a closed and silent one is certainly a simplified view. Newer studies suggest for example multiple different types of chromatin in *Drosophila melanogaster* based on the local proteome composition [94].

On a larger scale, different interphase chromosomes occupy different distinct *chromosome territories* in the nucleus [73, 74]. Instead of completely intermingled chromatin fibers, the overlap between different chromosomes are limited [41].

2.3.4 The Mitotic Chromosome

In mitosis, the chromosome undergoes a series of drastic changes with respect to their physical structure. In Figure 2.3, the different stages of mitosis are shown. The process starts with the *Prophase*, in which the mitotic spindle is formed and the cells start to be polarized. Furthermore, the chromosome condensation mechanism is started. In the following *Prometaphase*, chromosomes already have their X-formed shape with two sister chromatids connected to each other at the centromere. The nuclear envelope dissolves, the chromosomes attach to the *spindle fibers* (also called *microtubules*) via the *kinotochores* and move toward the *equatorial plane*. *Metaphase* begins when all chromosomes are properly aligned along the equatorial plane. Metaphase chromosomes have reached the highest degree of compaction [23]. Their condensed state allows them to be properly separated and checked for morphological errors.

The cell then passes into the *Anaphase* in which the sister chromatids are separated and pulled to different poles of the cell. This is accomplished by shortening of the microtubules and stretching of the *polar fibers* simultaneously. The final *Telophase* begins with the single chromatids reaching the poles. The nuclear envelopes build up and the condensed chromosomes unravel again. Mitosis is stopped by *Cytokinesis*, the actual cell division, and the daughter cells enter the G_1-Phase [56].

During mitosis, the dispersed interphase chromosomes undergo a transition into rigid, tightly compacted objects. The compaction ratio of DNA in mammalian metaphase chromosomes is of the magnitude of 10 000 - 20 000 [23]. The folding of the double helics into the 30 nm fibre accounts for a 40 fold compaction of the naked DNA. The higher order folding motifs that are responsible for the remaining approx. 500 fold compaction still remain largely unknown [95, 96].

Condensin and Cohesin

In the 1970s, fascinating images were produced in histone-depletion experiments under high salt conditions showing a protein core to which extended DNA loops of 30 to 90 *kb* were attached [97]. The protein core was observed to have similar sizes to untreated chromosomes [68]. Based on these studies, the *Radial Loop Model* was proposed where this central protein scaffold determines the chromosome shape with the DNA being bound to it in loops [58, 98]. The protein scaffold was found to be structurally independent of DNA with the main components being *condensin* and *Topoisomerase II* [99–101]. While Topoisomerase II seems to be required for the condensation process but not for its maintenance, the condensin complex is crucial for both [102–104].

Another important aspect of mitotic chromosome structure is the connection between sister chromatids. Sister chromatids are attached to each other until metaphase and subsequently pulled apart. Without the attachment, sister chromatid would segregate after condensation due to physical and entropic forces [28, 36]. The cohesin protein is the main factor responsible to hold sister chromatids together [105]. Different mechanisms were suggested for the way cohesin does this. One suggestion is that cohesin forms rings and encircles both sister chromatin fibers [106]. Another possibility is that two cohesin proteins each form a ring around one sister chromatin fiber. The fibers could then be attached to each other by binding of the two rings forming a handcuff-like structure [107]. In anaphase, cohesin is cleaved from chromosomes enabling the microtubules to segregate sister chromatids [24]

The interplay between condensation of the fiber which is facilitated by condensin and

the tethering of sister chromatids facilitated by cohesin govern the organization of mitotic chromosomes. Furthermore, condensin could be directly involved in cleavage of cohesin from chromosome arms aiding the eventual segregation of sister chromatids [108, 109].

Mechanical Properties of Mitotic Chromosomes

A different approach to the analysis of chromosome structure in metaphase was by measuring the elastic properties, as they can give insight into the organization and inner structure of objects [43]. Early qualitative measurements showed that human mitotic chromosomes and chromosomes from Xenopus egg extract are very elastic objects that can be stretched to many times of their native length [110, 111]. Houchmandzadeh and Dimitrov [112] measured for the first time bending and stretching elasticity of in vitro assembled mitotic chromosomes from Xenopus egg extract. The authors found the chromatids to be very flexible objects with the persistence length being only a few times the thickness of the chromatids. Based on these findings, the existence of thin rigid cores inside the chromatid was proposed [96].

For small extensions the authors reported a linear force-elongation behaviour and reversible deformability. Moreover, chromosomes were extensible up to 100 times their native length with a force plateau being observed at relative extensions larger than ≈ 15 [111,112]. This kind of elastic response was also confirmed for chromosomes from newt lung cells. The newt chromosomes showed reversible extension up to three times native length. For intermediate elongation, hysteresis was observed and for long extensions beyond 30 times native length, the linear force-extension curve decreased to a plateau [113].

While results on the stretching stiffness agree widely for chromosomes and chromatids of different species, this is not the case for the bending rigidity. As described above, metaphase chromosomes from Xenopus egg extract were reported to be very flexible [96, 111]. However, experiments on in vitro and in vivo assembled chromosomes from other animals yielded much higher rigidities, with persistence lengths being much larger than the lengths of the chromosomes themselves. Marshall et al [114] examined Drosophila chromosomes and found high bending rigidities. In many different experiments for Xenopus and newt chromosomes from prophase, prometaphase to metaphase, Poirier et al [115] also confirmed that chromosomes seemed to be very inflexible. This indicates that there is a difference in the inner structure between chromosomes from different species. It was also speculated that chromosomes from Xenopus egg extract were not fully condensed [43]. Also there could be differences between single chromatids and chromosomes with two connected chromatids [43].

Further experiments by Poirier et al [116,117] targeted the force-relaxation behaviour of stretched chromosomes, especially during nuclease digestion. Single chromosomes were put under strain using micro-pipettes and then treated with DNA-cutting enzymes. A following drop off of the strain could be observed, as DNA was cut due to the nuclear digestion. Therefore, the authors concluded that mitotic chromosomes do not have a mechanically contiguous protein scaffold, as otherwise the chromosome would be capable to support an applied tension even after nuclease digestion. In spite of this, a network model was put forward where the higher order structure was explained by the cross-linking of the 30 nm chromatin fibre [116, 118].

2.4 DNA Double-Strand Breaks

2.4.1 Causes of DNA Double-Strand Breaks

Cells are constantly exposed to many different sources of danger. Damage to DNA is a very frequent hazard and the cell possesses well adapted defense and repair mechanisms for them. Oxidative reactive agents produced by the metabolism and environmental chemicals can damage bases of nucleotides. Ultraviolet light from the sun is sufficient to alter the covalent bindings between nuclease and even thermal fluctuations can cause spontaneous reactions causing mutations. The abundance of damage to the genetic material that happens every day have caused cells to invest large amounts of resources to deal with them [54].

Single-strand breaks (SSB) are a very frequent type of damage to the DNA fiber [44]. For the repair of SSBs and other simple damages, the double-helical structure that our genetic material adopts is of greatly beneficial. The undamaged complementary strand can usually serve as a template for example to synthesize correct sequences of missing nucleotides. Repair becomes more challenging when both strands of the DNA fiber are damaged. This is for example the case when both strands of the DNA are broken and a *DNA double-strand break (DSB)* emerges.

DSBs can form when two SSBs occur in close proximity each other. Two major causes for DSBs are ionizing radiation and reactive oxygen species that are produced by the metabolism [119]. DSBs can also form following failure of the *Topoisomerase II* machinery to rejoin breaks that were induced before to enable fibers to pass through each other [120]. All organisms are constantly exposed to ionizing radiation, terrestrial from the decay of radioactive elements on earth and in the atmosphere as well as cosmic radiation. Since water is a main constituent of living organisms, ionizing particles that travel through living tissue can produce free radicals which in turn are able to damage and even break DNA [44]. An estimated 10 DSBs form in each cell every day [119].

DSBs threaten the existence of the cell in many different ways. The break can lead to general genomic instability and and even breakage of the chromosome. The loss of genetic material can also be a consequence of DSBs. Misrepair of DSBs can lead to chromosome rearrangements that alter the hereditary information leading to mutations and even the development of tumor tissue [27]. The correct processing of DSBs is therefore crucial for the maintenance of cellular function and genomic stability.

2.4.2 Cell Reaction to Double-Strand Breaks

Damage Recognition

Upon formation of double-strand breaks, a signaling cascade that involves various intricate, multi-level responses to the DSB is triggered in the cell. One of the earliest proteins that were observed to arrive at the site of DSBs in eukaryotic cells is the *MRN complex* which constitutes of Mre11, Rad50 and Nbs1 and has many functions related to DSB repair [121]. The three constituents are able to associate in different ways and form different kinds of structural complexes. Most importantly, MRN is abundant in the cell nucleus and able to bind to DNA [122]. Its fast association to areas with damaged DNA was shown in microscopy experiments [123]. MRN is then able to precipitate the accumulation of the kinase *ataraxia telangiectasia mutated (ATM)* to the break site [124, 125]. These quickly formed congregation of repair proteins are called *radiation induced foci (RIF)* and can be made visible by flurescence microscopy [126]. After arrival, ATM phosphorylates the tail

of histone variant H2AX at Serine 139 into γ-H2AX [127]. γ-H2AX can be found at up to one megabasepair around the break and is one of the most commonly used markers for DSBs [49, 128, 129]. ATM is also responsible for the further recruitment of various kinds of mediators and end processing proteins.

Main Repair Pathways and Repair Kinetics

There are two main repair pathways in prokaryotic and eukaryotic cells for DSBs: *Non-homologous end joining (NHEJ)* and *Homologous recombination (HR)*.

Homologous Recombination All homology-directed repair requires the nearby presence of a homologue template for the broken DNA. HR is therefore not available in haploid organisms and only available during the S and G2 phase of diploid organisms when the sister chromatid can serve as a template. During HR, the DNA fiber containing the break is first degraded and Rad51 loaded onto the single-stranded DNA. Rad51 further is able to search for the homologous template and facilitates strand invasion and the formation of a physical link between the broken DNA and the homologous template creating *heteroduplex DNA (D-loop)*. Different models exist for the subsequent repair of the DNA on the broken strand, including the formation of a *double Holiday Junction* and *strand displaced annealing*. HR is also involved in genetic recombination during meiosis. Due to the utilization of a template, HR is able to repair broken DNA with high accuracy and low chance for the loss of genetic information [130].

Non-homologous End Joining The sister chromatid is not available as a template in haploid cells and diploid cells prior to S phase. The repair pathway of choice is then non-homologous end joining. The Ku heterodimer plays a main role in NHEJ in eukaryotic cells. It forms a ring-like structure and is able to slide over the broken ends of the DNA double-helix after arrival at the site of the DSB [131]. NHEJ requires polymerases to re-synthesize DNA after the damaged nucleotides have been removed. In the final step, the broken ends are ligated, which is taken over by the DNA ligase IV complex in eukaryotes [119]. Polymerases and ligands are recruited to the DSB by the catalytic subunit of *DNA protein kinase* (DNA-PKcs) which is itself activated by Ku. Repair by the NHEJ pathway is by its nature more susceptible to errors and loss of genetic information. However, it is available in all stages of the cell cycle and embodies a reliable method for fast repair [46].

Chromatin Changes

In the process of DNA damage response, various alterations to chromatin structure arise. Changes on the nucleosomal scale involve for example the above mentioned phosphorylation of H2AX to γ-H2AX which is characteristic for DSBs. This phosphorylation itself was suggested to further promote chromatin rearrangements [127, 132]. Local expansion of surrounding chromatin after induced DSB was reported in a study of GFP-tagged H2 proteins [133]. The authors also noted the lack of large scale movement of the surrounding chromatin. Positional stability of DSB regions was also observed by Soutoglou et al [134]. However, local movement of sites of α-particle irradiation DSBs was reported by Aten et al [132] and Jakob et al [135]. A recent study confirmed the increased mobility of regions containing DSBs [136]. On the other hand, the recruitment of compaction promoting proteins such as HP1 and KAP1 seems to contradict a decondensation of chromatin surrounding DSBs [137, 138].

Clear differences have been detected for the treatment of DSBs in euchromatin (EC) and in heterochromatin (HC) in eukaryotic cells. In the highly condensed HC regions, γ-H2AX foci formation was much weaker and they were less frequently found compared to active EC [139–141]. Recently, a study showed that upon passing through HC, heavy-ion radiation does not induce formation of radiation-induced foci in their path but rather at the periphery of the heterochromatic domains [142]. Microscopy of Drosophila cells confirmed the early formation of γ-H2AX in HC domains and their quick relocation to the domain periphery [143]. Indeed, the tight compaction in heterochromatin may restrict the accessibility of strand breaks to repair proteins. Moreover, it might limit the physical effects of a DSB, thereby delaying its recognition. It was also speculated that quick transport of DSBs to the periphery of HC domains could be the explanation for the experimental observations [144].

Chapter 3

Physical Models for Chromosomes

3.1 Brief Introduction to Statistical Physics

In statistical physics, the properties of populations and large systems consisting for example of atoms or molecules in a gas are described with statistics and probability theory methods. Instead of the description of the detailed behaviour of each single component of the system, the goal of statistical physics is to characterize the system as a whole based on assumptions about the probabilities of microscopic states.

3.1.1 Ensembles

Statistical Ensembles are the means by which systems are described in statistical physics. An ensemble denotes a set of infinite virtual copies of the system. These virtual copies represent all possible configurations that the components of the system can take. Averaging over ensembles then yields the properties for the whole system and the connection between microscopic configurations and macroscopic properties. There are generally three different kinds of statistical ensembles.

In the **Microcanonical Ensemble**, the volume, energy and the number of particles of the system is conserved. All accessible states of the system have thus the same energy and the same particle number. The distribution over the different configurations is an uniform distribution.

$$p_i = \frac{1}{\Omega} \tag{3.1}$$

where p_i denotes the probability for state i and $\Omega = \sum_i 1$ is the *Microcanonical Partition Function*.

Energy conservation of the system requires its complete isolation from the environment. Such conditions are usually not met by real systems. The **Canonical Ensemble** assumes that the system is in thermal equilibrium with its environment which serves as a thermal bath. The system then can exchange energy with this bath until it has equal temperature. Instead of conserved energy, the temperature of the system in the canonical ensemble is conserved. The distribution over the different configurations is given by the Boltzmann Distribution:

$$p_i = \frac{1}{\mathcal{Z}} \exp\left(-\frac{E_i}{k_B T}\right) \tag{3.2}$$

with $\mathcal{Z} = \sum_i \exp\left(-\frac{E_i}{k_B T}\right)$ being the *Canonical Partition Function*.

A system that can not only exchange energy but also particles with its environment can be described by the **Grand Canonical Ensemble**. Here, the system has conserved volume V, temperature T and chemical potential μ while the particle number N is not fixed. The distribution over the possible configurations can be derived from the canonical ensemble and is given by

$$p_i = \frac{1}{\mathcal{Z}_\mathcal{G}} \exp\left(-\frac{E_i - \mu N_i}{k_B T}\right) \tag{3.3}$$

where $\mathcal{Z}_\mathcal{G} = \sum_i \exp\left(-\frac{E_i - \mu N_i}{k_B T}\right)$ is the *Grand Canonical Partition Function*. It can also be expressed in the form $\mathcal{Z}_\mathcal{G} = \sum_{N=1}^{\infty} e^{\mu N} \mathcal{Z}$ with the canonical partition function $\mathcal{Z}$.

3.1.2 Ergodic Hypothesis

When a thermodynamic system evolves in time, it explores different microstates in phase space. The *Ergodic Hypothesis* declares that the time spent in microstates of equal energy is proportional to the phase space volume of these states. On very long time scales, the system will eventually explore all accessible microstates. Let A be an observable. The time average of A following the trajectory of the system through time evolution is then given by

$$\overline{A} = \lim_{t\to\infty} \frac{1}{t} \int_0^t A(t')dt' \tag{3.4}$$

On the other hand, the ensemble average of the system is given by the probability distribution over the microstates in phase space

$$\langle A \rangle = \sum_i p_i A_i \tag{3.5}$$

with the sum going over all microstates i. Essentially, the Ergodic Hypothesis means that the time average over a system is the same as the ensemble average over the system: $\overline{A} = \langle A \rangle$ [145].

3.1.3 Statistical Physics in the Context of Chromosome Organization

The folding of the chromatin fiber is essentially a problem of statistical physics. The cell nucleus consists of a large number of particles including the DNA, many different types of RNA and a huge variety of proteins. It is clear that, with current resources, any attempt to describe this system on a detailed level is destined to fail. In fact, even the folding of a single chromosome on a coarse grained level can already be computationally demanding. At the same time, there is certainly not one single correct configuration of all the particles in the cell nucleus. Instead, we have to think of it as a statistical ensemble, where many configurations are accessible. The possible configurations of the cell nucleus can be accessed by time evolution which involves the movement of its particles. On the other hand, different cells can certainly be expected to adopt different configurations for their cell nuclei. These two scenarios thus exactly match the two scenarios in statistical ensembles - the virtual copies of the system and the time evolution.

However, the time scale of the progression of a cell nucleus through phase space and the explorations of a large enough number of accessible microstates is not clear yet. While local movement of chromatin is a frequent phenomena, large scale conformational changes could in fact require long-term progression through the cell cycle [146]. There have also been suggestions that the cell nucleus cannot be thought of as being in equilibrium [147]. This would mean that chromosomes are never equilibrated. However, the equilibration time scale could be well beyond the typical length of a cell cycle [147].

Many experiments in cell biology target the organization and function of chromosomes in the cell nucleus. Microscopy experiments usually examine single cells for which the structure was fixated. While the best way to characterize an ensemble is indeed to gain enough accessible configurations, it is experimentally also very hard to do this via microscopy. On the other hand, in high-throughput biochemical experiments such as the *Chromosome Conformation Capture* technology, a large number of cells are analyzed at once, often losing the information on each single cell but yielding a large sample size. Considering this, it is not sufficient for any modeling attempt of chromosomes to just provide a

single configuration of the chromosome [41]. Rather, the result of the model needs to be a whole ensemble of accessible conformations and the distribution of this ensemble has ideally to agree with experimental findings. Likewise, any reverse engineering attempt that aims at the reconstruction of the chromosome organization from experimental results must provide an ensemble of possible configurations instead of only one single configuration.

A major difficulty for experiments that work on a large number of cells is to ensure that each of the cells can really represent one state of the same system. Throughout the cell cycle, chromosomes can adopt hugely different structures that clearly do not belong to the same ensemble. Thus, for experimental analysis, cells need to be in an as similar condition as possible. Synchronized cells are therefore very important for analyzing chromatin organization in high-throughput experiments.

3.2 Physics of Polymers and Macromolecules

Polymers are macromolecules consisting of repeating subunits called *monomers*. The physical properties and conformational motion of polymers can be described by means of statistical physics. They are interesting for many different areas of science such as materials science, solid-state physics, biophysics or molecular biology. DNA can for example be seen as a polymer with the nucleotides being the repeat units. On a more coarse grained level, chromatin fibers of different hierarchies can be modeled by a polymer, too. Most importantly, the physical organization and mechanical properties of DNA and of chromatin can be efficiently described using polymer statistics. In this section a short introduction to basic polymer physics, different polymer models and quantities to determine polymer characteristics is given.

3.2.1 Freely-Jointed-Chains and Gaussian Chains

The most simple polymer model is the *freely-jointed chain* (FJC). The polymer is described by a number N of vectors $\mathbf{b}_i$ with fixed length b, connected to a chain. The direction of the vectors $\mathbf{b}_i$ are completely independent from each other. Thus the FJC corresponds to a *random walk* (RW) of $N+1$ steps or monomers with their position vectors denoted by $\mathbf{r}_i$.

End-to-end distance

The end-to-end distance vector denotes the vector from the position of the first to the position of the last monomer in the chain. It is an important measure for the size of a polymer.

$$\mathbf{R}_e \;=\; \mathbf{r}_{N+1} - \mathbf{r}_1 \;=\; \sum_{j=1}^{N} \mathbf{b}_j \tag{3.6}$$

The mean squared end-to-end distance over the thermal ensemble of a freely-jointed chain can be calculated analytically

$$\left\langle \mathbf{R}_e^2 \right\rangle \;=\; \sum_{i=1}^{N}\sum_{j=1}^{N} \langle \mathbf{b}_i \cdot \mathbf{b}_j \rangle \tag{3.7}$$

As all vectors are independent, we have $\langle \mathbf{b}_i \cdot \mathbf{b}_j \rangle = b^2\,\delta_{ij}$ and thus the mean squared end-to-end distance is simply given by

$$\left\langle \mathbf{R}_e^2 \right\rangle \;=\; N\,b^2 \tag{3.8}$$

Often this relationship is written in the form of a scaling law

$$\left\langle \mathbf{R}_e^2 \right\rangle \;\sim\; N^{2\nu} \tag{3.9}$$

where ν is called the scaling exponent and in the case of the freely-jointed chain, has the value $\nu = 0.5$. The exponent is different for different polymer models. Because of the independence of different chain segments, the freely-jointed chain is self-invariant [148]. Therefore, the mean squared distance between monomers j and k obeys the same scaling law

$$\left\langle (\mathbf{r}_j - \mathbf{r_k})^2 \right\rangle \;=\; |j-k|\,b^2 \tag{3.10}$$

Distribution of the end-to-end distance It is interesting to know not only the mean value of the squared end-to-end distance but the probability distribution function of the end-to-end vectors $p(\mathbf{R}_e)$. In case of the FJC, this function can be found easily by using the central limit theorem as the vectors $\mathbf{b}_i$ are all independent. In the limit of $N \to \infty$, the end-to-end vectors are distributed according to a Gaussian distribution [149]:

$$p(\mathbf{R}_e) \;=\; \left(\frac{3}{2\pi N b^3}\right)^{3/2} \exp\left(-\frac{3\mathbf{R}_e^2}{2Nb^2}\right) \tag{3.11}$$

This probability distribution is spherically symmetric and thus the probability to find an end-to-end vector with the length in $[R_e, R_e + dR_e]$ is given by

$$4\pi R_e^2\, p(R_e) \;=\; \left(\frac{6}{\pi N b^3}\right)^{3/2} R_e^2 \exp\left(-\frac{3R_e^2}{2Nb^2}\right) \tag{3.12}$$

Radius of gyration

Another important quantity is the radius of gyration. It describes the deviation of the monomers from the center of gravity of the polymer. Thus, the radius of gyration can be seen as a measure of the spatial latitude of the chain and its mean square value is given by

$$\langle R_g^2 \rangle \;=\; \frac{1}{N}\sum_{i=1}^{N} (\mathbf{r}_i - \mathbf{r}_{cm})^2 \tag{3.13}$$

with $\mathbf{r}_{cm}$ being the center of mass of the monomers. For the FJC the mean squared radius of gyration obeys the same scaling law as the mean square end-to-end distance. For polymers where all monomers have identical mass, the radius of gyration is equivalent to the moment of inertia.

While the length of all bond vectors in the freely-jointed chain are fixed, in the *Gaussian Chain Model* (GCM) the length of the bond vectors can fluctuate. More precisely, the bond vectors are Gaussian distributed

$$G(\mathbf{b}_i) \;=\; \left(\frac{3}{2\pi b^2}\right)^{3/2} \exp\left(-\frac{3\mathbf{b}_i^2}{2b^2}\right) \tag{3.14}$$

Note that the mean length of a bond vector is given by $\langle \mathbf{b}_i^2 \rangle = b^2$. The model still assumes freely-jointed bonds. Therefore, the mean square end-to-end distance is still given by $\langle \mathbf{R}_e^2 \rangle = N\, b^2$. The end-to-end distance distribution for the Gaussian chain model is also the same Gaussian distribution as for the freely-jointed chain. A Gaussian chain is equivalent to a chain where the monomers are connected by springs with spring constants $k = k_B T/b$. The Hamiltonian for such a chain has the form

$$H \;=\; \frac{3k_B T}{2b} \sum_{i=2}^{N} \left(\mathbf{r}_i - \mathbf{r}_{i-1}\right)^2 \tag{3.15}$$

3.2.2 Flexibility of Polymers

In the freely-jointed chain and the Gaussian coil, all bond vectors are uncorrelated. However for real macromolecules it is clear that this is not always the case. On the contrary, long polymers usually have bending rigidities that limit their flexibility, i.e. bending costs energy. While for a Gaussian coil the mean correlation between the bond vectors is $\langle \mathbf{b}_i \cdot \mathbf{b}_i \rangle = 0,\ i \neq j$, for real polymers this correlation is non-zero when the bonds are close to each other.

Kuhn Segment

A quantity that gives information about the stiffness of a polymer is the Kuhn Segment, named after Werner Kuhn. Kuhn studied the conformational properties of long filament-like molecules and compared them to freely-jointed chains by coarse-graining the filaments into segments with a certain length ξ_k. For an appropriate choice of ξ_k the coarse-grained polymer can be seen as a freely-jointed chain. Therefore ξ_k, called the *Kuhn Length*, is also a measure for the stiffness: stiff chains have large Kuhn lengths and more flexible chains have smaller Kuhn lengths. On length scales that are much larger than ξ_k a polymer can be seen as a random walk, but on length scales much smaller than ξ_k it behaves like a stiff rod. Introducing the Kuhn length allows the description of a real chain by a random walk (or a self-avoiding walk) on a larger length scale. Naturally the freely-jointed chain has per definition a Kuhn length which is equal to its bond length [149].

Persistence Length

To quantitatively describe the flexibility of polymers the directional correlation of different segments of the polymer can be used. A possibility to determine the directional correlation is by calculating the mean cosine of the angular correlation between different chain segments. Let $\mathbf{u}(s)$ be the direction of a chain segment at the contour length s. Then the correlation function between two segments separated by the contour length s' is

$$\langle \cos\theta(s') \rangle \;=\; \langle \mathbf{u}(s) \cdot \mathbf{u}(s+s') \rangle \tag{3.16}$$

Where the average means average over all positions s within the polymer and also the phase space of all polymer conformations in equilibrium. A quantity that measures the stiffness of the chain with respect to the orientational correlation is the persistence length ξ_p. It is defined as the integral width of the correlation function

$$\xi_p \;=\; \int_0^\infty \langle \cos\theta(s') \rangle ds' \tag{3.17}$$

The persistence length is proportional to the bending modulus B [150]:

$$B \;=\; k_B T \xi_p \tag{3.18}$$

The bending modulus can be described as the quantity which determines how much force is necessary to bend a segment of the chain to a certain curvature [148]. The SI-units of B are Nm^2.

Worm-Like Chain Model (WLC)

The *Worm-Like Chain Model*, which was put forward by Kratky and Porod, describes an elastic, homogeneous and continuous filament. Such a filament has a restricted flexibility as neighboring segments have influence on each other simply through mechanical interaction. The flexibility is therefore determined by Hooke's law. For finite temperatures the polymer bends due to thermal fluctuations, while for $T = 0\,K$ the filament is a rigid rod. In this model the directional correlation is assumed to be multiplicative

$$\langle \cos(s + s') \rangle \;=\; \langle \cos(s) \rangle \cdot \langle \cos(s') \rangle \tag{3.19}$$

This important attribute of the orientational correlation results directly in an exponential behaviour in dependency of the contour length s separating two segments

$$\langle \cos\theta(s) \rangle \;=\; \exp\left(-\frac{s}{\xi_p} \right) \tag{3.20}$$

with ξ_p being the persistence length. The WLC model is often used to describe real chain molecules.

In the continuous limit for the worm-like chain model, i.e. $N \to \infty$, $b \to \infty$, $Nb^2 \to L$, the trajectory of the chain can be described by the tangent vectors which are $\mathbf{u}(s) = \frac{\partial \mathbf{r}(s)}{\partial s}$. Here s denotes the arc length and $\mathbf{r}(s)$ is the position in space at a certain arc length. The total arc length is denoted by L. For this now continuous chain the end-to-end distance is given by the simply path integral

$$R_e \;=\; \int_0^L ds\, \mathbf{u}(s) \;=\; \int_0^L ds\, \frac{\partial \mathbf{r}(s)}{\partial s} \tag{3.21}$$

In order to obtain a relationship between the persistence length and the bending modulus B that is familiar from elasticity theory, the Hamiltonian of the persistent chain has to be derived. The bending energy at any point along the chain is given by

$$\frac{\partial E(s)}{\partial s} = \frac{B}{2} \left(\frac{\partial \mathbf{u}(s)}{\partial s} \right)^2 \tag{3.22}$$

and thus the Hamiltonian of the chain in the continuous limit is

$$\mathcal{H} \;=\; \frac{B}{2} \int_0^L \left(\frac{\partial \mathbf{u}(s)}{\partial s} \right)^2 ds \;=\; \frac{B}{2} \int_0^L \left(\frac{\partial^2 \mathbf{r}(s)}{\partial s^2} \right)^2 ds \tag{3.23}$$

where B is the bending modulus. Using this Hamiltonian, the relationship (3.20) is obtained and it can be shown that in three dimensions the bending modulus and the persistence length are proportional

$$B \;=\; k_B T\, \xi_p \tag{3.24}$$

Therefore for a persistent chain the bending modulus can be directly obtained by measuring the persistence length using (3.20).

Another possibility is to analyse a discrete worm-like chain. The bending energy in any joint is given by $E = B\ (\mathbf{u_i} \cdot \mathbf{u}_{i+1})$ and thus the Hamiltonian is

$$\mathcal{H} \;=\; -B\sum_{i=1}^{N} \mathbf{u_i} \cdot \mathbf{u}_{i+1} \tag{3.25}$$

We see that this Hamiltonian is exactly the same as for the one-dimensional Heisenberg model for ferromagnets. Calculations using transfer matrix and renormalization group theory for this Hamiltonian yield the same result for the directional correlation (3.20) and the linear relationship between bending modulus and persistence length is also obtained.

An expression for the mean squared end-to-end distance of the worm-like chain model can easily be derived. Again using (3.20) we get

$$\begin{aligned}
\langle R_e^2 \rangle &= \langle R_e \cdot R_e \rangle && (3.26)\\
&= \int_0^L ds \int_0^L ds' \langle \mathbf{u}(s) \cdot \mathbf{u}(s') \rangle && (3.27)\\
&= \int_0^s ds \int_0^{L-s} d(s-s') \exp\left(-\frac{s-s'}{\xi_p}\right) && (3.28)
\end{aligned}$$

The integration finally yields a Debye function

$$\begin{aligned}
\langle R_e^2 \rangle &= 2L\xi_p \left[1 - \frac{\xi_p}{L}\left(1 - e^{-L/\xi_p}\right)\right] && (3.29)\\
&= L^2 f_D\left(-\frac{L}{\xi_p}\right) && (3.30)
\end{aligned}$$

A look at the two limiting cases of very long and very short chains respectively, confirm the statements about Kuhn and persistence length of the previous section. In the limit of chains much shorter than the persistence length $L \ll \xi_p$, the exponential term can be expanded to $e^{-L/\xi_p} \approx 1 - \frac{L}{\xi_p} + \frac{L^2}{2\xi_p^2}$. We then get a constant mean squared end-to-end distance $\langle R_e^2 \rangle = L^2$. This means that the chain is a stiff rod with no bending. On the other hand, for long chains ($L \gg \xi_p$) we find $\langle R_e^2 \rangle = 2L\xi_p \sim N$ as $L = Nb$. This is the same behaviour as a freely-jointed chain.

$$\begin{array}{lllll}
L \ll \xi_p & \rightarrow & \langle R_e^2 \rangle = L^2 & \sim N^2 & \text{stiff rod}\\
L \gg \xi_p & \rightarrow & \langle R_e^2 \rangle = 2L\xi_p & \sim N & \text{freely-jointed chain}
\end{array}$$

We see that when coarse-grained with a suitable Kuhn length, worm-like or persistent chains can be treated like freely-jointed chains. It can be derived analytically that for worm-like chains the Kuhn length is approximately twice the persistence length $\xi_k \approx 2\xi_p$. We see that indeed both length scales are tightly connected. Here we would like to point out that in a recent study, it was shown that the concept of persistence length does not apply for real macromolecules [151].

3.2.3 Volume Interactions

Only in rare cases do real macromolecules behave like freely-jointed chains. An important feature that has to be included is the self-avoidance of the chain. When we keep in mind that each monomer represents a physical unit such as a group of atoms or molecules, then it becomes clear that two monomers cannot occupy the same space. There has to be a short-ranged repulsive potential preventing two monomers from hitting each other. This repulsion leads to a swelling of the chain in comparison to Gaussian coils. On the other hand, conformations occupying more space are entropically less favorable than more compact conformations. Thus polymers with excluded volume interaction have to equilibrate these two effects. The swelling can be quantitatively described by the swelling parameter α:

$$\alpha^2 \;=\; \frac{R^2_{e,\,vol}}{R^2_{e,\,gauss}} \tag{3.31}$$

It compares the spatial dimensions of the polymer with volume interactions to freely-jointed chains.

Theory According to P.J. Flory

First theoretical descriptions of polymers with volume interactions reach back to P.J. Flory [149]. The total free energy of a swollen coil is given by

$$F(\alpha) \;=\; F_{el}(\alpha) + F_{int}(\alpha) \tag{3.32}$$

where F_{el} is the entropic elasticity of the chain and F_{int} describes the volume interaction of the links. The first term can be approximated by

$$F_{el} \;=\; -\,T \cdot S \tag{3.33}$$

and using $S = k_B \cdot \log[p(\mathbf{R}_e)]$ this yields

$$F_{el} \;=\; \frac{3}{2} k_B T \cdot \frac{R_e^2}{N b^2} \tag{3.34}$$

This term represents the free energy of a coil without volume interactions, i.e. a Gaussian coil. It can be interpreted as an entropic spring with spring constant $\frac{3k_BT}{Nb^2}$.

The second term represents the interaction energy of the links. For low concentration systems only pairwise interactions need to be considered. Higher-order interactions, such as interactions involving three or more links can be neglected due to their low probabilities. The average number of pairwise interactions of N monomers with volume $\frac{4}{3}\pi b^3$ in a total volume $\frac{4}{3}\pi R_2^3$ is approximately $N^2 \frac{b^3}{R_e^3}$. Let β_0 denote the mean energy per interaction then the total contribution of interaction energy to the free energy is

$$F_{int} = \beta_0 \frac{N^2 b^3}{R_e^3} \tag{3.35}$$

The estimation of F_{int} using the average number of pairwise interactions corresponds to a *mean-field theory* of the problem. An average over all interactions is taken and fluctuations are neglected. The total free energy reads

$$F \;=\; \frac{3}{2} k_B T \cdot \frac{R_e^2}{N b^2} + \beta_0 \frac{N^2 b^3}{R_e^3} \tag{3.36}$$

The free energy can now be minimized in terms of R_e and we get

$$\langle R_e^2 \rangle \sim N^{2\nu}, \quad \nu = \frac{3}{5} \tag{3.37}$$

Although a mean-field theory, Flory's result of $\nu = 3/5$ is in good agreement with results from renormalization group theory and computer simulations which yield $\nu = 0.588$.

3.3 Computational Methods

3.3.1 Monte Carlo Simulations

If the system size is large, and consequently the number of accessible states i is very large, it is not feasible to calculate the canonical partition function $\mathcal{Z}$. This makes it impossible to obtain exact macroscopic properties of the system which are determined by ensemble averages

$$\langle A \rangle = \sum_i A_i \frac{1}{\mathcal{Z}} e^{-\frac{E_i}{k_B T}} \tag{3.38}$$

A way to obtain estimates for macroscopic system properties is to generate a large number of random samples of the ensemble, i.e. samples which are distributed according to the probability distribution of the ensemble. These samples can then be taken as an approximation of the configuration space of the system. Ensemble averages are estimated by the mean values over the sampled configurations.

The Monte Carlo Method is a numerical simulation method that can be used to sample configurations of a statistical ensemble, in particular the canonical ensemble for which the distribution is the Boltzmann Distribution. It is used for a great variety of different problems from various fields including physics, chemistry but also biology and finance [152–159].

Markov Chains

The Monte Carlo Method is based on the concept of *Markov Chains* or *Markov Processes*. A Markov Chain is a stochastic process in which the subsequent state depends only on the present state but not on any previous states. It is possible to progress in the process by considering the transition probabilities π_{ij} for the transition of state i to state j. These transition probabilities are normalized, which means that the sum of transition probabilities from one certain state i to all other states k is a probability distribution

$$\sum_k \pi_{ik} = 1 \tag{3.39}$$

If the global probability flux for transitions from one state to another is conserved, then the system satisfies **Global Balance**

$$\sum_{i, i \neq j} \pi_{ij} p_i = \sum_{j, j \neq i} \pi_{ji} p_j \tag{3.40}$$

The system satisfies **Detailed Balance** if each single probability flux from state i to state j is equal to the probability flux from state j to state i.

$$\pi_{ij} p_i = \pi_{ji} p_j \qquad \bigvee i, j \tag{3.41}$$

Detailed Balance requires Global Balance. Such a Markov Process can be tuned to produce states which are distributed according to the Boltzmann Distribution. Therefore, sampling of a Markov Process can be an efficient way to obtain configurations of a system in the canonical ensemble.

The Metropolis-Hastings Algorithm

In 1953, Metropolis, Rosenbluth, Rosenbluth, Teller and Teller [160] suggested an algorithm in their famous publication to obtain random samples from a probability distribution. In this algorithm, Markov chains are utilized to generate the samples. Starting out with an arbitrary state, the algorithm suggests local modifications to the state to generate the new state in the Markov process.

Assuming detailed balance, the transition probability from the current to the new suggested state is given by the relative probabilities of both states. In the case of a canonical ensemble, this relation of the transition probabilities from state i to j and from j to i is given by

$$\frac{\pi_{ij}}{\pi_{ji}} = \exp\left(-\frac{E_j - E_i}{k_B T}\right) \tag{3.42}$$

Using the following transition probability for the transition from state i to j

$$\pi_{ij} = \min\left\{1, \exp\left(-\frac{E_j - E_i}{k_B T}\right)\right\} \tag{3.43}$$

the algorithm generates random samples which are Boltzmann distributed. This algorithm is know as the Metropolis-Hastings Algorithm and belongs to the Markov Chain Monte Carlo (MCMC) Methods.

Autocorrelation Time

In Markov Chain Monte Carlo Methods (MCMC), each state is generated from its previous state. The samples obtained from MCMC simulation can thus be highly correlated to each other. In order to obtain uncorrelated conformations from the simulation, the *autocorrelation time* has to be considered. Moreover a certain number of steps are required to reach thermal equilibrium from the start configuration, which is usually chosen randomly. The integrated autocorrelation time τ_{int} can be used to calculate the autocorrelation time. τ_{int} is determined by the autocorrelation function $C(t)$ and the normalized autocorrelation function $\rho(t)$ respectively. These functions measure the correlation of a certain observable for conformations which are separated by t Monte Carlo steps. Let $A(t)$ be an observable, then the unnormalized autocorrelation function for $A(t)$ is given by

$$C(t) \;=\; \langle\, A(s+t) \cdot A(s) \,\rangle_s \;-\; \langle\, A(s) \,\rangle_s^2 \tag{3.44}$$

and the normalized autocorrelation function is simply $\rho(t) \;=\; \frac{C(t)}{C(0)}$, where $\langle\cdot\rangle_s$ denotes the average over the ensemble at step s. For finite samples the average values can be estimated by mean values. The windowing procedure described in [161, 162] is a method that can be used to obtain an estimate of the integrated autocorrelation time

$$\tau_{int} \;=\; \frac{1}{2} \sum_{t=1}^{M} \rho(t). \tag{3.45}$$

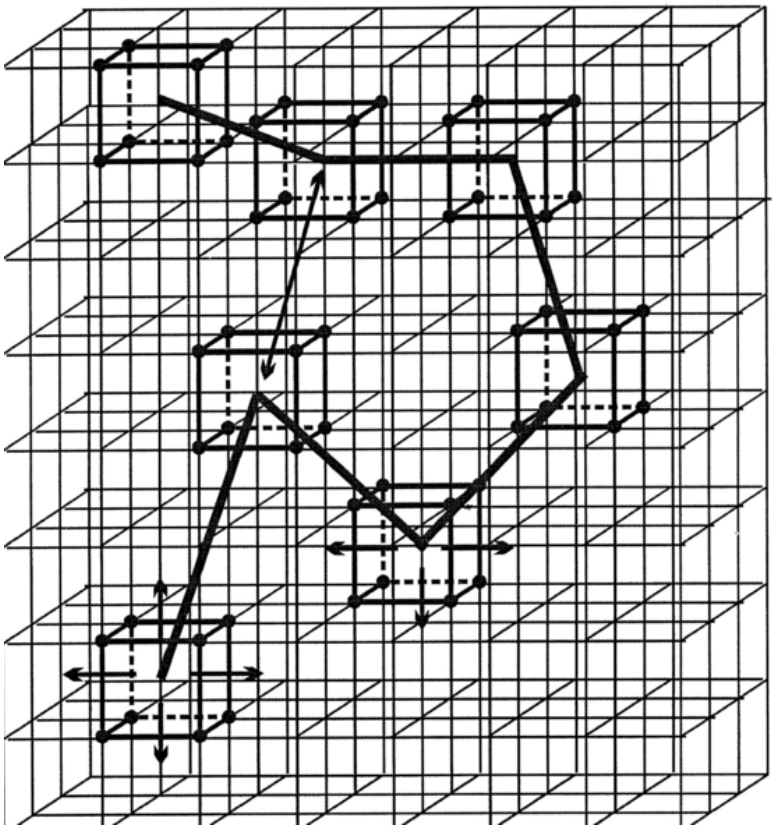

Figure 3.1: Illustration of the Bond Fluctuation Model. The Bond Fluctuation Model is implemented on a three dimensional cubic lattice. Each monomer occupies 8 lattice sites and is connected to its neighbors on the polymer via bond vectors. There are 108 different allowed bond vectors with lengths between 2 and $\sqrt{10}$.

The integer M is chosen such that $M > c \cdot \tau_{int}$ [162]. c can vary between 4 for exponential decaying $\rho(t)$ to 10 for slower decay [161, 162]. Two subsequent conformations are considered to be uncorrelated when more than $5\,\tau_{int}$ steps are between them. In the beginning of the simulation, $10\,\tau_{int}$ steps are considered to be enough for equilibration of the starting configuration.

The Bond Fluctuation Model

The Bond Fluctuation model (BFM) is a well established *lattice model* for polymers. The BFM incorporates excluded volume interactions and preservation of the topological state of the polymers by prohibiting bond crossings. It is especially well suited for high density polymer systems as it provides low rejection rates in Monte Carlo simulations. The BFM is ergodic, which is very important for Computer simulations, since it means that thermodynamic averages of the system can be calculated either from many parallel simulations or from a single simulation that runs for a long time or even a combination of both [163, 164].

The BFM is implemented on a cubic lattice. For simplicity, the lattice constant is set to the value 1. In the BFM, each monomer of the polymer is represented by a cube with side length 1 on the lattice. In other words, each monomer occupies 8 lattice sites, where the

front left corner serves as reference point of the monomer. In a polymer, each monomer is permanently linked to its neighbors by bond vectors. The bond vector between two monomers is the vector between the two reference points on the lattice. These bond vectors form a set $\{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_{N-1}\}$. To ensure excluded volume interactions between monomers, bond vectors have a minimum length of 2, as monomers could otherwise overlap. In order to a void bond crossings, there are further restrictions on the bond vectors, ensuring that the topological state of the polymer is preserved. There are in total 108 different bond vectors allowed. These vectors can be obtained by permutation of the base vectors $(2,0,0), (2,1,0), (2,2,0), (3,0,0)$ and $(3,1,0)$. The representation of monomers as a box and the possibility for different lengths of the bond vectors is a key element of the BFM. This allows monomers more ways to move than in a simple lattice model, where monomers occupy nodes on the lattice, and eventually faster equilibration of the model.

Sampling of the conformation space is done through updating of the respective current state by local moves of the monomers. A monomer is chosen randomly and a move by one lattice site in a random direction is suggested for this monomer. The move is principally allowed if the conformation of the polymer after this move does not violate any rules of the BFM. For instance the new position of the monomer cannot overlap with any other monomer and the new bond vectors have to be from the set of allowed bond vectors in the BFM.

3.3.2 Molecular Dynamics Simulations

A very important numerical method to study the behaviour of many-particle systems are *Molecular Dynamics Simulation*. The basic principle of Molecular Dynamics Simulations is to solve the equation of motion for a system of particles to uncover their dynamical behaviour. When the system is in thermodynamic equilibrium, Molecular Dynamics Simulations can be used to sample different states of the system, by following the trajectory of the system through phase space. According to the ergodic hypothesis, time averages over the system are equal to ensemble averages. Therefore, these states can then be used to calculate macroscopic thermodynamic quantities of the system [165].

Simulation at Conserved Energy

In the microcanonical ensemble, the total energy E of the system is conserved. The equation of motion of particle i with coordinates $\mathbf{x}_i$ is then given by the Newton's equation of motion

$$m_i \ddot{\mathbf{x}}_i = -\nabla U(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N) \tag{3.46}$$

where m_i is the particle mass and U is a potential function that depends on all particles in the system. This gives a set of coupled differential equations. They can be numerically integrated for example with the *Verlet algorithm* [166]. The potential U is often referred to as *forcefield*.

Langevin Dynamics

A system, in which the particles move according to Newton's equation of motion has conserved energy. But molecular systems are usually not completely isolated from their environment. Cells for example can exchange energy with the rest of the organism. Often, the temperature of the system is conserved instead of the energy, moving the system to

the canonical ensemble. To simulate many-particle system at constant temperature T, thermostats can be used. One of the most commonly used thermostats sit the *Langevin Thermostat*, which produces *Langevin Dynamics* within the system.

In Langevin dynamics, the particles are imagined to be embedded in a viscous solvent. Besides the potential U, two additional forces act on each particle: (1) a *Stokes* friction force which is proportional to the velocity of the particle and (2) a noise term drawn from a Gaussian distribution. The system of differential equations for an N-particle system with masses M and particle coordinates $X(t)$ is then given by the *Langevin Equation*:

$$M\ddot{X} = -\nabla U(X) - \gamma\dot{X} + \sqrt{2\gamma M k_B T}R(t) \tag{3.47}$$

with $R(t)$ being a *stationary Gaussian process* with zero mean and delta correlation that simulates Gaussian white noise for the system.

$$\langle R(t)\rangle = 0 \tag{3.48}$$
$$\langle R(t)R(t')\rangle = \delta(t-t') \tag{3.49}$$

Through these two terms, the system gains the ability to exchange energy with the implicit solvent. The system loses energy through the friction forces with the solvent. On the other hand, energy is added to the system by the Gaussian noise which models random collisions of the system particles with solvent particles. Gaussian white noise is a reasonable model to simulate noise in the system under the assumption that many forces contribute to it and that the microscopic fluctuations in the system decay fast compared to the macroscopic time scales. Integration of the Langevin equation for the N-particle system leads to dynamics that conserves the temperature of the system.

Chapter 4

Mitotic Chromosomes and Their Mechanical Properties

References

The content of this chapter was adapted from the following publication

- Y. Zhang and D.W. Heermann, *Loops determine the mechanical properties of mitotic chromosomes.*. PLoS ONE 6(12): e29225. doi:10.1371/journal.pone.0029225. 2011.

Chapter Summary

We introduce a new polymer model for mitotic chromosomes. The key assumption of the model is the ability of the chromatin fiber to cross-link to itself due to binding proteins. These protein-chromatin interactions are included by a probabilistic and dynamic mechanism. The hypothesis is motivated by the observation of high repulsive forces between ring polymers. We performed computer simulations to validate our model. Our results show that the presence of loops leads to a tight compaction and contributes significantly to the bending rigidity of chromosomes. Moreover, our qualitative prediction of the force elongation behaviour is close to experimental findings. The model presented here indicates that the internal structure of mitotic chromosomes is based on self-organization of the chromatin fiber rather than attachment of chromatin to a protein scaffold. It also shows that the number and size of loops have a strong influence on the mechanical properties. We suggest that changes of the mechanical characteristics of chromosomes for example in different stages of the cell cycle can be explained by an altered internal loop structure.

4.1 Introduction

During mitosis the dispersed interphase chromosomes undergo a transition into rigid, tightly compacted objects. This condensation mechanism and the inner structure of the chromosomes in this phase has been the target of many studies so far. The lengthwise compaction ratio of DNA in mammalian metaphase chromosomes is of the order of 10 000 - 20 000 [23]. On the lowest folding level, double-helical DNA wraps around histone octamers and thus forms nucleosomes about every 200 base pairs [54]. In a next step the coiling of this 10 nm thick beads-on-a-string fiber into a 30 nm thick filament was suggested [70, 72, 167]. These first two levels of folding account for a 40 fold compaction of the naked DNA. However, the existence of the 30 nm fiber is still under debate. The higher order folding motifs that are responsible for the remaining approx. 500 fold compaction still remain largely unknown [95, 96].

Many models have been put forward for the description of the chromatin structure in mitosis, including radial loop models, hierarchical folding models and network models [168]. In an early model Bak et al [169] suggested a helical folding of a 400 nm thick chromatin fiber. The composition of the chromosome of a thin fiber of 200-300 nm in diameter was also proposed by Sedat and Manuelidis [170]. Most textbooks feature the radial loop model which is based on histone-depletion experiments. It assumes the chromosome shape to be essentially governed by an axial non-histone protein scaffold to which chromatin loops are attached [97, 98]. Condensins and Topoisomerase II were found to be the main components of the protein core and are therefore the main candidates for the driving forces of the condensation [100, 101]. However the radial loop model has been put more and more into question. Different experiments report that Topoisomerase II and condensins are highly mobile within mitotic chromosomes [171, 172]. Kireeva et al [173] showed that axial staining of condensins cannot be seen until late prophase when considerable condensation has already taken place. Instead of an axial protein scaffold, the authors suggest a hierarchical folding of the chromatin fiber. These kind of models predict different folding levels from the 30 nm fiber into the $\approx 1\,\mu$m thick chromosome. Possible folding levels are a 100-130 nm fiber and subunits in the size of 250 nm [170, 174, 175].

Another approach to the analysis of mitotic chromosome structure are micromechanical manipulation experiments which target the elastic properties of chromosomes [43]. Human chromosomes and chromosomes from newt lung cells were found to be very elastic objects which can be stretched to several times of their native length [110, 111]. Houchmandzadeh and Dimitrov measured the bending rigidity and the stretching stiffness of single in vitro assembled chromatids from Xenopus egg extract. They found the chromatids to be very flexible objects with the persistence length being only a few times the thickness of the chromosomes. For small extensions the authors reported a linear force-elongation behaviour and reversible deformability. Furthermore chromosomes were extensible up to 100 times their native length with a force plateau being observed at relative extensions larger than ≈ 15 [112]. This kind of elastic response was also confirmed for chromosomes from newt lung cells. The chromosomes showed reversible extension up to three times of their native length. For intermediate extensions, hysteresis was observed and for long extensions beyond 30 times of the native length, the force-extension curve decreased to a plateau [113]. Further experiments by Poirier and Marko [117] targeted the force-relaxation behaviour of stretched chromosomes, especially during nuclease digestion. The authors concluded that mitotic chromosomes do not have a mechanically contiguous protein scaffold but rather proposed a network model, where the 30 nm chromatin fiber is cross-linked to itself by binding proteins [116]. While results on the stretching stiffness agree widely for chromo-

somes of different species, this is not the case for the bending rigidity. As described above, Xenopus egg extract chromatids were reported to be very flexible [96, 112]. However, experiments on in vitro and in vivo assembled chromosomes from other animals yielded much higher rigidities with persistence lengths being much larger than the length of the chromosomes themselves [114, 115]. Recent investigations showed that human mitotic chromosomes have very similar mechanical properties to mitotic newt chromosomes and thus likely smiliar structures, too [118].

All studies on mitotic chromosome structure indicate that chromatin loops play an important role in its organization. Loops can compact the chromatin fiber and be in part responsible for the mechanical properties. Especially the size and the number of loops were suggested to be closely connected to them [176, 177]. Moreover, there is evidence that looping of the chromatin fiber is crucial for chromosome compaction during all stages of the cell cycle. FISH experiments and new 3C/4C/5C/HiC experiments showed that loops of all length scales can be found in the interphase chromosome, possibly connected to transcriptional activity and genome function [73, 178–181]. The Random Loop Model and its further development, the Dynamic Loop Model, assume dynamic formation of loops on all length scales. They predict the confined folding of the chromosome without spatial constraints [182–184]. However, in contrast to the interphase, the chromatin fiber does not show long range interactions between distant segments in mitosis. Estimates for loop sizes here are in the range of 20 to 90 kb [97, 117]. Marko [185] pointed out that local coiling of a polymer along its length while long-range cross-linking is absent can be responsible for a lengthwise condensation.

In this work we investigate how the formation of loops can account for the condensation and mechanical properties of chromosomes during mitosis. Polymer rings have been found to repel each other much stronger than linear polymers [186]. Therefore, looping alone can already be responsible for a considerable stiffening up of the chromosome. However, our model does not impose an ordered structure on the chromatin fiber. Rather, cross-links and thus loops are formed upon collisions of fiber segments. Condensins and Topoisomerase II were suggested to be responsible for the cross-linking of chromatin during mitosis [23]. The probabilistic nature of our model, where loops form and dissolve dynamically, can account for the mobility of these proteins within the chromosome.

For the mitotic chromosomes we assume a restricted interaction range for the formation of cross-links in order to achieve a lengthwise compaction of the chromatin fiber. The restricted interaction range of the cross-links is motivated by the geometrical shape of mitotic chromosomes, which appear to be rod-like and thus very different to the more spherical shapes of chromosome territories in interphase. On the other hand, a strong compaction in length especially in eukaryotic mitotic chromosomes can be observed. Therefore, the model assumes a short-ranged folding of the chromatin fiber that results in a length-compaction and condensation of the fiber into rigid rods. This kind of folding also guarantees that genes are aligned linearly along the mitotic chromosome. Essentially, the coiling of the chromatin fiber can be seen as the folding of a thin fiber into a thick fiber.

Our results suggest that mechanical properties can be explained by self-organization of the chromatin fiber without the need of any axial protein scaffold. With this dynamic formation of loops, the resulting structure of the chromatin fiber is similar to a chromatin network. Moreover, our model can be seen in the context of a hierarchical folding model.

4.2 Methods

4.2.1 Dynamic Loop Model

The main idea of the model is that the tight condensation of the mitotic chromosome, which is presumably facilitated by condensin proteins and Topoisomerase II, can be modeled by a dynamic looping mechanism of the chromatin fiber. The model assumes, that genomically distant sections of the chromatin fiber can cross-link for a fixed amount of time when they come into physical proximity of each other. This self-tethering mechanism mimics the dynamics of binding proteins such as Topoisomerase II and condensins that have been found to be significant for metaphase chromosome structure. Although the exact role of Topoisomerase II and condensins in mitotic chromosomes is still unclear, it is ascertained that they are able to bind to chromatin and to cross-link the fiber [177]. However, the important element in the model is the probabilistic nature of the cross-linking mechanism. Rather than being a fixed structure, the organization of the fiber is dynamic. This accounts for the fact that proteins in the surrounding solvent of the chromatin fiber are mobile. Therefore also the binding sites are subject to fluctuations in space and time, which mirrors the effect of protein concentration and binding affinity.

4.2.2 Monte Carlo Simulations

The behaviour of the chromatin fiber is simulated using a lattice Monte Carlo algorithm based on the well-established Bond Fluctuation Model (BFM), which incorporates excluded volume interactions and preservation of the topological state of the polymer [163]. The Monte Carlo algorithm for the Dynamic Loop Model consists of two main steps. In the first step, local moves for the single monomers are proposed and accepted if the constraints of the bond vectors are not violated. These local moves make sure that the algorithm produces correct Rouse dynamics for the polymer [164]. The key feature of the Dynamic Loop Model is the ability of the fiber to cross-link with itself, which is comprised in the second step. When two fiber segments come into the proximity of each other by diffusion, there is a certain probability κ that they form an additional bond between each other and thus a loop in the chromatin fiber. The size of the loops, i.e. the contour length between the bound fiber segments, is restricted by a maximum length C. The loop also has a restricted lifetime which is drawn from a Poisson distribution with mean value τ. After this lifetime, the cross-link between the fiber segment dissolves.

Conformations were sampled from the equilibrium distribution using the Monte Carlo algorithm described above. The algorithm sweeps the space of possible conformations with equal probability. Simulations were performed on a lattice with periodic boundary conditions so there was no spatial confinement of the polymer. As in all Metropolis Monte Carlo algorithms, subsequent conformations are highly correlated to each other. The integrated autocorrelation time τ_{int} was used in this work to calculate the autocorrelation time. Let $A(t)$ be an observable, then the normalized autocorrelation function $\rho(t)$ for $A(t)$ is given by

$$\rho(t) \;=\; \frac{C(t)}{C(0)} \tag{4.1}$$

where $C(t)$ denotes the unnormalized autocorrelation function

$$C(t) \;=\; \langle\, A(s+t) \cdot A(s) \,\rangle_s \;-\; \langle\, A(s) \,\rangle_s^2. \tag{4.2}$$

We used the windowing procedure to estimate the integrated autocorrelation time τ_{int} [161]:

$$\tau_{int} \;=\; \frac{1}{2}\sum_{t=1}^{M}\rho(t). \tag{4.3}$$

In this study we used $M = 10\tau_{int}$ for all simulation runs. After $5\,\tau_{int}$ Monte Carlo steps we considered conformations as being independent. In our simulations, two initialization stages were run. In the first stage the looping mechanism was still turned off, the fiber thus equilibrated from the start configuration to a self-avoiding walk. In the second stage the looping mechanism was switched on and $10\,\tau_{int}$ Monte Carlo steps were performed to reach equilibrium for the model chromosomes.

Polymer chains consisted of N monomers. Simulations with various values for N between 400 and 800 were performed. To properly investigate the influence of the cutoff length C and the number of loops per chain length k_p, we conducted runs with varying parameters. C took values between 20 and 80 and for k_p values between 0.5 and 1.4 were considered. Typically, the autocorrelation time for a self-avoiding walk scales with the square of the number of statistical segments N. With the high densities in our model chromosomes and consequently the high rejection rates, this made it hard to model longer chains. The exact autocorrelation times in the simulations were also depending on the cutoff length and mean loop concentration. As an example, the autocorrelation time for a configuration with $N = 650, C = 30$ and $k_p = 1.22$ was about $1.7 \cdot 10^8$ MC steps. Around 5000 independent conformations could be sampled in 96 hours by parallel simulations running on 64 processors.

For the simulation of the stretching of model chromatids a force F was included via a potential U_{pull}. The direction of the force is parallel to the end-to-end vector of the model chromatid, so the fiber can move without spatial constraints. U_{pull} is given by

$$U_{pull} \;=\; F \cdot |\mathbf{x}_N - \mathbf{x}_1| \tag{4.4}$$

The force F is a parameter in the simulations. The potential U_{pull} has then the effect that local moves of one of the end monomers, which increase the end-to-end distance, are only accepted with probability $e^{-F\cdot\Delta R_e}$. In the stretching simulations the polymers are also in thermal equilibrium. The mean relative extension at force F is given by

$$\epsilon^F \;=\; \frac{\langle R_e^F\rangle - \langle R_e^0\rangle}{\langle R_e^0\rangle} \;=\; \frac{\langle \Delta R_e\rangle}{\langle R_e\rangle} \tag{4.5}$$

Here $\langle R_e^0\rangle$ denotes the mean end-to-end distance without any pulling force and $\langle R_e^F\rangle$ is the mean end-to-end distance for a configuration with forces F.

4.2.3 Bending Rigidity and Persistence Length

Long polymers usually have bending rigidities that limit their flexibility. While for simple models such as the ideal chain or the Gaussian coil the mean correlation between the bond vectors is $\langle \mathbf{b_i} \cdot \mathbf{b_j}\rangle = 0,\ i \neq j$, for real polymers this correlation is non-zero. To quantitatively describe the flexibility of polymers the directional correlation of different segments of the polymer can be used. Let $\mathbf{u}(s)$ be the direction of a chain segment at the contour length s. Then the correlation function between two segments separated by the contour length s' is

$$\langle \cos\theta(s') \rangle \;=\; \langle \mathbf{u}(s) \cdot \mathbf{u}(s+s') \rangle \tag{4.6}$$

The averaging is done over both, all positions s within one conformation and the ensemble of all conformations in thermal equilibrium. A quantity that measures the stiffness of the chain with respect to the orientational correlation is the persistence length ξ_p. It is defined as the integral width of the correlation function [148]

$$\xi_p \;=\; \int_0^\infty \langle \cos\theta(s') \rangle ds' \tag{4.7}$$

It can be shown that the persistence length is proportional to the bending modulus B [150]

$$B \;=\; k_B T \xi_p \tag{4.8}$$

The bending modulus can be described as the quantity which determines how much force is necessary to bend a segment of the chain to a certain curvature. In classical elsticity theory, the bending modulus is connected to Young's modulus Y which determines the elastical behaviour of a material. In the case of a homogeneous cylinder with radius R, the relationship is

$$B \;=\; \frac{\pi}{4} Y R^4 \tag{4.9}$$

4.2.4 Estimating Backbones, Directional Correlation and Radial Density

In this work the bending stiffness of the model chromatids were estimated via the directional correlation of segments of the chromatids. For this, for each conformation a backbone which represents the alignment of the model chromatid was calculated. The polymer chain which is given by the position vectors of the monomers $\mathbf{r}_1 \ldots, \mathbf{r}_N$ was divided into N_k segments of k monomers each. The center of masses of these segments $\mathbf{r}_1^k, \ldots, \mathbf{r}_{N_k}^k$ then represented a new, coarse-grained chain that approximated the alignment of the coiled chromatin fiber (Figure 4.3).

With the imaginary backbone the mean directional correlation between distant segments of the model chromatids were determined. Let $\mathbf{b}_1^k \ldots, \mathbf{b}_{N_{sec}-1}^k$ denote the bond vectors of the backbone, where $\mathbf{b}_i^k = \mathbf{r}_{i+1}^k - \mathbf{r}_i^k$. Then the mean directional correlation $\langle \cos\theta(s) \rangle_{conf}$ between segments that are separated by the arc length s for one single conformation is given by

$$\langle \cos\theta(s) \rangle_{conf} = \frac{1}{k} \sum_{j=1}^{N_{sec}-s} \frac{\mathbf{b}_j^k \cdot \mathbf{b}_{j+s}^k}{|\mathbf{b}_j^k| \cdot |\mathbf{b}_{j+s}^k|} \tag{4.10}$$

Let M be the sample size of the Monte-Carlo simulation, then the thermal average was estimated by

$$\langle \cos\theta(s) \rangle_{thermal} = \frac{1}{M} \sum_{m=1}^{M} \langle \cos\theta(s) \rangle_{conf_m} \tag{4.11}$$

The directional correlation was used to test if the coarse graining level k for the calculation of the backbone was chosen correctly. In this case the directional correlation showed a exponential decaying behaviour, whereas for coarse graining levels that were too low or too high, the behaviour would be non-exponential.

The backbone that is determined using the method described above is a coarse grained backbone and suitable to analyse the mean directional correlation between segments of the chromatid. However, for the analysis of other properties such as the thickness of the chromatid, a less coarse grained backbone is more advantageous. Such a nearly smooth backbone is obtained using a similar method. A coarse graining level k is selected again and the backbone is build in the following way: the first point of the backbone chain is the center of mass of monomers $1, 2, ..., k$, the second point of the backbone is the center of mass of monomers $2, 3, ..., k + 1$ et cetera. Thus, a chain is obtained where the beads are spatially very close to each other and the backbone can be considered as a smooth trajectory. These backbones were used to estimate the length and the thickness of the chromatid by calculating the mean radial density perpendicular to the backbone. The chromosome radial thickness r_d was estimated as the distance for which 90 % of all monomers were aligned closer to the backbone than this distance.

To estimate the thickness of model chromatids under an external force, backbones were calculated using the same coarse graining level as for the model chromatids without stretching force. As the total number of loops in the linear elongation region only changes marginally, it is justified to assume that the same coarse graining level yields the correct backbone.

4.3 Results

4.3.1 Dynamic Loop Model for Mitotic Chromosomes

Condensins and Topoisomerase II are presumably the proteins that establish cross-links of chromatin in mitotic chromosomes [100, 101]. Christensen et al [171] found Topoisomerase II to be mobile in human mitotic chromosomes. High mobility of condensin I in Drosophila metaphase chromosomes was reported by Oliveira et al [172]. Hence, the loop structure of mitotic chromosomes is not fixed but rather subject to fluctuations in loop sizes and positions of the loops. However, the complex dynamics of proteins and their interactions with the chromatin fiber are too complicated to be modeled in detail on the scale of a complete chromatid. Coarse grained approaches are therefore used to model the chromatids.

We present a model, where the cross-linking due to Topoisomerase II and condensins is incorporated by a dynamical looping mechanism of the fiber. This mechanism consists of the ability of distant fiber segments to form cross-links when they come into physical proximity of each other by diffusion. The shape of mitotic chromosomes is rod-like, as opposed to the more spherical shape of interphase chromosomes. Bohn et al [184] have shown before that long-range interactions unevitably lead to spherical shaped objects if the number of cross-links is high, which it has to be in the case of highly condensed mitotic chromosomes. On the other hand, short ranged interactions and the lack of long-range interactions at the same time were discussed to be responsible for a lengthwise condensation of the chromatin fiber [185]. Therefore, in this model for mitotic chromosomes, we included a restriction for the interaction range of the chromatin fiber and thus for the maximum loop size. Below this limit, all loop sizes are equally possible.

Such a lengthwise condensation also accounts for the fact that genes are aligned linearly in mitotic chromosomes, whereas long-range interactions can easily lead to mixing of distant chromosomal parts, bringing genomically distant genes into physical proximity of each other. A restriction of the length of loops is also consistent with experimental observations which do not give any indications for the existence of long-range interactions

in mitotic chromosomes. Consequently, all proposed models for the folding of mitotic chromosomes implicitly include restricted interaction ranges with the estimates for loop sizes ranging from 20 kb to 90 kb.

Additionally, in order to mimic the dynamics and mobility of the involved proteins, the cross-links have limited lifetimes, after which they dissolve again. The two important parameters of the model are therefore the restriction on the interaction range, the cutoff length C which determines the maximum size of loops, and the number of loops divided by the number of statistical segments that the chromatin fiber consists of, the loop concentration k_p. A value of $C = 50$ means that segments can only form cross-links if they are separated by no more than 50 statistical segments. A value of $k_p = 1$ means that there is on average one cross-link and thus one loop per statistical segment.

The dynamic crosslinking mechanism in our model implicitly mimicks the presence of binding proteins like condensins in the surrounding solvent of the chromatin fiber. The cutoff length for the size of loops accounts for the fact that long-range interactions are not present in mitotic chromosomes. However, we do not wish to imply that proteins like condensins would not be able to crosslink two chromatin regions that are genomically far away. Our presented model does not include explicit binding mechanisms of condensins and other proteins to chromatin. Instead, we suggest that condensins and Topoisomerase II cause a local folding up of the chromatin fiber leading to the formation of a thicker and strongly compacted filament. The emerging chromatin-solvent interfase following this compaction could then be a reason for the prohibition of long range interactions within mitotic chromosomes. The cutoff length is meant to account implicitly for such phenomenons which prevent the formation of large loops.

One of the major questions addressed in this work is whether this dynamic looping mechanism could be responsible for the condensation of the chromatin fiber into the mitotic chromosome. Furthermore, it was shown experimentally that mitotic chromosomes have high bending rigidities while still being very elastic. The other important question that is addressed here is to which extent entropic effects invoked by the formation of chromatin loops can account for the mechanical properties of mitotic chromosomes.

4.3.2 Dynamic Looping Mechanism Promotes Condensation into Rigid Objects

We analysed the shape and morphology of the model chromatids. In Figures 4.1 and 4.2 typical conformations for different parameter configurations can be seen. When no loops are present, the fiber behaves like an ordinary self-avoiding walk. For low loop concentrations, cross-links at different positions along the chromatin fiber are formed. In these regions a compaction and a formation of blobs can be observed. These blobs are connected by fiber sections with no loops. However, when the settings for the looping probability and the mean loop lifetime are increased, the chromatin fiber condenses into a thicker, rod-like filament. The structure then resembles a flexible rod and is homogeneous along its contour. Thus, for high loop concentrations, the Dynamic Loop Model produces coiled fibers with a strong resemblance to mitotic chromatids.

We used coarse grained polymer chains consisting of up to 800 segments to represent the underlying chromatin fiber. Assuming a DNA content of approx. $10^8\,bp$ in the chromosome we obtain the size of one statistical segment to be approx. $125\,kb$. We investigated configurations with loop concentrations up to $k_p = 1.4$. We observed that a value of at least $k_p \approx 0.9$ is needed for condensation of the fiber into a rod that resembles a mitotic chromatid. A value of $k_p = 0.9$, which means on average 0.9 cross-links per statistical

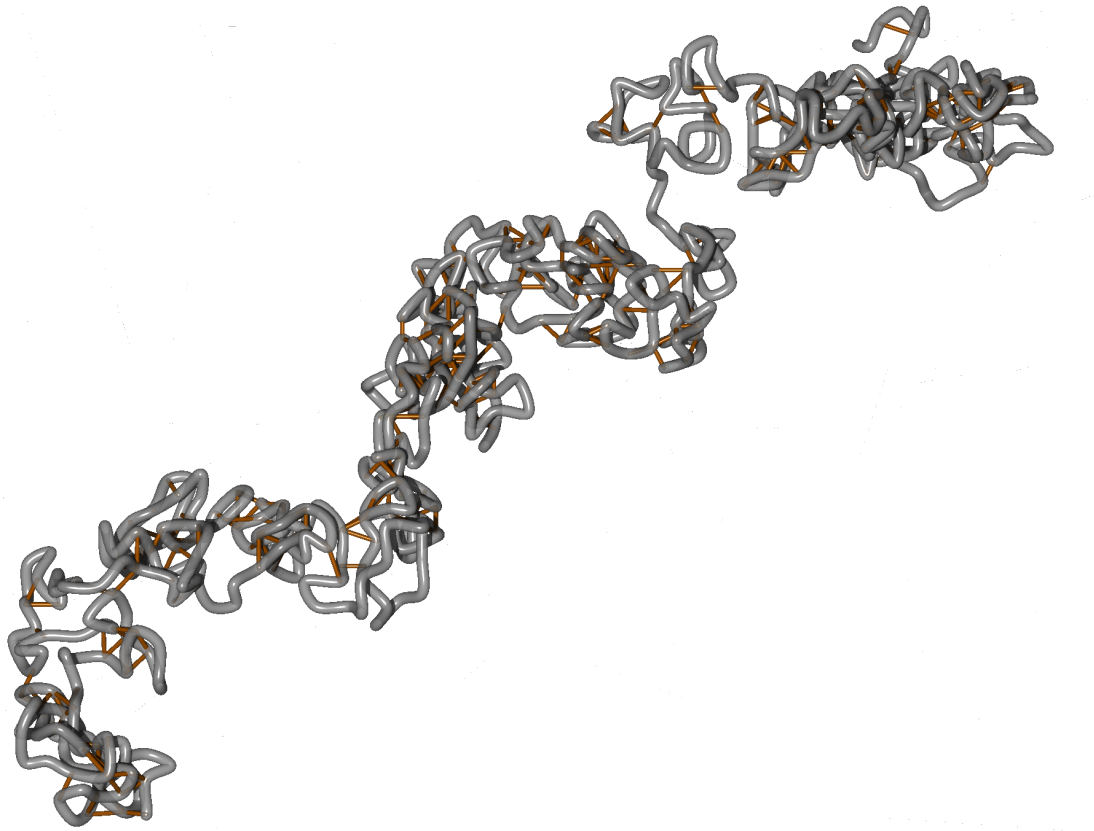

Figure 4.1: Example Conformation. The grey tube represents the chromatin fiber. The orange sticks visualize the cross-links between distant fiber segments. Each chain embodies a single chromatid. The chromatin fiber in both examples consisted of $N = 650$ segments. The cutoff length for the loop sizes is $C = 50$. It means that fiber segments which are separated by a genomic distance greater than 50 monomers cannot form an additional bond. **A**. In this example, the mean loop concentration is $k_p = 0.7$. At these low loop concentrations the conformations are non-homogeneous. Rather, a2 formation of blobs can be observed in regions with many cross-links. These regions are connected by fiber section with no or only few loops.

segment, would thus correspond to one cross-link every approx. $140\,kb$. Likewise, $k_p = 1.4$ would mean one cross-link every approx. $90\,kb$.

In order to analyse the shape and mechanical properties we calculate backbones which represent the alignment of the model chromatids. Each backbone can be seen as a coarse grained polymer that describes the large scale properties of the model chromatid without the details of the coiling on the local scale. Figure 4.3 illustrates this fact. The backbones are used to estimate the geometrical properties and the directional correlation between different segments of the chromatids. The mean chromatid lengths are calculated and compared for different settings of cutoff length C and loop concentration k_p. When the maximum loop size C is increased, the length of the rod decreases as large loops condense the fiber more efficiently than small loops. The compaction is also tighter when the mean loop concentration k_p is higher. Therefore, the chromatid length decreases with higher loop concentrations, too.

The estimation of the chromatid thickness involves the calculation of the average ra-

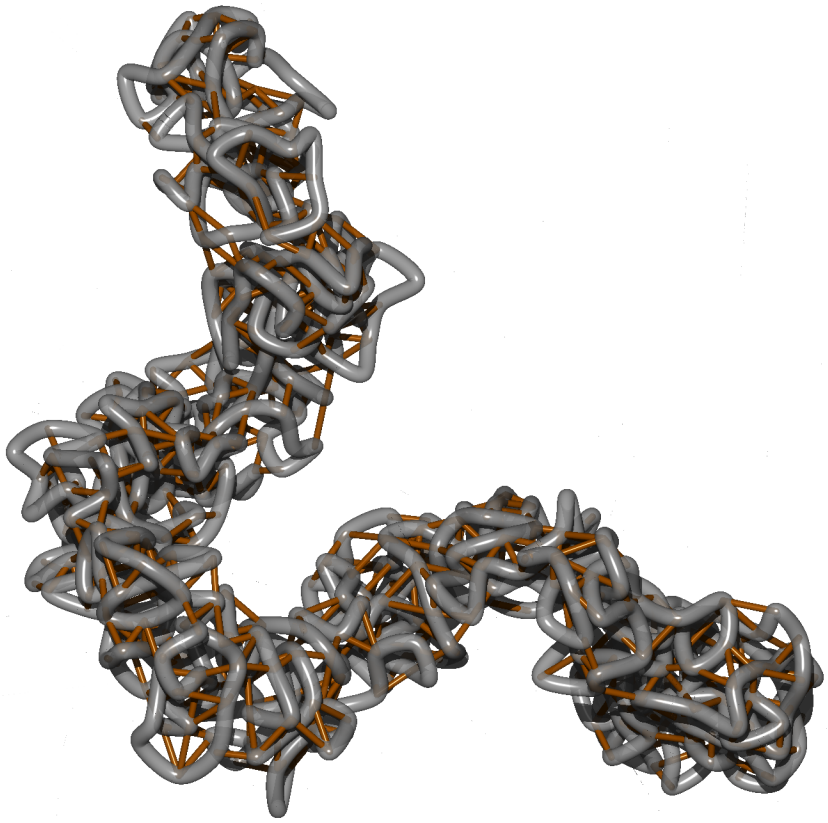

Figure 4.2: Example Conformation. When the loop concentration is high enough, a condensation of the chromatin fiber into thick, homogeneous rods can be observed. In this configuration the loop concentration is $k_p = 1.2$. Cross-links are distributed homogeneously along the chain.

dial monomer density functions. Radial in this case means perpendicular to the calculated backbone. Figures 4.4 and 4.5 show radial monomer density functions and their dependency on the cutoff length C and the loop concentration k_p. At the central axis the density has a minimum, but increases quickly with the distance from the backbone, until a plateau area is reached. This broad plateau is then followed by the expected decay for large distances. The drop off at the backbone indicates that the fiber tends to form rings around a central axis and is roughly organized in a helical-like structure. However the organization is much more complex than a simple helix. The broad plateau region suggests that the chromatid is radially homogeneous on a large scale.

Next we look at the dependency of the density and the width of the model chromatids on the parameters C and k_p. As expected, larger cutoff lengths result in a more extended plateau and thus thicker chromatids. On the other hand the number of cross-links has only little influence on the thickness. For the same cutoff length, conformations with more cross-links only yield higher monomer densities but have the same widths as conformations with fewer cross-links. The dependency of the chromosome thickness on the cutoff length C is displayed in Figure 4.6. A linear behaviour can be observed.

We used many different parameter settings (C, k_p) to investigate how they influence the geometry of the resulting model chromosome. To illustrate that the model chromosomes do match the geometry of real chromosomes, for example human metaphase chromosomes, we consider the result of a setup with $N = 650$ statistical segments, a cutoff length $C = 50$

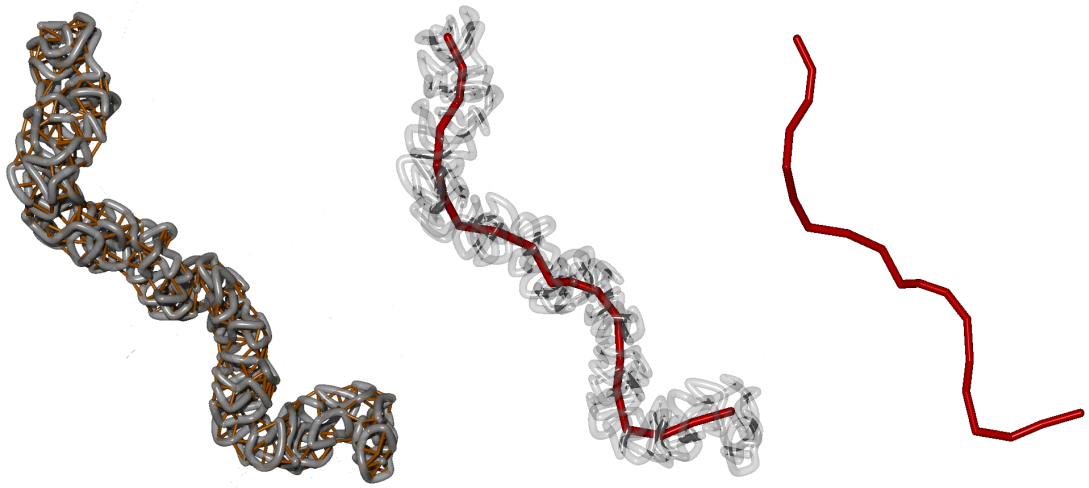

Figure 4.3: Visualization of the backbone for a model chromatid. In order to analyse the shape and the mechanical properties of the condensed rods, it is necessary to calculate backbones which represent the alignment of the model chromatids. These backbones are obtained by a coarse graining method that is applied to each single conformation. The original polymer chain is divided into N_{sec} sections that contain k statistical segments each and the center of masses of each section is calculated. The new chain consists of the center of masses of the sections. The degree of coarse graining, characterized by the parameter k is an important parameter in this method. It has to be chosen correctly in order to guarantee that the backbone truly represents the alignment of the chromatid.

and a loop concentration of $k_p = 1.37$. Assuming a lattice constant of $0.05\,\mu m$, the mean length of the model chromatid would then be $4.75\,\mu m$ and the thickness would be $1.07\,\mu m$, corresponding to a cross-section of $0.90\,\mu m^2$. This example demonstrates that our model chromosomes have indeed the dimensions of real chromosomes.

Our results for the geometry of the model chromatids show that the Dynamic Loop Model covers a broad range of different geometries, depending on the parameters C and k_p. When the maximum loop size is higher, the chromatid is thicker and shorter. On the other hand, the mean loop concentration influences the compaction and length of the chromatid but not the thickness. Therefore, the loop structure, the number and size of loops, obviously plays an important role for the shape of chromosomes. In Table I an overview of results on the length and width of model chromatids is given. The length-to-width ratios match those of natural mitotic chromosomes in different stages of mitosis.

4.3.3 Presence of Loops Enhances the Bending Rigidity due to Entropic Repulsion

The analysis of mechanical properties, especially the bending rigidity and the elastic response have been important parts in the experimental examination of mitotic chromosomes. Therefore, we analyse results from our model and from models without loops such as the self-avoiding walk with respect to the directional correlation using the calculated backbones. Figure 4.8 shows a comparison between the mean directional correlation of both models when the same degree of coarse graining for the calculation of the backbone is

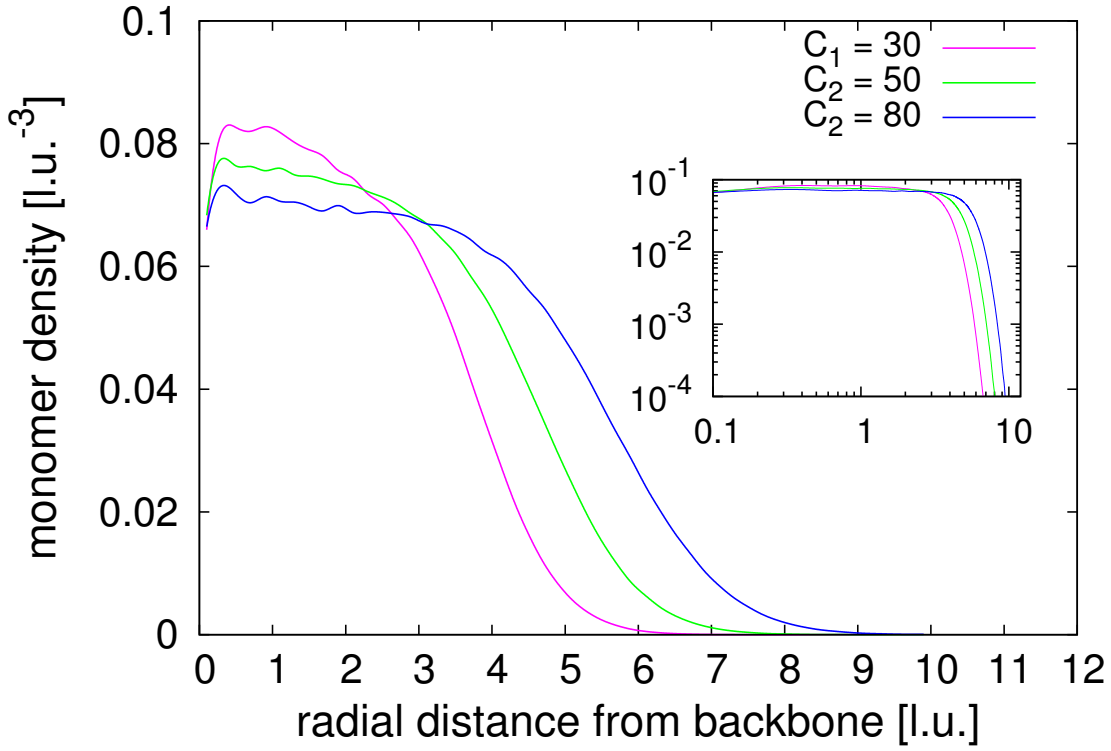

Figure 4.4: Mean radial distance function $\rho(r)$. The Figure shows $\rho(r)$ for conformations with the same mean loop concentration but different cutoff lengths. Larger cutoff sizes give thicker model chromatids but smaller densities and lengthwise compaction ratios.

used. The directional correlation function for the Dynamic Loop Model shows an exponentially decaying relationship with the separating genomic distance. Therefore the backbone of the model chromatid behaves like a worm-like chain on this length scale. This result is consistent with experimental findings of Houchmandzadeh and Dimitrov [112] who found chromatids from in vitro assembled Xenopus laevis egg extract to show an exponentially decaying mean directional correlation for one order of magnitude. Furthermore, when compared to a self-avoiding walk, the Dynamic Loop Model polymer has a much higher bending rigidity. This is a very important finding as it shows that simply the existence of loops enhances the bending rigidity of the chromatin fiber. The entropic repulsion between polymer rings is responsible for this observation. In the presence of a large number of rings within the chain as in the case of the Dynamic Loop Model, bending of the chromatid will reduce the distance between closely aligned loops. Hence the energy required to bend the chromatid is higher than in the case where no loops are present, leading to an enhanced stiffness of the filament.

However, it should be noted that already the calculation of a backbone for the self-avoiding walk is not meaningful since the conformations of self-avoiding walks do not have the shape of mitotic chromosomes. On the other hand, the rescaling for the Dynamic Loop Model generates worm-like backbones which truly represent the overall alignment of the model chromatid.

To evaluate the persistence length it is more convenient to measure the distances within the backbone in lattice units rather than genomic distance of the original chain. This was

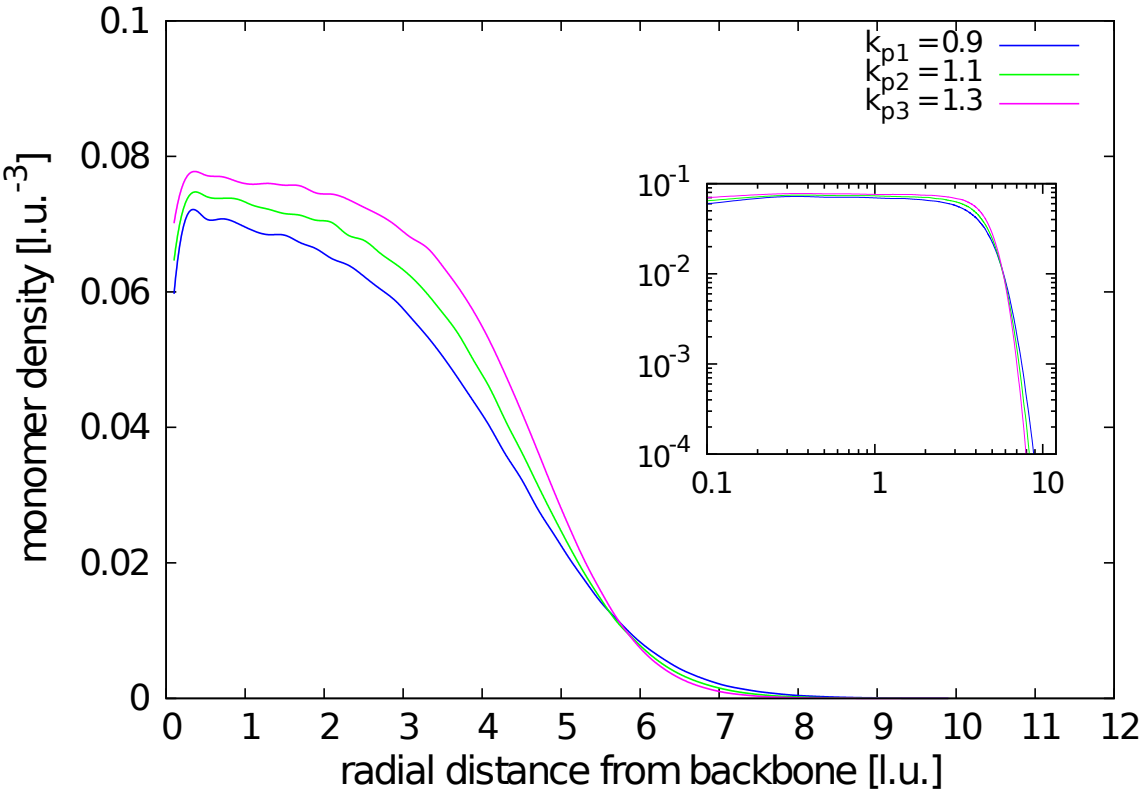

Figure 4.5: Mean radial distance function $\rho(r)$**.** The Figure shows $\rho(r)$ for conformations with the same cutoff length but different mean loop concentrations. The number of loops has only little influence on the chromatid thickness. However, chromatids with more loops show tighter compaction ratios and hence higher monomer densities.

done by taking the average distance d_m between adjacent beads of the backbone. The contour length of the backbone is then simply $d_m \cdot c$ with c being the contour length in units of center of mass points.

Simulations with different cutoff lengths and loop concentrations show that the bending rigidity is very sensitive to both parameters. Directional correlation functions for different values of cutoff length and loop concentration can be seen in Figures 4.9 and 4.10. Increasing the cutoff length results in a higher mean loop size. This in turn leads to chromatids with larger thickness and thus reduced flexibility of the filament. For a homogeneous cylinder, the bending rigidity is proportional to the fourth power of the radius [150]. Furthermore, large loops within the chromatin fiber tie parts of the fiber together which would normally be farther apart. This tightening also contributes to the enhanced stiffness. Consequently, the flexibility of a chromatid with higher cutoff length is reduced compared to a chromatid with smaller cutoff length. The bending rigidity is also influenced by the number of loops within the polymer. Higher mean loop concentrations k_p are associated with stronger compactions of the chromatids. Hence, the loops or chromatin rings are spaced closer to each other increasing the entropic repulsive forces between them. Bending becomes therefore even more energy consuming.

The persistence lengths for model chromatids with different parameters are shown in Figure 4.11. Clearly the bending rigidity increases with the number of loops in the chain and also with higher cutoff lengths. We find typicalvalues of the persistence length to

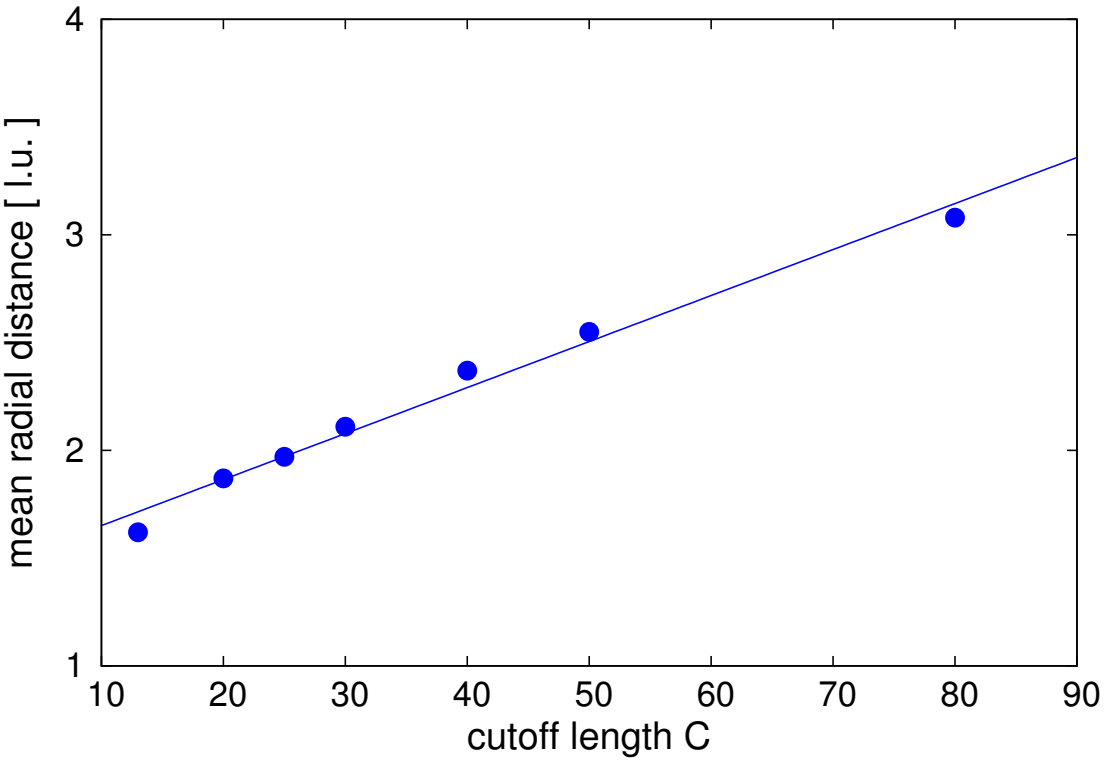

Figure 4.6: Influence of the parameter settings on the spatial dimensions. A linear relationship between the average distance of monomers from the backbone and the cutoff length is found. Together with the observed drop off of the monomer density at the central axis we conclude that the fiber coils around the backbone in a helical-like folding manner. However this gives just a general tendency and the exact structure is much more complicated.

be in the range between 1.5 to 4 times the diameter of the chromatid, depending on the parameter settings. This is in good agreement with results on mitotic chromatids from Xenopus egg extract [112].

4.3.4 Variations in Size and Number of Loops Evoke Different Elastic Responses

Measuring the elastic response is one possibility to study the internal structure of mitotic chromosomes. Simulations of model chromatids under an external pulling force are done to examine their elasticity. In the pulling simulations, model chromosomes are first subjected to the Monte Carlo algorithm of the Dynamic Loop Model until they are fully condensed. Then a constant pulling force $\mathbf{F}$ is applied to the chromatid ends directed along the end-to-end vector. A corresponding pulling energy U_{pull} is added to the energy of the conformation. Conformations are then sampled from the equilibrium distribution including the additional pulling potential. Thus, the pulling can be viewed as adiabatic. For fixed parameter sets, we analyse the mean end-to-end distances of the model chromatids at different pulling forces and calculate the mean relative extensions ϵ.

We have to point out that in our coarse grained model no additional potential between the segments exist apart from the dynamic cross-linking mechanism. Furthermore, the

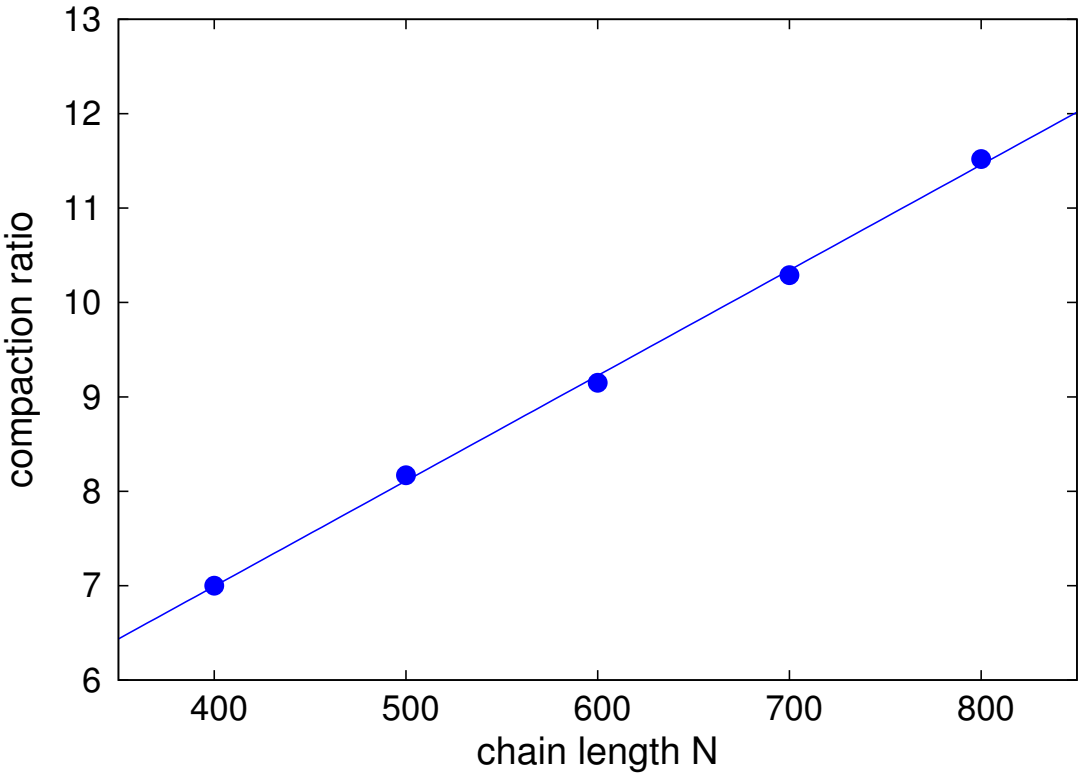

Figure 4.7: Influence of the parameter settings on the spatial dimensions. Interestingly we can see that at constant values for C/N the lengthwise compaction ratio still increases linearly with the chain length N. In this example $C/N = 0.05$ and the mean loop concentration is $k_p = 1.0$. When extrapolating the linear curve to a compaction ratio of 500-fold, a polymer with $N \approx 24\,000$ statistical segments and cutoff length $C \approx 1200$ would be needed.

stretching in the model is performed in equilibrium. The stretching forces in our model are therefore much smaller than the forces that are measured in experiments. Assuming a lattice constant of $0.05\,\mu m$ in our model, the forces are in the range of $10^{-15}\,N$ which is several orders of magnitude smaller than forces measured in micromechanical experiments on mitotic chromosomes. However, using this kind of coarse grained modeling it is possible to make qualitative tests and predictions.

The elastic response of the model chromatids shows different domains. Furthermore, the size and number of loops play a crucial role for the elasticity. Force elongation curves for different settings of cutoff length C and loop concentration k_p can be seen in Figures 4.13 and 4.14. For relative extensions of up to twice the native length of the model chromatid, we observe a linear relationship between force and relative extension. This means that in this region the chromosome has the elasticity of a homogeneous, elastic material. Such a behaviour of mitotic chromosomes was found in numerous experiments [112, 113, 118]. In this region no significant decrease in the total number of loops can be seen. The chromatid is stretched but its looping mechanism still efficiently cross-links different chromatin segments. However, the analysis of the loop size distribution shows that although the total number of loops does not change, there is a reorganization of the loop domains. The number of small loops increases while the number of large loops decreases, hence there is a shift from large to small loops. Apparently, when only small

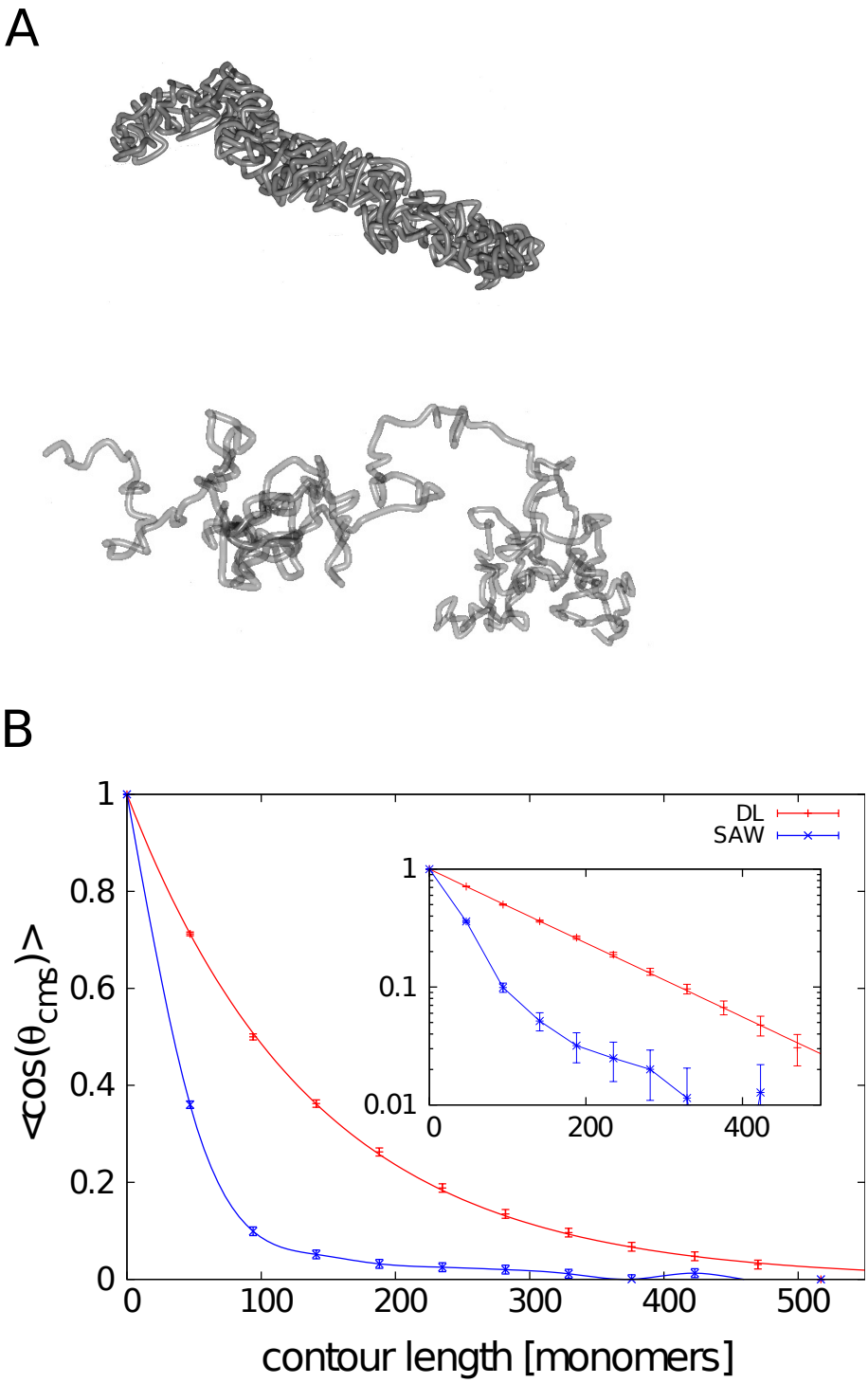

Figure 4.8: Comparison between Dynamic Loop Model and self-avoiding walk. A. The upper conformation is a Dynamic Loop Model chromatid with $N = 650$, cutoff size $C = 50$ and mean loop concentration $k_p = 1.4$. For comparison, a conformation without loops with the same chain length $N = 650$ is shown below. **B.** For both, self-avoiding walk and Dynamic Loop Model the coarse-graining method is applied and the directional correlation is calculated. The same degree of coarse-graining is used for both models. The figure shows an exponential decay of the directional correlation function of the Dynamic Loop Model, while the the self-avoiding walk does not show this behaviour. Most importantly, the Dynamic Loop Model chromatid is much stiffer than the self-avoiding walk. This shows that the entropic repulsion of the chromatin loops that are generated by the cross-linking mechanism leads to a considerable stiffening up. Error bars represent the standard error of the sampled conformations.

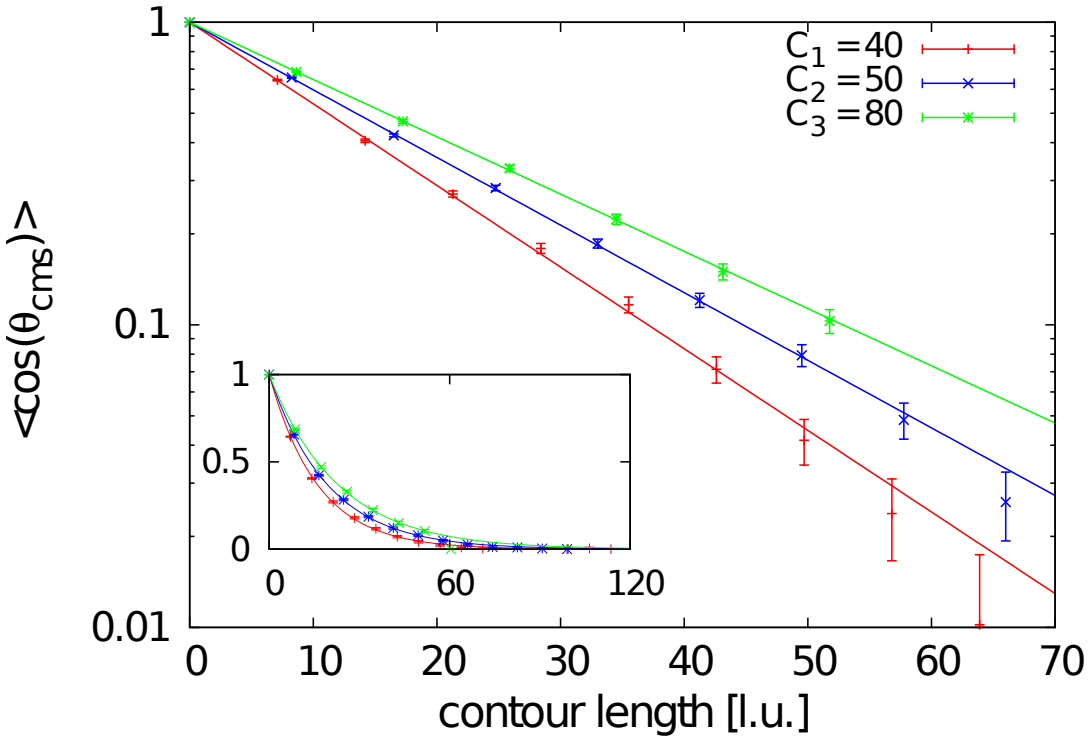

Figure 4.9: Directional correlation functions for model chromatids with different cutoff lengths. The chain length for all configurations is $N = 650$. Error bars represent the standard error. For fixed mean loop concentration $k_p = 1.0$ we can see that the stiffness increases with the cutoff length C. Larger cutoff lengths result in thicker chromosomes and in turn less flexibility.

forces are applied, the formation of larger loops is inhibited because distant chromatin segments are pulled apart. However on the local scale the looping mechanism is still intact and thus more small loops are formed, keeping the total number of loops constant.

At larger pulling forces the loop formation on all scales is inhibited. The formation of large loops is still obstructed stronger than the formation of small ones. The force extension curve is very flat in this region and resembles a plateau. At this strength, the applied force destroys the cross-linked structure of the polymer. The region of $\epsilon \approx 2$ to approx. 20 times elongation of the native length can be characterized as the decondensation domain of the model chromatid. In this domain the internal structure of tightly condensed loops is destroyed. The slope of the force-extension curve is very low in this region and a small increase of the force results in a vast stretching of the chromatid. Such force plateaus were also found in the experiments, although at larger relative extensions [111, 113].

We evaluate the slope of the force extension curves in the linear domain from relative extensions of $\epsilon = 0$ to $\epsilon = 2$ for several different parameter sets. The slope is then used to calculate Young's modulus that characterizes the elasticity of an elastic material. Almagro et al [96] showed that mitotic chromosomes do not have a homogeneous elasticity but that rather different segments show different elastic moduli. Furthermore the elastic behaviour of mitotic chromosomes can be changed by exogenously added agents, such as trypsin, proteinase K or Topoisomerase I and II [187, 188]. These alterations of the mechanical properties were suggested to be related to changes in the internal chromatin structure for example by reducing the number of protein cross-linkers.

Results of our simulations on stretched chromosomes show that the elasticity is highly

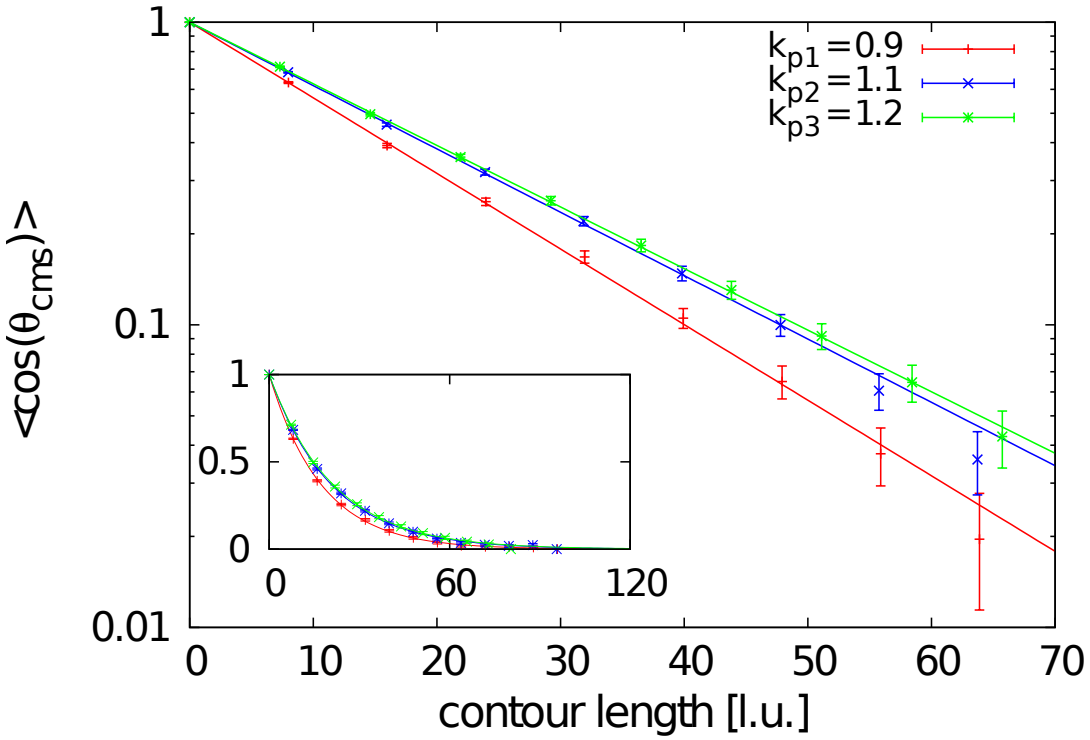

Figure 4.10: Directional correlation functions for model chromatids with different cutoff lengths. The chain length for all configurations is $N = 650$. Error bars represent the standard error. Shown are results for $C = 50$ and different mean loop concentrations. An increased number of cross-links is associated with a more densly packed chromosome. Thus, the distance between loops is decreased and the repulsive forces between them are stronger. Consequently higher bending rigidities are obtained.

dependent on the mean loop concentration. We performed simulations with the same cutoff length $C = 50$ for polymers consisting of $N = 650$ monomers, but different loop concentrations. The results show that when the mean loop concentration is increased from 0.9 to 1.2, the Young's modulus of the model chromatid increases by a factor of two. Tighter compaction is thus associated with a strongly decreased elastic response. This result is plausible as more cross-links within the fiber means that segments are glued to each other more efficiently and hence are harder to be stretched by an external force. Therefore, the loop formation is also responsible for the elastic response and the number of loops controls the stretching stiffness of the chromatid. We calculated the bending moduli that would be associated with the obtained Young's moduli if chromatids were cylinders made of a homogeneous material. The calculated bending moduli are three to five times smaller than the ones obtained from direct measurement of the flexibility. Table II gives an overview over results obtained from both, elasticity measurement and direct measurement of bending fluctuations.

Estimation of the thickness of elongated chromatids show that the widths decrease when the chromatids are pulled. This is consistent with the observed shift of the loop size distribution from larger to smaller loops and the decrease of the mean loop size. We calculated Poisson's ratio to quantify this finding. The inset in Figure 4.14 shows the relative change in width to the relative change in length of the chromosomes. The relationship is not linear from the start but rather converges to a linear curve. The

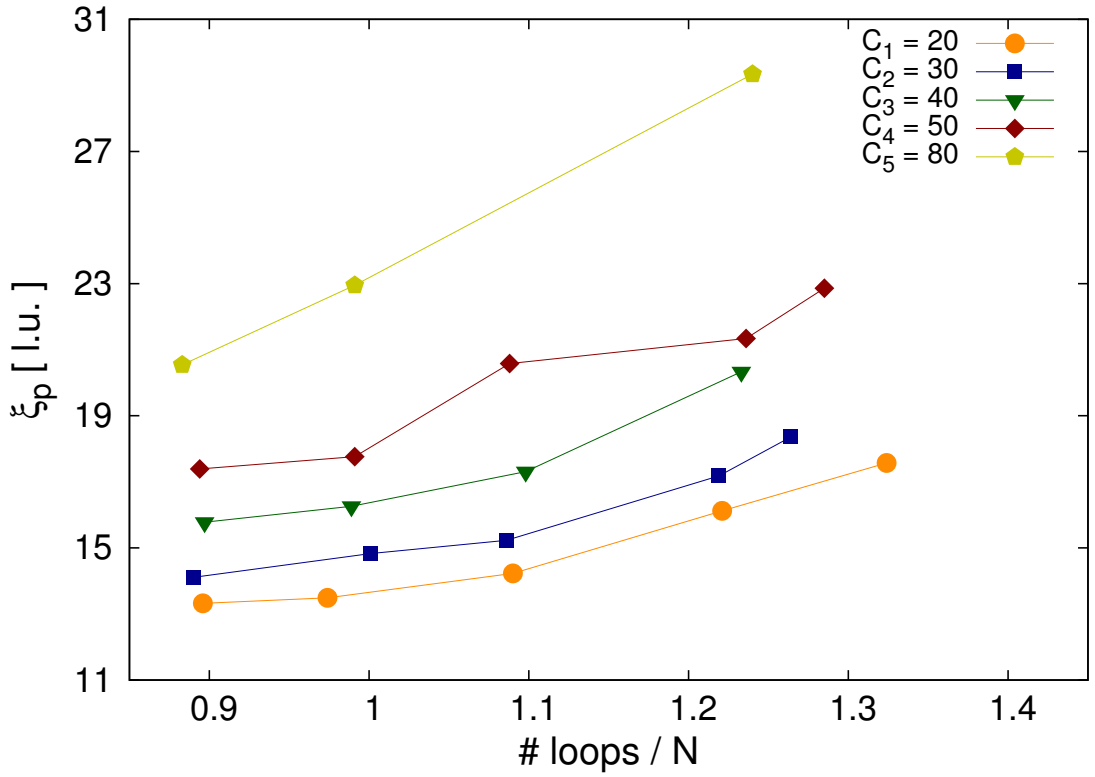

Figure 4.11: Persistence length in dependency of the cutoff length C and the mean loop concentration k_p. The persistence length and thus the bending rigidity increases with higher cutoff length and higher mean loop concentration. However, no simple dependency can be derived from the results. As the internal structure of the model chromatids is complex, the persistence length also has a complicated relation to the parameters.

Poisson's ratio is determined by fitting the linear region of the curves. Experimental studies by Poirier et al [113] resulted in a Poisson's ratio of $\nu = 0.069 \pm 0.005$ for newt lung cells. For our studies, different Poisson's ratios were obtained for different mean loop concentrations. With values between $\nu = 0.045$ and $\nu = 0.065$, our results are in the same range as the experimental findings.

4.4 Discussion

In this work we used a coarse grained polymer model to investigate if the condensation during mitosis can be understood by a probabilistic, locally restricted cross-linking mechanism of the chromatin fiber. We showed that this mechanism results in a tight compaction of the chromosome. The restriction of the loop sizes by a cutoff length in our model implicitly describes the fact that long range interactions cannot be formed in mitosis while the dynamical formation and dissolution of crosslinks implicitly accounts for the dynamics of the binding proteins in the surround solvent.

Although we do not describe explicitly the binding of proteins to DNA at special binding sites and do not want to state that condensins or Topoisomerase II could not link distant segments of chromatin to each other, we believe that it cannot be excluded that there could be principles that allow binding proteins to distinguish between different

Figure 4.12: Simulation of pulling of the model chromatids. Model chromatids at different pulling forces. For small elongations the chromatid is stretched but the total number of cross-links does not change. For higher elongation the number of cross-links decreases rapidly and the chromatid becomes inhomogeneous.

chromosomal regions. We believe that the cohesion of sister chromatids can serve as an good example. It seems that cohesin proteins have the ability to distinguish between the chromatin strands of the sister chromatids. Therefore, in the same way, condensins could have a principle after which they distinguish between chromatin segments that are genomically close and those that are genomically far away. For example the chemical composition of different chromatin segments, e.g. through histone modifications, could play a role at this.

In order to validate our model we compared the geometry and especially the mechanical properties, i.e. flexibility and elasticity, of the model chromatids to experimental findings. With our model we obtained objects that matched the shape of mitotic chromosomes and flexibility of chromatids assembled from Xenopus laevis egg extract [112]. In particular, we observed a much increased bending stiffness compared to simple polymer models such as self-avoiding walks, which can be explained by the entropic repulsion between the chromatin loops that are formed by the cross-linking of the fiber. Simulations of applied stretching forces revealed changes in the loop structure with a reorganization for small forces, followed by breakage of loops at large forces. We found that the loop structure, the size and the average number of cross-links within the chromatin fiber are essential for the mechanical properties. Therefore we suggest that altered physical dimensions and mechanical properties in different stages of mitosis and across different species can be

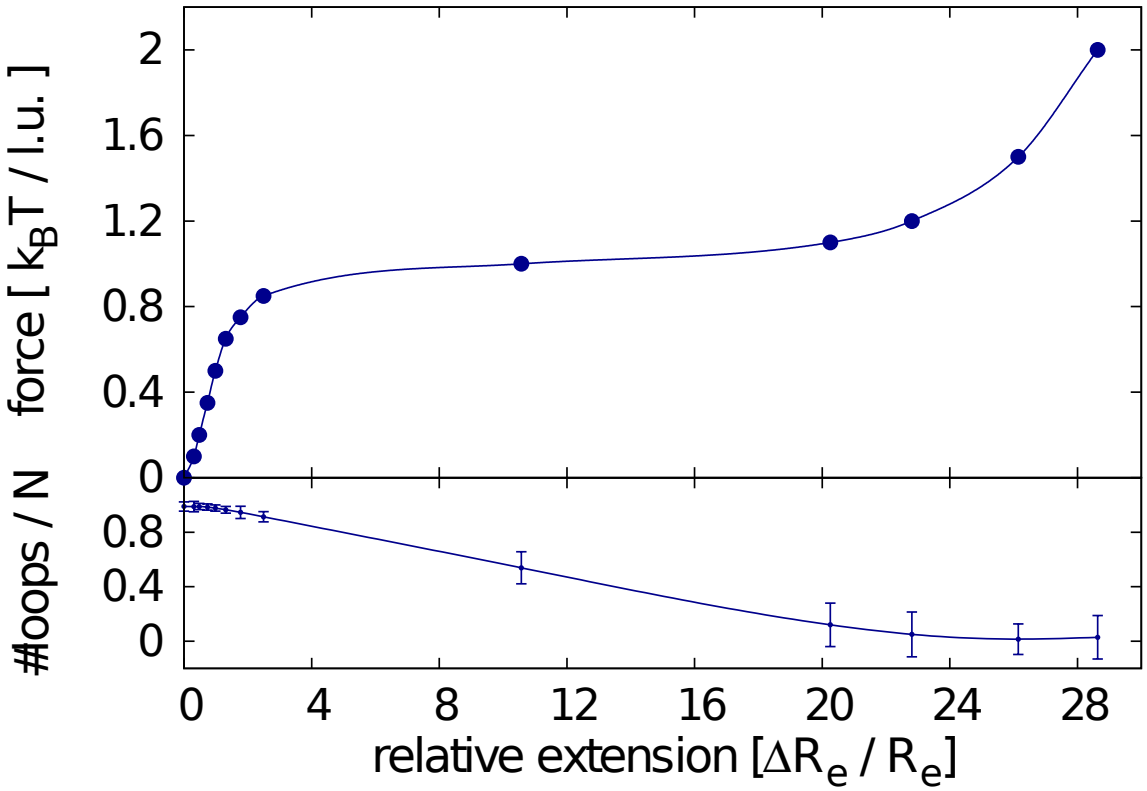

Figure 4.13: Force extension curve. Detailed look at the force elongation curve for the configuration $N = 650, C = 50, k_p = 1.0$. In the range of extensions up to two times of the native length, a linear dependency can be observed, where the total number of cross-links remains nearly constant. For higher extensions a force plateau is reached. Here the number of cross-links decreases and the chromatid is unfolded rapidly. This region corresponds to a decondensation region.

explained by different internal loop structures.

At high looping probabilities, the cross-linking of the fiber results in a condensation into a homogeneous, rod-like shape. The lengths and widths of the model chromatids match observations of chromosomes of numerous species, in particular those assembled from Xenopus egg extract [112, 189]. We found that the length and the thickness of the model chromatids are governed by the size restriction for chromatin loops and the number of loops within the chromatin fiber. The linear dependency of the chromatid thickness with the cutoff length is consistent with the simplified assumption of a helical folding of the chromatin fiber, which was in fact one of the earliest propositions for metaphase chromosome structure [111, 169]. However, the Dynamic Loop Model for mitotic chromosomes is not a simple helix but rather resembles a chromatin network with the tendency of the fiber to form rings around the central axes. A chromatin network was considered as the structure of mitotic chromosomes before [117, 176].

In our coarse grained description we obtain lengthwise compaction ratios between 10 and 30 fold of the native length, depending on the upper restriction for the loop size. As the compaction from the $30\,nm$ fiber to the mitotic chromosomes is in the range of 500-fold, this would suggest that our coarse-grained chromatin fiber has a diameter much larger than $30\,nm$ but still well below $1\,\mu m$. Our model can therefore be seen in the context of a hierarchical folding model for the mitotic chromosome [23, 173]. Here the dynamic

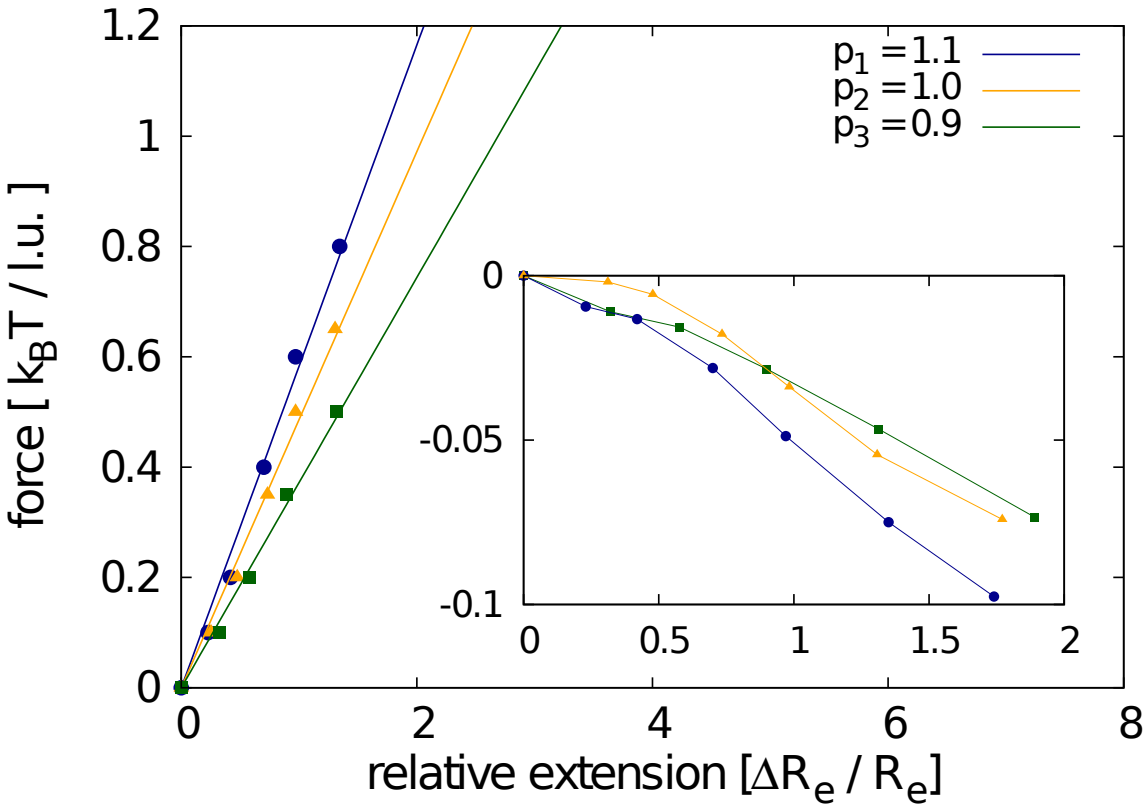

Figure 4.14: Linear region of force extension curve. The slope of the force elongation curve in the linear part depends strongly on the mean loop concentration. Here we show results for $N = 650, C = 50$ and three different values for k_p. The force modulus for configurations with mean loop concentration $k_p = 1.1$ is more than double than the modulus for configurations with $k_p = 0.9$. Hence we conclude that different elastic responses can be explained by altered loop structures. The inset shows the relative change in chromatid thickness against the relative extension. Similar Poisson's ratios are obtained for the different configurations. The values are in the range $\nu = 0.04$ and $\nu = 0.07$ and therefore close to experimental findings.

formation of cross-links would account for the compaction in one hierarchy level. On the other hand we find that for fixed C-to-N values and fixed k_p, the lengthwise compaction ratio increases with the chain length as the loop structure becomes finer. Therefore the Dynamic Loop Model is able to produce high compactions when the chain length is large enough. The network model that was put forward by Poirier and Marko assumes the cross-linking of the $30\,nm$ fiber [113]. However, the simulation of such high compaction ratios requires the equilibration of very long polymers that is computationally not viable.

Experimental results suggest that the flexibility of chromosomes is subject to the species and to the stage of mitosis [43]. In our simulations we found ratios of persistence length to thickness of the chromatids in the range of 1.5 to 4, depending of the choice of parameters. This is consistent with the experimental findings of Houchmandzadeh and Dimitrov [112] on Xenopus egg extract who reported a ratio of approx. 3.4. Hence, the bending rigidity of the chromatin structure in the egg extract can be explained by the loop formation alone, without the assumption of a protein scaffold. However, other experiments of in vitro and in vivo assembled chromosomes from Xenopus cells, newt lung cells, the newt TVI cell line and Drosophila cells found much higher bending rigidities

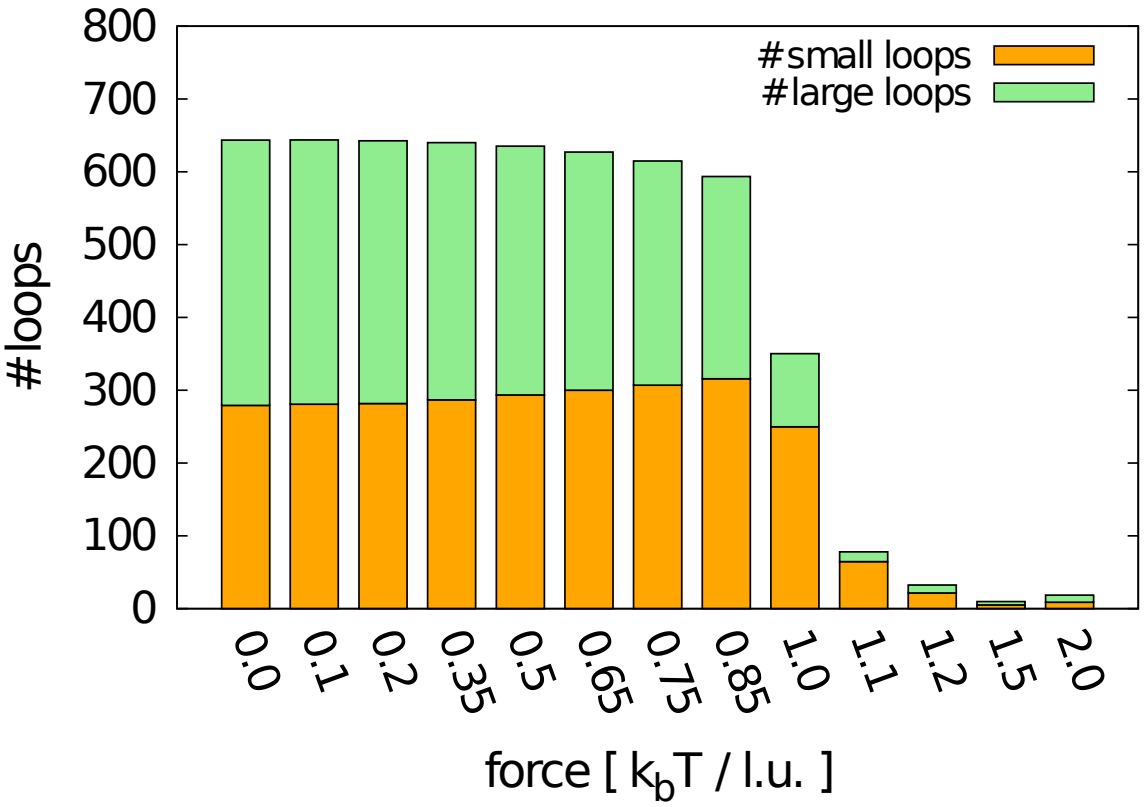

Figure 4.15: Reorganization of the loop structure . Although the total number of loops is constant in the linear region, there are changes in the loop structure. Shown are results for $N = 650, C = 50, k_p = 1.0$. The loop domains are reorganized due to the pulling force, with the proportion of small loops (size ≤ 15) increasing at the expense of the number of large loops (size > 15). As the thickness is essentially determined by the size of the loops this finding indicates that chromosome width decreases in this area which is consistent with experimental results.

with persistence lengths that are many times of the actual chromosome length [114, 115]. Poirier et al [115] suggested that the differences between the egg extract and in vivo assembled chromosomes arise from different chromatin organization in both systems. This could be connected to the different functions of egg extract and somatic tissue culture cells or because egg extract chromatids are not completely condensed. We observed for the Dynamic Loop Model that increasing looping probabilities resulted in much higher bending rigidities, thus supporting the argument that not fully condensed chromosomes are more flexible. Moreover, condensins, which are the main candidates for the binding proteins, were found to be able to dimerize and also to form heterodimers with other proteins [190, 191]. When cross-links can cluster in this way, it has to be assumed that the loop concentrations in real chromosomes are much higher than it is possible to model in our coarse grained approach. Therefore, consideration of such protein-protein interactions in the model could also account for an enhanced stiffness.

Furthermore, we have to point out that the entropic repulsion between chromatin loops is not the only factor that determines the flexibility of chromatids. Rather we suggest that these entropy effects contribute to the bending rigidity, and in some cases, such as for chromatids from Xenopus egg extract, are sufficient to explain them. However, there are certainly other factors, such as the surrounding solvent, that do also contribute to the

mechanical properties.

The stretching simulations revealed that the looping mechanism results in a very elastic chromatid that can be stretched to many times of its native length. For elongations of up to three times of the native length, a linear relationship between stretching force and relative extension was found. This is in agreement with experimental findings where chromosomes as well as single chromatids behave like a homogeneous elastic material [111, 113]. Our results show clearly that the number of loops is of great importance for the elastic response of the chromatid. For fixed chain length $N = 650$ and fixed cutoff length $C = 50$ we observed that the increase of the initial average loop concentration from $k_p = 0.9$ to $k_p = 1.1$ is associated with a doubling of the Young modulus. Experimental evidence for this dependency was given by Almagro et al [96]. The authors measured the elastic response of Xenopus egg extract chromatids after cleavage of SMCs with trypsin. It was found that chromosomal domains containing SMC proteins had a much higher stretching stiffness (up to four times) than domains where parts of these SMC proteins were cleaved. As SMCs are subunits of condensin proteins which are most likely to be responsible for chromatin cross-linking, our results confirm this experimental finding, as higher loop concentrations in the Dynamic Loop Model are also associated with an increased stretching stiffness.

Closer examination of the loop structure in this region of small relative extensions showed that a reorganization takes place when the chromatid is stretched. Such a behaviour was proposed before in the network model of Poirier et al [116]. The loop size distribution shows a shift from large loops to small loops and therefore leads to a thinning of the chromatid. The measurement of values of 0.045 to 0.065 for the Poisson's ratio in our model is in good agreement with experimental results from Poirier et al [113] with a value of 0.069. We have to point out that our Monte Carlo algorithm simulates chromosomes in thermal equilibrium and the stress is introduced by a pulling potential representing the force. However, it might be that this kind of approach does not match experimental conditions as we do not impose a constant stretching rate. Due to the thermal equilibrium situation, the forces in the simulation were much lower than what one would get if the pulling process was assumed to be a non-equilibrium process.

For extensions higher than $\epsilon \approx 2$, a strong leveling off of the force extension curve occurred, resulting in force plateaus. Force plateaus were also observed in stretching experiments although only for long extensions of $\epsilon > 15$ and more. The fact that in our simulations the plateau regions started much earlier can be explained by the coarse grained character of the polymer model. In reality, the structure of the chromosomes are certainly much finer. In addition, the loop concentrations in real chromosomes are most probably also much higher, considering the effect of SMC dimerization and heterodimerization as has been pointed out before. However, at present, simulations of much finer systems are still computationally not feasible. Moreover, our model does not include elasticity of the underlying coiling chromatin fiber itself, which could also contribute to the total elasticity of the whole chromatid [192].

Different chromosome states after retraction from extension into the plateau region were reported for chromosomes from different animals and different ways of assembling (in vivo or in vitro). While Poirier et al [113] observed a swollen ghost state, Houchmandzadeh et al [111, 112] witnessed non homogeneous chromatids with alternating thick and thin regions and which are up to five times longer than originally. Such inhomogeneous chromatids are obtained in the Dynamic Loop Model when the number of loops are small and consequently cross-links are not located homogeneously along the chain (see Figure 4.1). A possible reason for this is that binding sites are destroyed when the elongation

is too far. Therefore, in the retraction process the looping probability might be much lower than in the original chain which in turn results in the longer and inhomogeneous chromosomes.

We have shown that the dynamic cross-linking mechanism leads to the condensation of the chromatin fiber. The loops within the fiber evoke an increased bending stiffness by entropic repulsive forces. Our model is able to explain the shape of mitotic chromosomes and the flexibility of mitotic chromosomes assembled from Xenopus egg extract. Furthermore, simulations of stretching forces showed good qualitative match of our results with experimental findings. We therefore conclude that the structure and mechanical properties of mitotic chromosomes are in a great part invoked by internal formation of loops of the chromatin fiber.

Chapter 5

Binding Dynamics and Mechanics of Sister Chromatids

References

The content of this chapter was adapted from the following publication

- Y. Zhang, S. Isbaner and D.W. Heermann, *Sister Chromatid Mechanics Studied with a Polymre Model*. Front. Physics, doi: 10.3389/fphy.2013.00016. 2013.

Chapter Summary

Sister chromatid cohesion denotes the phenomenon that sister chromatids are initially attached to each other in mitosis to guarantee the error-free distribution into the daughter cells. Cohesion is mediated by binding proteins and only resolved after mitotic chromosome condensation is completed. However, the amount of attachment points required to maintain sister chromatid cohesion while still allowing proper chromosome condensation is not known yet. Additionally the impact of cohesion on the mechanical properties of chromosomes also poses an interesting problem. In this work we study the conformational and mechanical properties of sister chromatids by means of computer simulations. We model both protein-mediated cohesion between sister chromatids and chromosome condensation with a dynamic binding mechanisms. We show in a phase diagram that only specific link concentrations lead to connected and fully condensed chromatids that do not intermingle with each other nor separate due to entropic forces. Furthermore we show that dynamic bonding between chromatids decrease the Young's modulus compared to non-bonded chromatids.

5.1 Introduction

In the interphase, eukaryotic chromosomes are replicated and two identical copies of each chromosome, called sister chromatids, are present in the nucleus. In mitosis, chromosomes undergo a condensation into very compact, rod-like objects that have a high stiffness. Chromosome condensation is necessary for the error-free separation of different chromosomes since their territories are overlapping in interphase [23, 73, 168]. To further ensure that sister chromatids are properly distributed to the two daughter cells, they are connected to each other, a phenomenon called sister chromatid cohesion. Cohesion is resolved in anaphase, after chromosome condensation is completed and all chromatid pairs are aligned at the equator of the mitotic spindle [54].

Without factors that facilitate cohesion, sister chromatids would quickly segregate due to physical properties. In particular, excluded volume interactions and entropic conditions that favor separated sister fibers would be sufficient to drive this segregation [28, 36]. On the other hand, the mitotic condensation process involves the formation of cross-links within the chromatin fibers [117] leading to the presence of a large number of loops. This can even further facilitate the segregation process due to the entropic repulsive forces between loops within sister chromatids [186]. Therefore attachment points between the two sister fibers are necessary. However, the abundance and position of attachments could have a profound influence on the conformational properties on sister chromatids and their condensation process. The main question that we target in this work is therefore: How does the combination of attachments between two sister chromatids and intra-chromatid cross-links determine the conformational properties of the sister chromatid system? We also address how the mechanical properties of a system of two connected chromatids is changed compared to single chromatids or non-bonded chromatids.

Cohesin is believed to be the main factor for the tethering of sister chromatids [105]. This protein complex is composed of Smc1 and Smc3 subunits of the SMC family and Scc1 and Scc3 [193]. It is believed to form ring-like structures when associated with chromosomes [194]. Different models exist to explain the exact mechanism by which the cohesin complex attaches sister strands to each other. A common interpretation is that cohesion forms a ring around both strands [106]. Another suggestion is that two cohesin rings each surround one strand of the chromatid pair and cohesion is established by binding of the two rings to each other [107]. A recent study has shown that cohesin could also passively facilitate chromatid cohesion by maintaining intertwining between sister chromatids in addition to its active tethering mechanisms [195].

Experimental studies showed that the location of cohesin binding sites along chromosomes are not fixed. Although cohesin is enriched at the centromere region, sister chromatid cohesion is spread also along chromosome arms [196, 197]. In particular, cohesin is mobile in the chromosomal domain and along the chromatin fiber, which in turn means that sites of cohesion are flexible and possibly transcription-dependent in interphase [198, 199]. Dynamics of cohesin on the chromatin fiber could be possible through sliding of the cohesin ring along the fiber [106] or binding and unbinding of rings in the handcuff model [107]. Interestingly, cohesion is established or reinforced genome-wide following DNA damage, thereby indicating that bonding between sister chromatids can be dynamically restructured [200–202].

While sister chromatid cohesion is important for the correct distribution of chromatids to the daughter cells, condensation of chromosomes in mitosis plays a key role for their error-free separation [23]. The condensin complex and Topoisomerase II have been identified as key proteins facilitating proper condensation of chromosomes. Micromechanical

experiments using micro pipettes were conducted to assess the internal organization of mitotic chromosomes. Direct measurements of the flexibility of single chromatids extracted from *Xenopus laevis* eggs showed a worm-like behavior of the chromatids [112]. Pulling experiments revealed a high extensibility of chromatids and chromosomes extracted from cells including human chromosomes [96, 112, 113, 116, 118]. However, chromosomes extracted from cells possess a much higher bending rigidity than egg extract chromatids, which could be due to different internal structures [43]. Additionally, the influence of cohesion between sister chromatids on the mechanical properties of chromosomes is also not well understood.

In this work we introduce a polymer model for mitotic chromosomes that includes mechanisms for the condensation of each chromatid as well as cohesion between sister chromatids. We model the cohesion between sister chromatids by dynamic binding and unbinding between the two sister fibers. The condensation of each of the chromatids is realized by dynamic intra-chromatid looping, which accounts for the presence of binding proteins such as condesins. We use computer simulations to sample possible conformations for different model parameters. Our results show that inter-sister bonds and intra-fiber cross-links can act together to realize condensation and cohesion at the same time. However, we also show that the inter-sister and the intra-fiber bonds compete with each other due to entropic constraints. We only observe condensed and aligned sister chromatids for a small and sensitive range of model parameters.

In pulling simulations we further study the mechanical properties of sister chromatid systems at different levels of cohesion and compare the results with simulations of single chromatids. We show that binding between sister fibers lead to an increase of the elasticity of the chromosome and facilitates unfolding upon stretching forces. In contrast to our model, simple polymer models are not able to explain the experimental observation of force plateaus following linear regions.

5.2 Methods

5.2.1 Model for Mitotic Chromosomes

The sister chromatids are modeled as polymer chains consisting of typically $N = 200$ or $N = 400$ monomers each. Such coarse-graining approaches have proven useful, since it is not necessary to know the exact configuration on a detailed level when simulating the structure of a complete chromatid. The coarse-graining allows us to neglect interactions on smaller scales such as electrostatic interactions or van-der-Waals forces and is also necessary for computational feasibility. The chromosomes were simulated as polymers on a lattice based on the Bond Fluctuation Model (BFM), a lattice model incorporating self-avoidance [163, 164]. The BFM has recently been extended to the Dynamic Loop Model, which has been successfully applied to inter- and metaphase chromosomes [184, 203].

In the BFM, monomers occupy a cube of 8 lattice sites and are connected to other monomers via bonds of fluctuating length (but otherwise static) allowing a maximal bond length of $\sqrt{10}\,l.u.$ (lattice units) [163]. With the Dynamic Loop Model, an additional binding mechanism has been introduced: monomers may temporarily establish a bond to other monomers nearby. In each Monte-Carlo step, all monomers are tried to move in a random direction. The move is accepted if the new site is unoccupied and the new bond vectors are allowed. If the monomer is now close enough to another monomer, a temporary bond is established with probability p_{bond}. The lifetime of the bond is drawn from a Poisson distribution, with the simulation parameter τ_{bond} as its mean value. The

bond dissolves again when its assigned lifetime expires. Each of these additional bonds between non-adjacent monomers forms a new loop of the chromatin fiber. The size of the loop is then determined by the separation of the two monomers along the fiber. The dynamic looping of the chromatin fiber results in a mean number of loops n_{loop} and a mean loop concentration $k_{loop} = \frac{n_{loop}}{N}$. It models how binding proteins such as the condensin complex can temporally bind chromatin segments to each other.

For the folding model of mitotic chromosomes we introduce a limitation to the size of the loops called cutoff length C. This means that monomers can only form a loop bond if their separation along the fiber is smaller than C. This cutoff length is firstly based on the observation that mitotic chromosomes form rod-like objects instead of spherically shaped clumps, which they do without limitation of the cutoff length. The entropic forces that are exerted by the loops determine the mechanical properties of the model chromatid. Details of the model for single chromatids can be found in an earlier work [203].

For the sister chromatid systems we allowed not only the monomers of one strand to bond to each other and thus form loops within the chromatin fiber, but also for monomers belonging to different strands to bond to each other forming interlinks. The mechanism for these interlink bonds are essentially the same as for the loop bonds within one strand. If two monomers from both strands come into physical proximity of each other in the Monte Carlo process, they can form an additional bond with probability p_{link}. A lifetime which is drawn from a Poisson distribution with mean τ_{link} is assigned to this bond. Model sister chromatids can bind to each other through this dynamic link formation which results in a mean number of interlinks n_{link} and a mean link concentration $k_{p,link} = \frac{n_{link}}{N}$. Just like the looping mechanism within one strand models condensin binding, this linking mechanism models how the cohesin complex binds sister chromatids to each other.

5.2.2 Pulling Simulations

For the pulling simulation, a force is introduced by a pulling potential $U_{pull} = -F \cdot |\mathbf{r_N} - \mathbf{r_1}|$. Here, $\mathbf{r_1}$ denotes the position of the first monomer of the chain and $\mathbf{r_N}$ the position of the last monomer. The force F is a simulation parameter in units of $k_BT/l.u.$. The Boltzmann factor $\exp\left(-\frac{\Delta U_{pull}}{k_BT}\right)$ with $\Delta U_{pull} = U_{pull}(\text{current step}) - U_{pull}(\text{proposed step})$ replaces the probability to move for the start and end monomer. In the pulling simulations, the ends of the two sister chromatids are concatenated permanently to each other in order to obtain a wels defined pulling direction.

To avoid abrupt high pulling forces and too fast pulling of the fiber, we increase the pulling force gradually by small steps starting with a small value. After applying a new force, chains are first equilibrated and conformations then sampled from the equilibrium distribution. After typically sampling a few thousand conformations we then increase the pulling force by a small step again. Thus, in every point in the stress-strain diagram the chains are in equilibrium which means that our pulling simulation is a reversible and adiabatic process.

5.2.3 Autocorrelation Time

In one Monte-Carlo step as described above, the conformation changes only locally. Since we want to calculate ensemble mean values and corresponding fluctuations, independent samples have to be analyzed. We use the autocorrelation time $C(t)$ to determine when two subsequent conformations in the Monte Carlo simulations are independent. The auto correlation function for an observable $A(t)$ is defined as

$$C(t) = \langle A(s+t) \cdot A(s) \rangle_s - \langle A(s) \rangle_s^2 \tag{5.1}$$

and is usually normalized to $\rho(t) = \frac{C(t)}{C(0)}$. It measures whether samples are correlated ($\rho(t) = 1$) or uncorrelated ($\rho(t) = 0$). The auto correlation function goes to zero exponentially with time (i.e. Monte-Carlo steps). We use Sokal's automatic windowing algorithm to compute the integrated autocorrelation time τ_{int} [161]. Conformations separated by 5 τ_{int} time-steps are treated as independent samples. As observable A we used the radius of gyration.

5.2.4 Radial Distribution Function

The radial distribution function (RDF) is a measure for the probability to find a pair of monomers at a separation $\mathbf{r}$. It is defined as

$$g(\mathbf{r}) = \frac{1}{N} \left\langle \sum_i \sum_j \delta(\mathbf{r} - \mathbf{r_{ij}}) \right\rangle \tag{5.2}$$

where $\mathbf{r_{ij}} = \mathbf{r_i} - \mathbf{r_j}$ denotes the separation of monomers i and j. The sum is taken over all relevant monomers and the average is taken over the whole sample of conformations that we obtained with the MC simulations. Assuming an isotropic system the relevant measure becomes only dependent on the distance r but not the direction. In this work we calculate the RDF by taking all the distances between pairs of monomers and create a normalized histogram with them. Thus we obtain the probability distribution function to find two monomers at the distance of r from each other. We distinguish between the RDF calculated for monomers on the same chain giving information on the size of an individual chain, and the cross RDF for monomer pairs each belonging to a sister chain, which yields information on the distances between the two chains. With the cross RDF it is possible to distinguish between chromatids that are intermingled and those that are aligned but separated. Intermingled chains have a well localized RDF, whereas the RDF for separated chains is smeared out to larger distances.

5.2.5 Chromatin Density Distribution

The chromatin density distribution denotes the distribution of the average density of chain monomers that can be found in the vicinity of a single monomer. We calculate this property by counting the number of chain monomers in a sphere with radius r_S around each monomer in the simulation and then averaging over all monomers in the system. We perform this calculation for all conformations that we sampled with the MC simulations yielding a probability distribution function for the average density. In the BFM, the bond length between monomers can have a distance of up to $\sqrt{10}$. Therefore we choose a larger radius for the calculation of the monomer density and set a value of $r_S = 6$.

Furthermore, we distinguish between the average density of monomers that belong to the same chain as the monomer and the average density of monomers that belong to the sister chain. Both distributions are compared to each other to determine if sister chromatids are intermingled or separated. In the case of intermingled sister chromatids, both distributions are the same, since around all monomers, the average density of monomers belonging to the same chain is the same as the average density of monomers belonging to the sister chain. On the other hand, sister chromatids that are not intermingled and thus distinguishable from each other have different distributions. In this case, the average

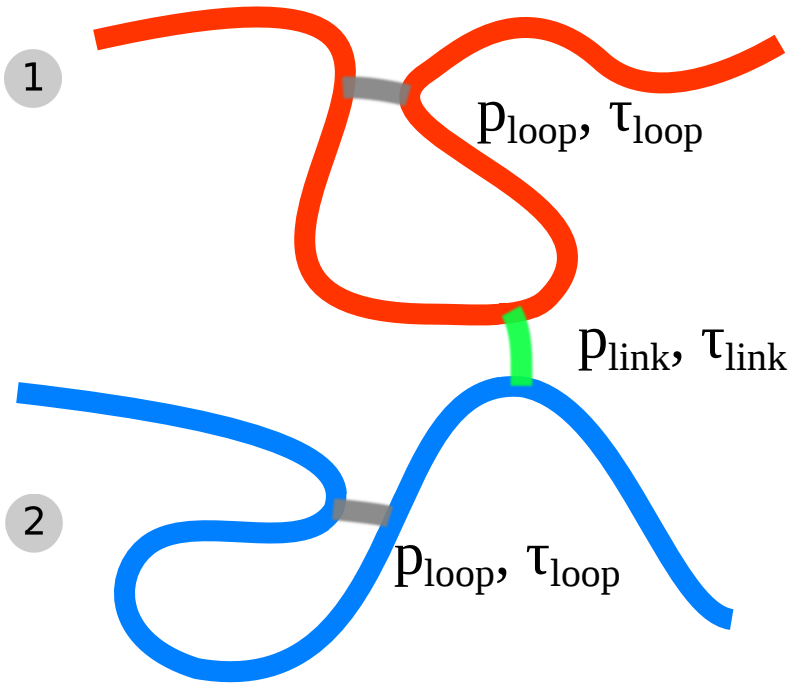

Figure 5.1: Schematic of the Dynamic Loop Model for mitotic sister chromatids. Each chain represents a coarse grained sister chromatid (fiber 1 and 2). The folding of each single chromatid is modeled by internal cross-linking of each of the chromatids forming chromatin loops (grey cross-links). Furthermore the model sister chromatids can be tethered to each other by inter-chromatid cross-links (green cross-links). We model both kind of cross-links with a dynamic mechanism.

density of other monomers that belong to the same chain in the surrounding of a specific monomer is much higher than the average density of other monomers that belong to the sister chain because the distance to the sister chain is much larger.

5.3 Results

5.3.1 Model

The folding behavior of the chromatin fiber cannot be feasibly modeled on an atomistic scale. Instead, we pursue a coarse grained approach for the description of chromosomes in metaphase. The chromatin fiber is represented by a polymer chain with N monomers. Each monomer can be seen as an effective substitute for a statistical segment which have on average the same behavior on a more detailed scale. However, the small-scale details do not contribute to the large-scale folding properties and thus can be neglected [41].

In mitosis, chromosomes undergo a condensation into very compact, rigid and rod-like objects. This condensation is believed to be facilitated by different proteins, in particular the condensin complex [204]. On the other hand, condensin was observed to be highly mobile within chromosomes in different stages of mitosis [172]. To account for this phenomenon, we introduce a dynamic and probabilistic cross-linking mechanism of the chromatin fiber for single chromatids in mitosis. In our model, two non-adjacent monomers belonging to the same fiber can form an additional bond between each other when they come into close proximity by diffusion. The probability of the bond formation is given

by a model parameter p_{loop}. A lifetime τ drawn from a Poisson distribution with mean τ_{loop} is assigned to each bond. The formation of a bond means at the same time that a loop of chromatin fiber is established. In order to account for the observation that for example the ends of a chromatid can never be bound to each other we exclude arbitrarily large loops by introducing a size restriction C. This dynamic looping mechanism results in the condensation of chromatid into a rod-like object when the model parameters are chosen such that the average concentration of loops $k_{p,loop}$, which denotes the average number of loops n_{loop} over the number of monomers N is high. The motivation in the model for the incorporation of the size restriction for intra-fiber loops is based on the observation that long-range interactions of the chromatin fiber do not exist in mitosis. A possible reason for this lack of long-range interactions in mitosis could be that chromosomes fold up locally first when entering mitosis, for example through a length-wise condensation of the fiber [185], forming local, compact blobs. Such blobs would give rise to a chromatin-solvent interface which was not present before (as the blobs are much more compact than the loose interphase chromatin). A surface like this could prevent the formation of cross-links between chromatin segments that are in different blobs, effectively inhibiting long range contacts. Moreover, the chromatin fiber is not homogeneous along the genome but rather has variations in many different quantities such as gene density, different types of histone modifications or DNA methylation. These chemical variations along the chromatin fiber could also make it possible for binding proteins to distinguish between segments that are far away along the genome and segments which are close. This could provide a possible biological mechanism for the establishment of a cutoff length for the loop size in the chromatin fiber. Details and results for single chromatids can be found in a previous work [203].

In this work, each of the sister chromatids is modeled by such a dynamically looping fiber. Additionally we include the effects of sister chromatid cohesion by introducing a similar dynamic binding activity between the two sister fibers. Two segments, each belonging to one of the sister chromatids can form a bond upon collision with each other by diffusion. The rate of such associations is controlled by the probability p_{link} while the dissociation rate is controlled by the lifetime of the cohesion bond that is drawn from a Poisson distribution with mean value τ_{link}. This dynamic association and dissociation results in a mean concentration of sister bonds $k_{p,link}$ that is dependent on p_{link} and τ_{link}. Figure 5.1 shows a schematic description of the model highlighting the cross-linking and interlinking mechanism.

5.3.2 High Number of Attachment Points Prohibits Condensation of Chromatids

We sampled conformations for different parameter settings of $p_{loop}, \tau_{loop}, p_{link}, \tau_{link}$ resulting in different loop and link concentrations $k_{p,loop}$ and $k_{p,link}$. Figure 5.2A shows a conformation with a large number of interlinks between the model sister chromatids. Such a high interlink concentration results in sister fibers that are highly intermingled and the overall shape of the indistinguishable mixture of the two fibers is rather spherical. Clearly, such kind of conformations do not resemble eukaryote chromosomes in mitosis after prophase. In Figure 5.2B we show the bonds between sister chromatids. If two segments from different sister fibers are bound to each other we visualize this by a red connection. The high number of interlinks prevents the sister chromatids from condensation and adoption of a rod-like shaped structure. Since such high interlink concentrations will inevitably result in such kind of intermingled fibers, we conclude that the number of

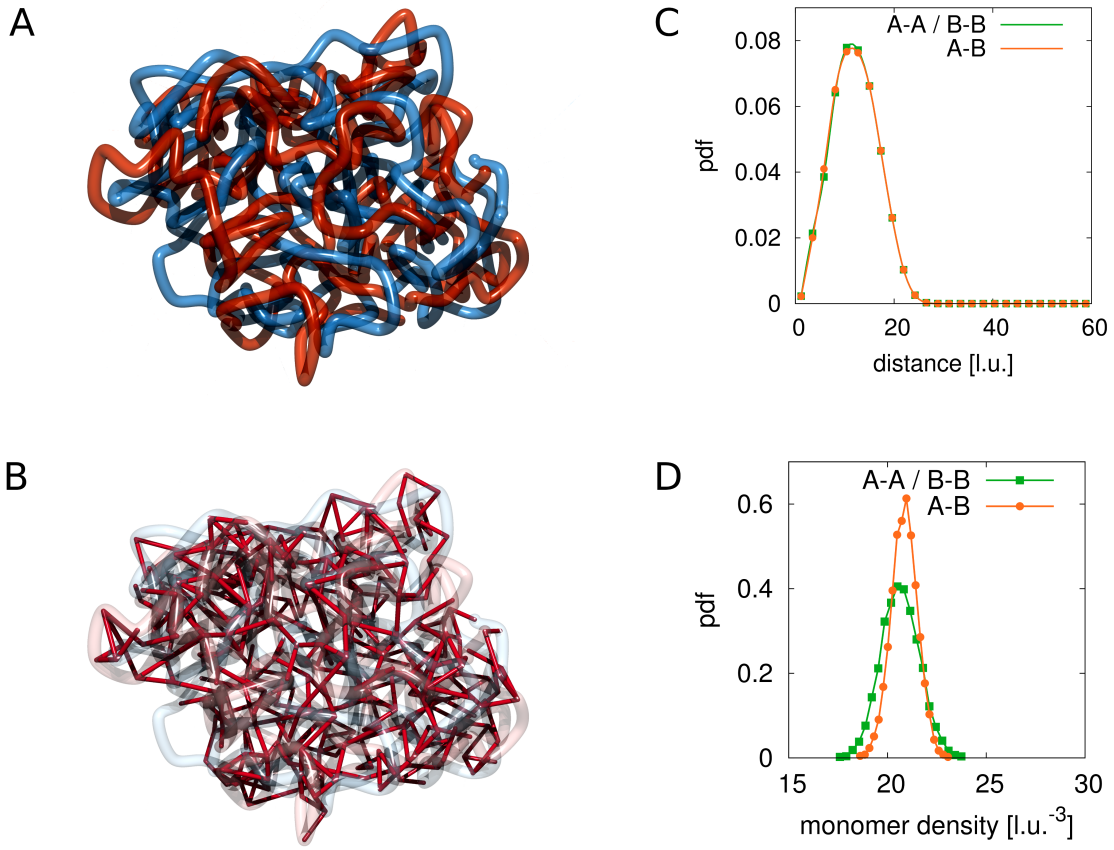

Figure 5.2: Example conformation for completely intermingled sister chromatids. The red and blue strands each represent one sister chromatid and the red links visualize the interlinks between the two fibers. **A.** For large numbers of interlinks the two sister chromatids are strongly intermingled due to the many tethering points. Such conformations resemble the situation after chromosome replication but before mitotic condensation has taken place. **B.** In this image we highlight the interlinks between the two sister chromatids. We observe that the large number of interlinks leads to an strong intermingling of the two chromatids. Since links between the sister fibers are randomly established upon collision of fiber segments they can be found anywhere along the fibers. **C.** In this panel we show the radial distribution function between segments of the two sister chromatids. In the case of completely intermingled chromatids, the correlation between polymers belonging to the same chromatin fiber (A-A, B-B) and the correlation between monomers from the two different chains (A-B) are exactly the same. **D.** This panel shows the density distribution of other monomers surrounding each monomer. In the intermingled state, the distribution of monomers of the sister chromatid in the vicinity of a monomer is the same as the distribution of monomers of the own fiber around this monomer. It shows that the environment of each single monomer does not indicate its membership to either one of the chains.

tethering points between sister chromatids must be limited.

To assess the degree of intermingling between the two sister strands we calculate the radial distribution function for monomers belonging to each of the fibers and a cross-pair radial distribution function between monomer belonging to different sisters. Figure 5.2C shows these radial distribution functions for the completely intermingled state. All three distributions are identical, which means that the average positioning between monomers

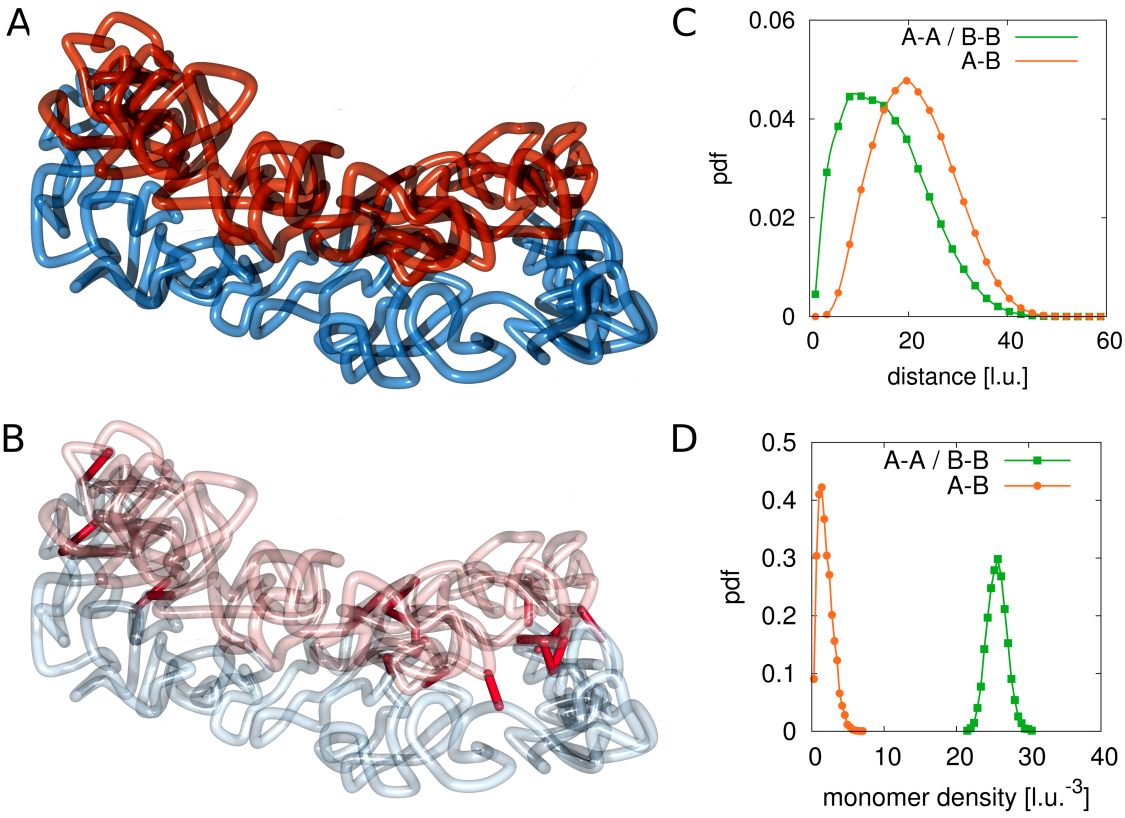

Figure 5.3: Example conformation for different configurations. A. At certain values of the interlink concentration, sister chromatids segregate due to entropic repulsion but are still concatenated by a few interlinks. These configurations resemble the situation found in sister chromatid systems in metaphase. When the interlink concentration is further decreased, the two chromatin fibers separate completely from each other. **B.** This figure highlights the present interlinks between the sister fibers. Interlinks are found along the contour of both model chromatids. **C.** The radial distribution function between monomers of different fibers is shifted to larger values compared to the function between monomers of the same fiber. **D.** The concentration of monomers of the own fiber is much higher than the concentration of monomers of the sister fiber around one monomer.

of different chains is the same as between monomers of the same chain. Additionally, we calculate the chromatin density distribution around each segment of the fibers. The results are shown in Figure 5.2D. The green curve shows the density distribution around a statistical segment that is produced by its own fiber. The orange curve shows the density distribution that is produced by the sister fiber. In the intermingled state, the same distribution can be found in the environment of all segments. Therefore, the two chains cannot be distinguished from each other in this intermingled state.

We performed simulations for settings with low linking probabilities and thus low ratios between association and dissociation rate for sister fibers. The results show that below a critical value for this rate, the entropic repulsion between the two condensed sister chromatids cannot be compensated by the dynamic linking mechanism. The sister chromatids become untethered and eventually drift away from each other as completely disconnected individual chains. This is verified by the radial distribution functions. In the

case of disconnected sister chromatids, the distribution for monomer pairs belonging to the same fiber is the same in both sister chromatids. The cross-sister radial distribution function however shows that distances between the sisters is highly variable and unlimited.

Only in a small window for the association and dissociation rates for interlinks can the model chromatids been observed in a condensed state where they have the shape of mitotic chromosomes while still being tethered to each other and not intermingled or completely separated. Figure 5.3A shows an example conformation for chromosomes in this state. This situation highly resembles the state of mitotic chromosomes after condensation and segregation in prophase until metaphase. Here, the average number of interlinks is very small compared to the number of loops within each of the models sister chromatids. This allows each single chromatid to be in the condensed rod-like state. The two sisters are then held together by only a few links along the contour of the rod-like chromatids without forcing an intermingling of the fibers. The fact that sister chromatids are not intermingled is verified by the radial distribution function (Figure 5.3C) and the chromatin density distribution (Figure 5.3D). The radial distribution function between monomers from the same chain has its maximum at a much smaller distance than the radial distribution function between monomer pairs from sister fibers thus indicating that sister fiber monomers have on average a much larger distance to each other than monomers from the same chain. Furthermore, the average density of other monomers from the same chain around any monomer is much higher than the average density of monomers from the sister fiber.

In Figure 5.4 we show a phase diagram for the different states of sister chromatids in this model. The diagram contains all the tested simulation setups with respect to mean interlink concentration between sister chromatids and mean loop concentration within each sister chromatid. For small interlink association to dissociation rates, sister chromatids separate since the entropic repulsive forces are stronger than the effective attractive forces by the dynamic interlinks. These setups result in separated sisters where the interlink concentration is zero. Large interlink concentrations result in intermingled sisters that do not have the characteristic rod-like shape. Only in a limited range of interlink concentrations, sister chromatids are both clearly distinguishable from each other and still connected.

5.3.3 Exclusive Permanent Linkage at the Centromere Does not Guarantee Alignment of Sister Chromatids

In this model, we do not restrict the sites of binding between model chromatids. However, this means that the resulting sister chromatids are not necessarily aligned in parallel. Instead, conformations where one end of one chromatid is connected to the center part of the other chromatid are also possible. Also, some kind of torsion where sisters are wrapped around each other can also be observed in some conformations. These kind of conformations naturally form due to the entropic freedom of the chains.

A well established view in sister chromatid cohesion is that the sister chromatids are permanently linked to each other at the centromere region. In particular the concentration of cohesin has been found to be enhanced at the centromeres. To assess how such a permanent linkage can affect the conformational dynamics of model chromatids we perform calculations at which both sister strands are bonded to each other at the middle forming a star-like polymer. In the polymer models permanent links of monomers represent an infinitely high binding potential. Such a potential is assumed for example between genomically adjacent beats of the chain. In this work we assume that cohesin concentration is considerably higher at the centromere than at chromosome arms result-

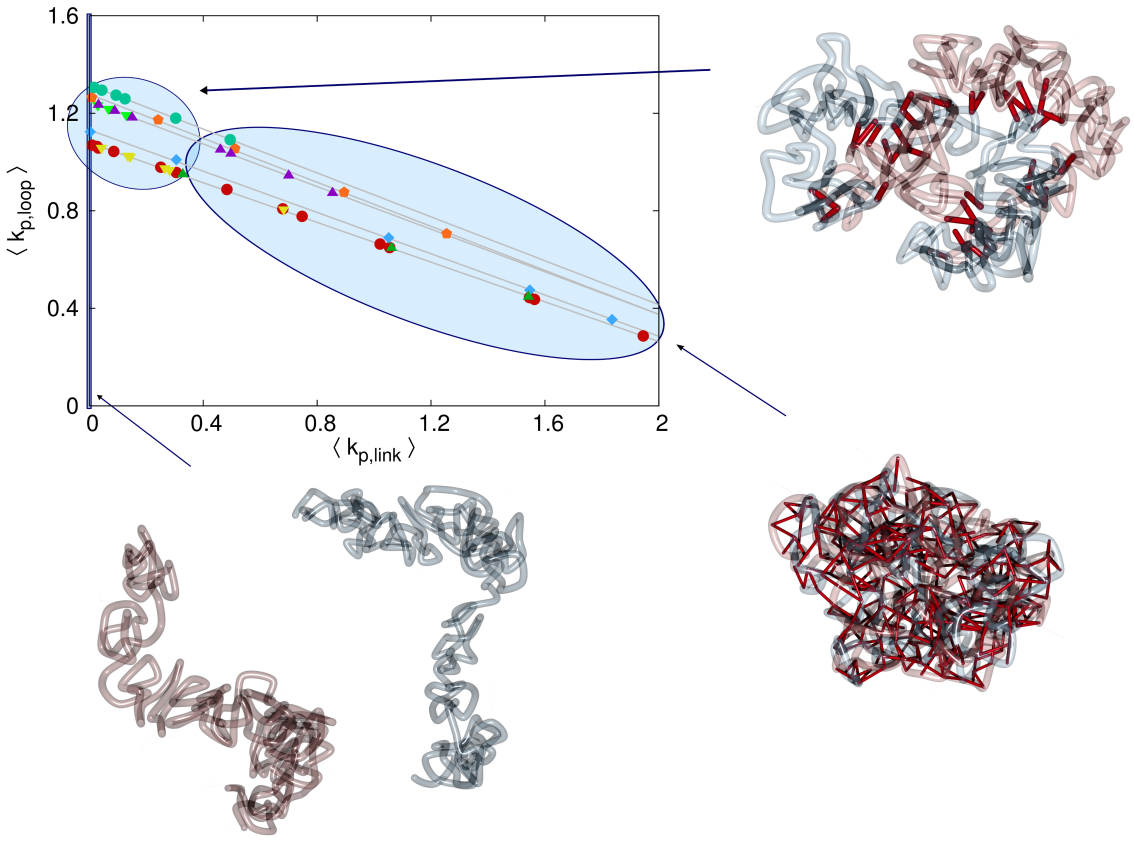

Figure 5.4: Phase diagram for the different possible configurations. The mean concentration of interlinks between the sister fibers is denoted by $\langle k_{p,link} \rangle$ and the average concentration of intrafiber cross-links is denoted by $\langle k_{p,loop} \rangle$. The different symbols in the diagram denote the different series of simulation runs. Note that the link concentrations are governed by the linking probabilities p_{loop}, p_{link} and the mean lifetime of links τ_{loop}, τ_{link}. If the probability p_{link} that one segment of the first sister chromatid forms a link with a segment of the second sister chromatid is very small, then the rate of interlink formation is not high enough to keep the two chromatids together. Entropic forces will then drive them away from each other and consequently the interlink concentration is $k_{p,link} = 0$. For very high p_{link} or long lifetimes τ_{link} the interlink concentration becomes so high that the two sister fibers are completely intermingled. Only in a sensitive intermediate region of interlink concentrations do the sisters segregate properly but are hold together by some interlinks preventing them to drift away from each other.

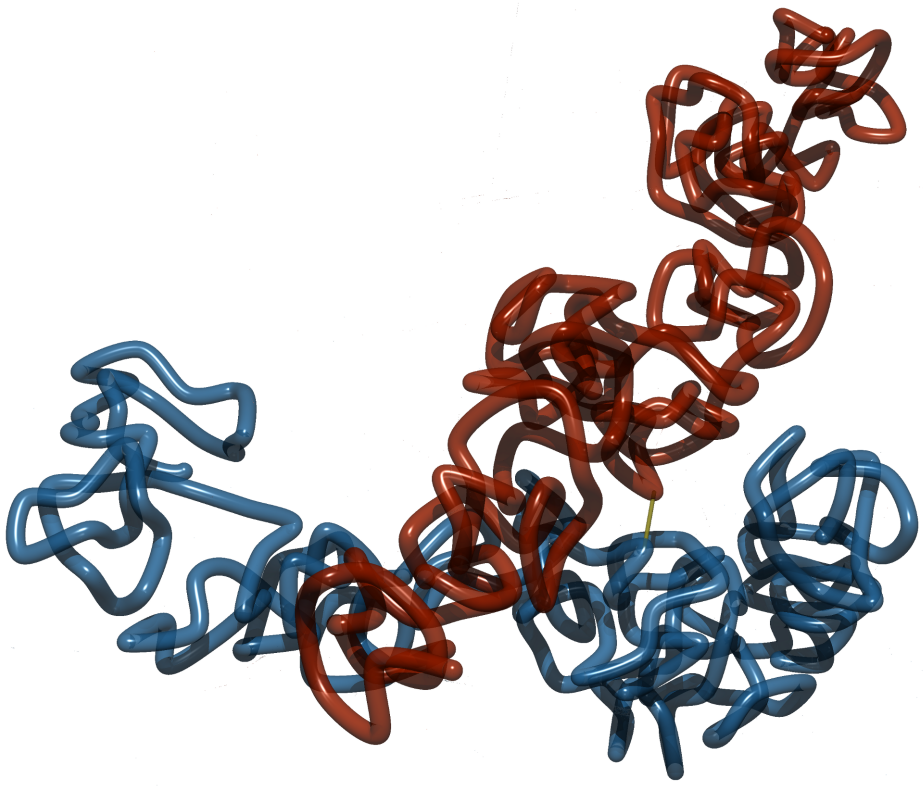

Figure 5.5: Two sister chromatids with a permanent link at the middle and no links otherwise. The single link at the middle holds model sister chromatids together. However, without further links at the arms, the entropic repulsive forces between the folded fibers makes it unfavorable for them to align in parallel to each other. Instead, a more cross-like conformation is preferred.

ing in strong cohesion in this region. Therefore a permanent link between the two center monomers of each model sister chromatid efficiently accounts for this enhanced bonds at the centromere. Intra-fiber cross-linking for chromatid condensation is included as in all other simulations, too. We test the alignment of sister chromatids for different interlink concentrations ranging from $k_{p,link} = 0$ to $k_{p,link} = 0.4$.

Our simulation results show that permanent linkage at the centromere without any other regions of cohesion, holds the chromatids together but does not maintain parallel alignment of the model chromatids. Due to the entropic repulsion between the looping fibers, sisters take up configurations rather resembling crosses. On the other hand we observe that chromatids permanently linked to the each other in the middle are much more likely to align in parallel for small link concentrations. An example conformation is shown in Figure 5.5.

5.3.4 Elastic Behavior of Tethered Chromatids

Micromechanical experiments on extracted chromosomes in mitosis intend to study the elasticity of mitotic chromosomes and thereby draw conclusions on the internal folding behavior of the chromosomes. Such studies have let to the suggestion of a network model

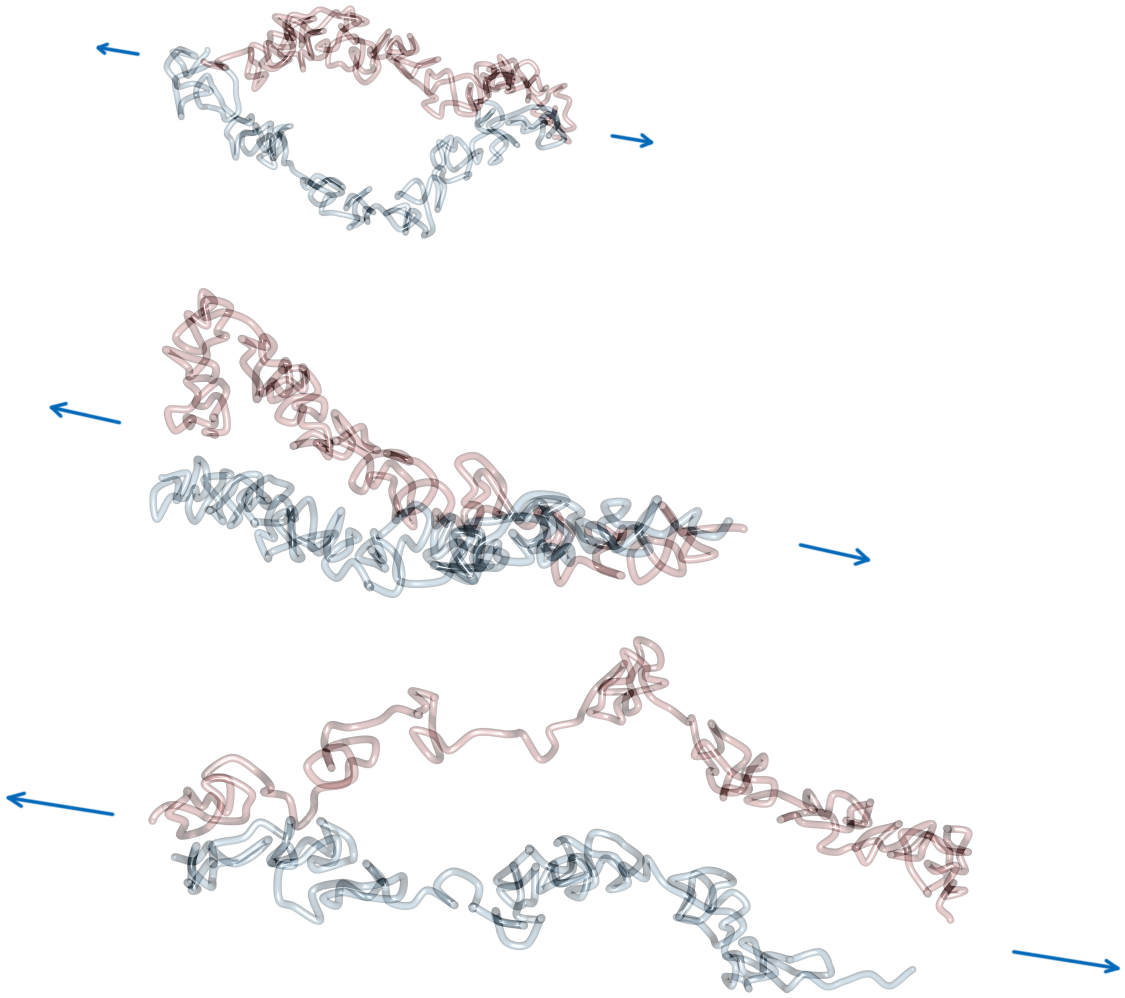

Figure 5.6: Force extension of non-interlinked model sister chromatids. Model chromatids are permanently linked together at the ends of the chromosome. The linking probability for interlinks is set to 0, which means that there are no interlinks between model sister chromatids. Therefore, segments of both fibers are not in close proximity along the contour and also not necessarily aligned even at higher pulling forces. The force was applied via a pulling potential U_{pull} that is proportional to the end-to-end distance of each model sister chromatid and the given applied force F. In this figure three example conformations at forces $0.6, 1.2$ and $1.8\, k_bT/l.u.$ are shown. The figure shows how the pulling force first cause a elongation of the model chromatids and eventually lead to a dissolution of the rod-like shaped structure of each of the two sister chromatids at large pulling forces. Note that in the visualization the chain ends where the chromatids were linked to each other are not shown.

for the chromatin fiber in mitosis and to our model of a dynamically folding chromatin fiber [117]. Micromechanical experiments are performed in vitro on chromosomes that can be isolated from cells or from egg extracts [43]. Especially for cell extracted chromosomes it can be expected that chromosomes consist of two tethered sister chromatids which often cannot be distinguished from each other [43]. Egg extracts on the other hand consist of single chromatids [111].

In this work we assess the mechanical properties of tethered sister chromosomes by measuring the elongation of model chromosomes under an external force. Model sister

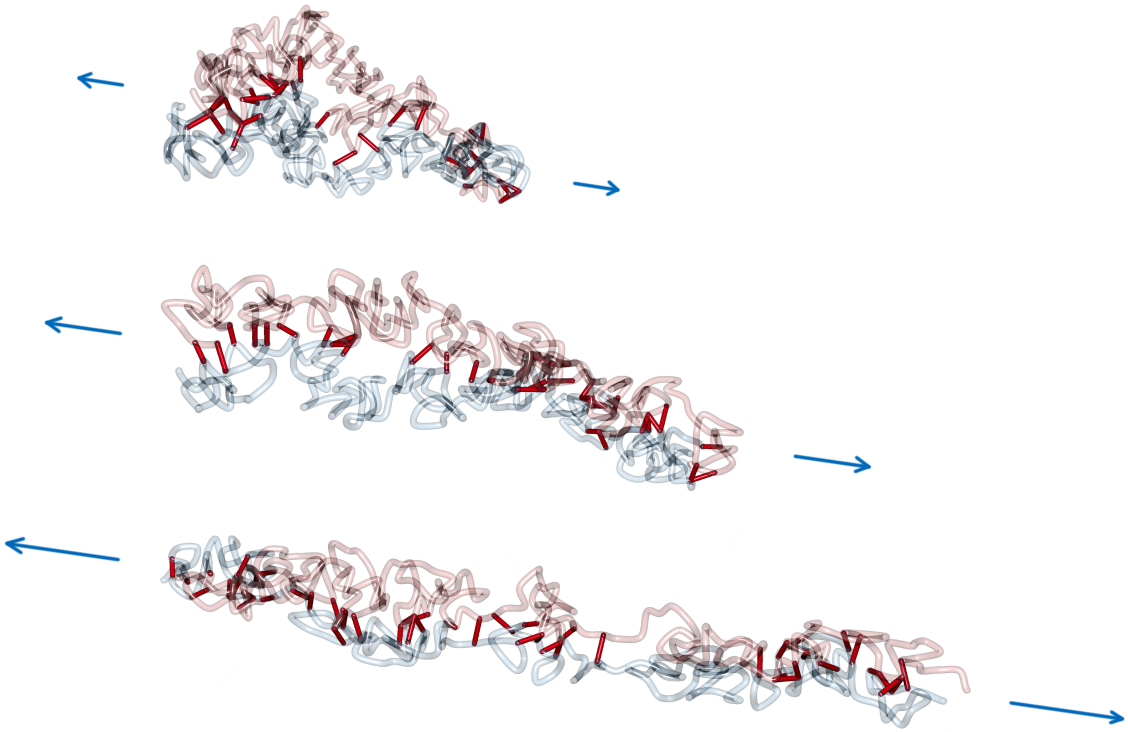

Figure 5.7: Force extension of interlinked model sister chromatids. In this scenario, a non-zero linking probability p_{link} and mean link lifetime τ_{link} is assumed. Consequently, the sister chromatids are aligned to each other during the whole pulling process. For large pulling forces the model chromatids disintegrate, but are still attached to each other.

fibers are permanently linked to each other at the ends. This is done to ensure that the chromatids have the same end-to-end distances. Also it prevents them from drifting apart from each other even in the case that the tethering probability is set to zero. The pulling force is included by a potential $U_{pull} = F \cdot |\mathbf{r_N} - \mathbf{r_1}|$ where $\mathbf{r_1}$ denotes the position of the first and $\mathbf{r_N}$ the position of the last monomer in each fiber. Forces F are gradually increased and conformations are sampled at each value of the force. The mean end-to-end distances of the two fibers are then calculated from the sampled conformations. Upon induction of the stress, model chromosomes begin to restructure their internal organization, with regard to both, the cross-links and interlinks until they reach a new equilibrium situation.

Figure 5.6 and 5.7 show example conformations of chromosomes under tension. In Figure 5.6 the situation of non-tethered sister chromatids is displayed. It can be observed that both chromatids try to avoid contact with each other due to the entropic repulsive forces. In Figure 5.7 stretched sister chromatids that are tethered to each other can be seen. The cohesin-mediated bonds cause the fibers to be close to each other and to align. In both cases it can be observed that for intermediate forces, only an elongation of the model chromatids can be observed while for larger forces the average number of intra-fiber cross-links is reduced and sister chromatids become inhomogeneous.

The behavior of sister chromatids under tension is shown in Figures 5.8 and 5.9. As in the case of single chromatids, the stress-strain curve shows the characteristic behavior that was also observed in micromechanical experiments [43]. For small forces, a linear dependency between force and relative elongation can be observed for the chromosomes.

In this linear region, the average concentration of intra-fiber cross-links for both sister chromatids stays nearly unchanged. This means that for moderate forces, the chromatids are elongated but do not essentially change their average internal folding behavior. The elongation is also in part due to the straightening of chromosomes as well as the slight increase in average bond lengths between statistical segments.

Comparing the stress-strain curves between single chromatids and non-tethered sister chromatids shows that the slope in the linear region is different in both situations. In the linear elongation region, each of the model sister chromatids is an entropic spring with a certain spring constant. Two identical, parallel springs would then show the behavior of a spring with a doubled spring constant. This is not the case in our simulations. The elasticity for the double-chromatid system is increased by only approx. 50 % because we incorporate steric repulsion between the statistical segments of our model chromatids. This steric repulsion plays a role since it decreases the number of accessible conformations for two polymers that are very close to each other. Thus, it effectively changes the elasticity of the sister chromatid system.

When the sister fibers are tethered to each other by the dynamic linking mechanism, the slope in the initial linear region further decreases. This means that the Young's modulus for tethered fibers is smaller than that of untethered sister chromatids. In Figure 9 a close up of the linear region of the stress-strain curve is shown. We fitted the curves to determine Young's modulus Y which is given by

$$Y = \frac{\sigma}{\epsilon} \tag{5.3}$$

where $\sigma = \frac{F}{A}$ denotes the stress and $\epsilon = \frac{\Delta L}{L}$ denotes the strain. We observed that the presence of tethering between the sisters decrease the overall slope of this region. However, our results also show that this part of the stress-strain curve does in fact deviate from a linear relationship between force and extension. The cohesion between sister chromatids thus have a profound influence on the mechanical properties of chromosomes. Especially the level of cohesion between sister chromatids strongly influences the elasticity. We find that the Young's modulus decreases with increasing inter-sister link concentrations.

For large forces, the chromatids are not able to maintain the loop structure along their whole contour and the chromosomes become inhomogeneous as no intra-chain cross-links can form anymore in certain areas. Due to the high strain, each of the sister chromatids disintegrates as its internal loops dissolve. Thus, the chromatids can be extended without significantly increasing the pulling force resulting in a force plateau. The level of this force plateau is much lower in the case of the single model chromatid compared to sister chromatid systems. It is plausible that less force is needed to disintegrate a single chromatid than to disintegrate a system of two chromatids. More interesting is the observation that the force plateau decreases with increasing link concentrations. This means that sister chromatids that are connected to each other are also more easily disintegrated than unconnected sisters.

Figure 5.9 shows the corresponding link and loop concentrations as a function of the relative extension. In the linear force elongation region at small elongations, the loop concentration of single chromatids and unconnected sisters do not change. The loop concentration in connected chromatids even slightly increases. In the force plateau region, the loop concentration decreases rapidly as chromatids are pulled apart and the internal loop structure cannot be maintained along the complete chromosome anymore.

An interesting observation is that the concentration of interlinks strongly increases upon pulling chromosomes into the plateau region. This can be explained by the fact

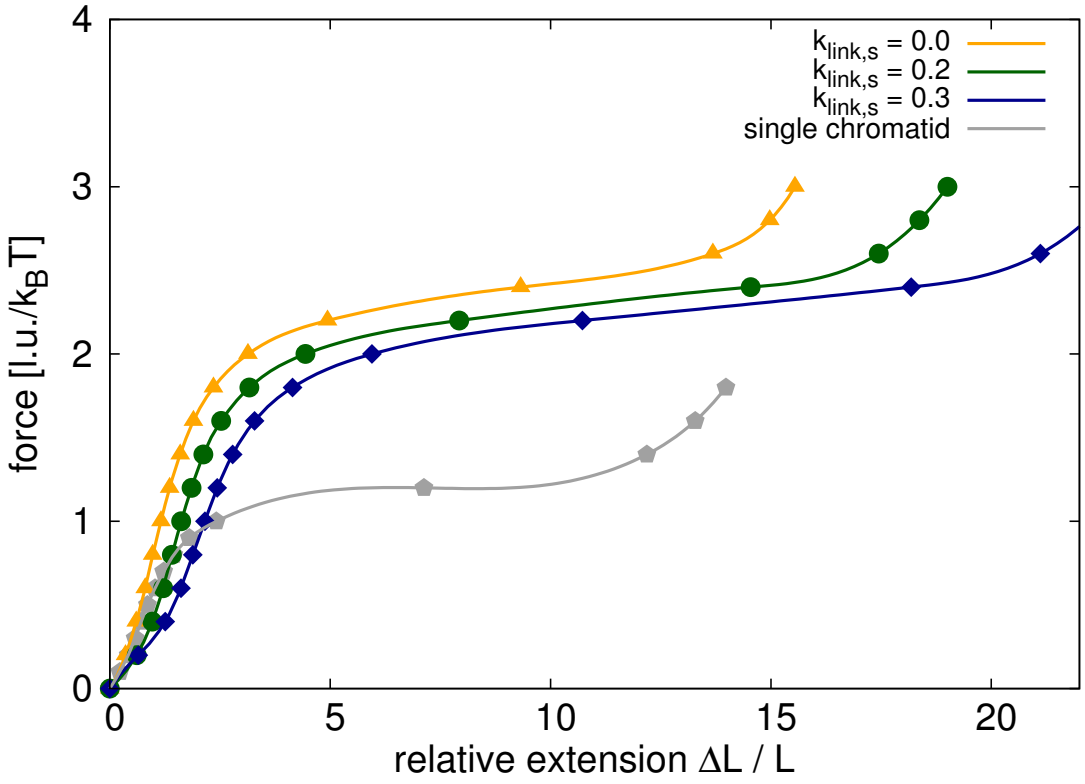

Figure 5.8: Force extension behavior for model sister chromatids. The figure shows the stress-strain curves for model chromatid systems with different link concentrations compared to the single model chromatid. The stress-strain curves all show a characteristic behavior with a linear elongation region for small extension followed by a force plateau for larger extensions. The slope of the curves in the linear region is proportional to Young's modulus which is a measure of elasticity. The Young's modulus for a system of two sister chromatids which are attached at the ends but nowhere else is higher than the Young's modulus of a single model chromatid. The Young's modulus decreases again, if the dynamic bonding mechanism between sisters is switched on and the link concentrations is increased. The plateau region for sister chromatids is significantly higher than for a single chromatid but again decreases with increased link concentrations.

that the strain facilitates the alignment of sister chromatids. In turn, aligned sister chromatids are easier to be bonded to each other by links that are created upon collision of chromosomal parts. In a configuration where the mean concentration of interlinks is high before the pulling starts, the increase of the mean concentration of interlinks is also high. For $k_{p,link,s} = 0.3$ (blue curve in Figure 5.9) the final concentration is $k_{p,link,f} \approx 1.4$ while for initial concentration of $k_{p,link,s} = 0.2$ the final concentration is $k_{p,link,f} \approx 1.0$. This shows that bonding between aligned sister chromatids could be strengthened upon physical stress.

For comparison we perform simulations for a simpler polymer model, the self-avoiding walk (SAW). The force extension behavior for a single SAW and for double polymer systems with tethered SAWs is shown in Figure 5.10. Simple polymers have obviously completely different force extension behaviors. In fact, for a Gaussian chain, which does not have excluded volume, the stress-strain curve is given by a Langevin function and the spring constant in the linear region is inverse proportional to the chain length. The inset in the panel shows the linear regions of the stress-strain curve for a single SAW and

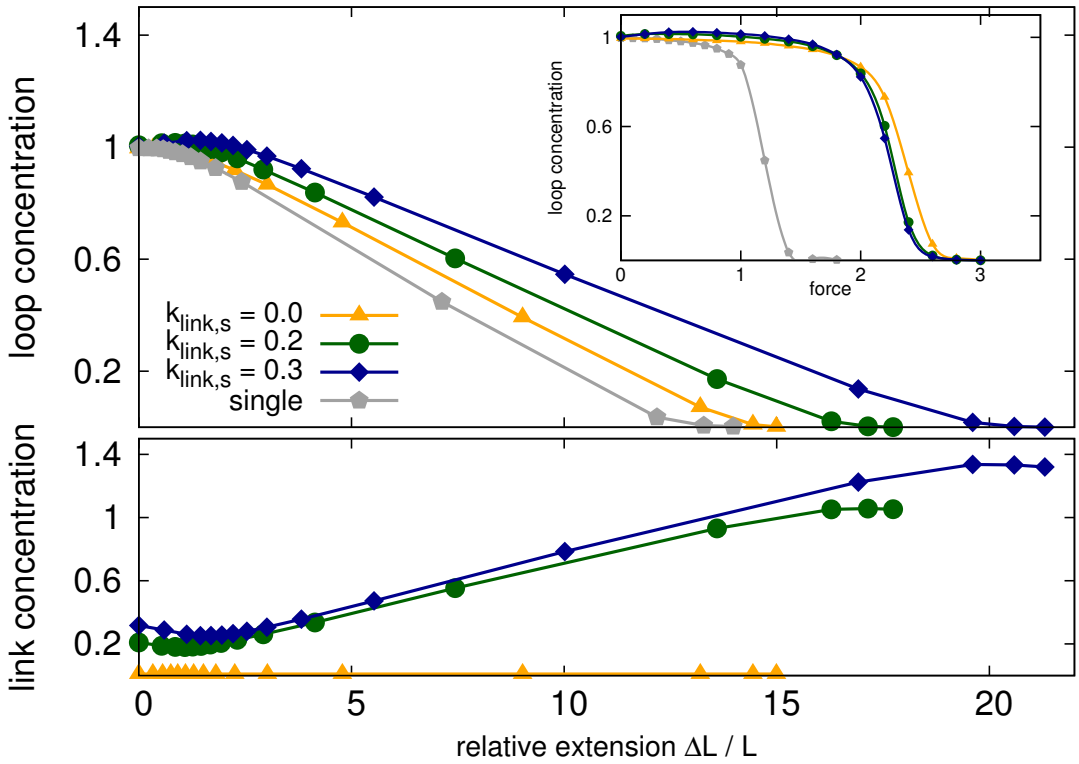

Figure 5.9: Force extension behavior for model sister chromatids. This figure shows the loop and link concentrations in the systems upon stress. The upper panel shows the loop concentrations within the chromatids in dependency of the relative extension. For a single chromatid and untethered sister chromatids, the loop concentration remains fairly constant in the region of small extensions and decrease for larger extensions. For tethered sister chromatids, the loop concentration in fact first increases slightly upon stress and then also decreases with larger extensions. The inset shows the loop concentrations in dependency of the applied pulling force. Here we show that for small forces the loop concentrations for single chromatids and untethered sister chromatids at small forces are exactly the same and only differ slightly in the force plateau region. However, the relative extensions at the same forces are quite different for single chromatids and sister chromatid systems.

two SAWs that are only tethered to each other at the chain ends. We also performed simulations where SAWs could dynamically bind to each other. The corresponding stress-strain curves are also shown in 8C. As in the case of our model chromatids, bonding also increases the elasticity for SAW systems. For high link concentrations, the two SAWs are also intermingled and the force extension changes its characteristics. Instead of a Langevin function-like behavior, we then first observe a initial sharper increase followed by a plateau area which then goes over to a Langevin-like tail for large elongations. Figure 5.11 shows the interlink concentrations depending on the relative elongation. It shows how upon small forces, the link concentration is reduced first because small forces de-mingle the SAWs. For large forces however, polymers are again brought to an elongated and aligned state where they can form links more easily and thus the link concentration increases again.

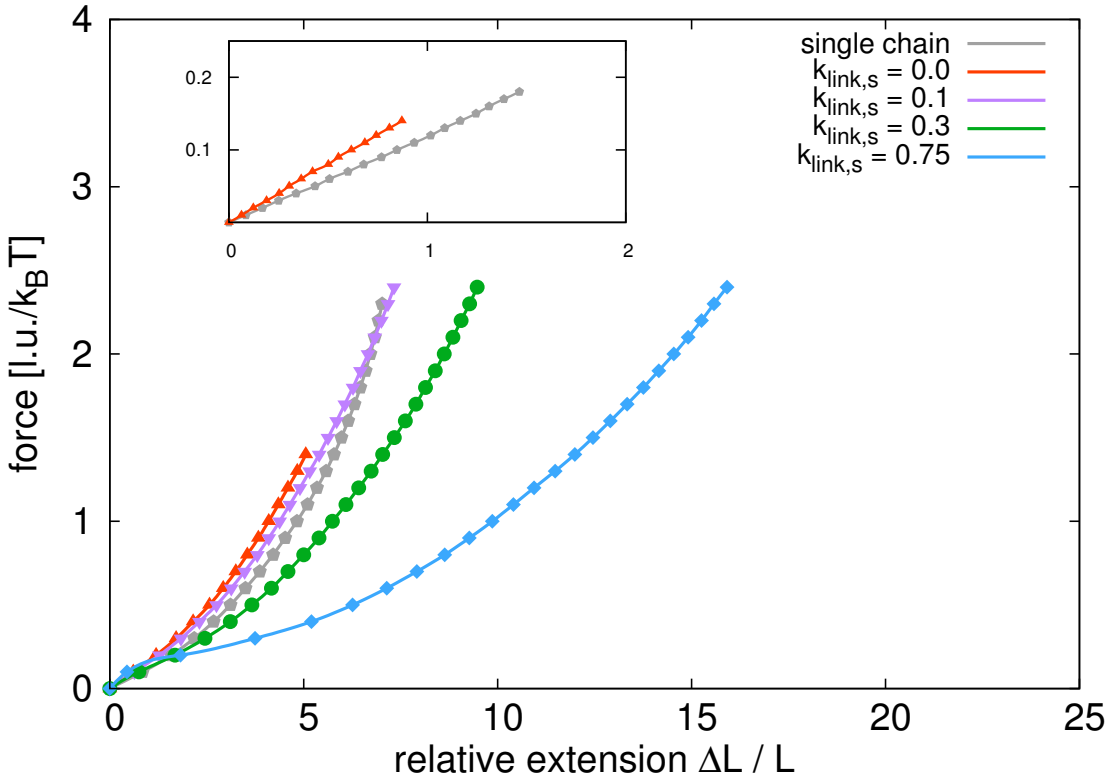

Figure 5.10: Force extension behavior for self-avoiding walks. For comparison with sister chromatids we also analyzed the force extension behavior of self-avoiding walks (SAWs) and systems of two tethered SAWs. The stress-strain curves have a different characteristics and especially do not show a long force plateau. In the initial linear region the entropic spring constant in the double chain system is higher than for the single chain but does not reach its doubled value due to excluded volume interactions between the chains.

5.4 Discussion

In this work we analysed how sister chromatid conformation in mitosis is governed by the interplay between condensation and cohesion. In our model, each individual chromatin fiber can dynamically form size-restricted loops which can result in its coiling into rod-like objects. Additionally, sister fibers can dynamically establish interlinks between each other leading to a mean number of bonds. We explored the parameter space for the looping probability p_{loop} and linking probability p_{link}. For each parameter setting we sample equilibrium conformations with Monte Carlo simulations. Depending on the looping and linking parameters our model yields different loop and link concentrations for the model fibers. We thus show that the combination of these two mechanism can result in vastly different conformational states of sister chromatids.

We were able to characterize the resulting conformations of the sister chromatid system by three main types. Firstly, there is a minimum threshold for the ratio of association rate and dissociation rate for links if sisters are to stay bonded. Below this threshold, the entropic repulsive forces between sister chromatids exceeds the effective binding force by the dynamic linking. Sister chromatids would then drift away from each other. Furthermore, we found that in order to obtain a system of two clearly distinguishable chromatids

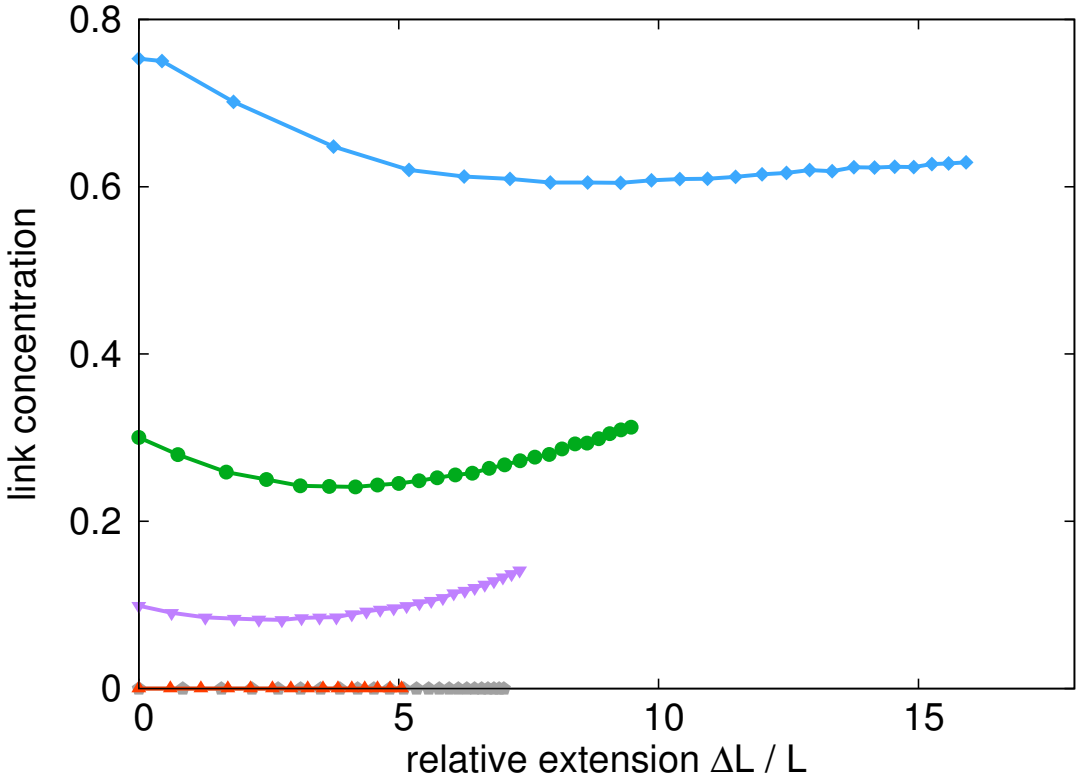

Figure 5.11: Force extension behavior for self-avoiding walks. The link concentration in dependency of the relative extension decreases first for small extensions and increases again for larger extensions. It has therefore the same tendency as for model chromatids but is not as pronounced due to the lack of internal loops that can be unfolded.

there must be a cap in the mean concentration of links. For higher link concentrations, model sister chromatids are completely intermingled and not distinguishable from each other. From our results we can conclude that the mean number of links by which sister chromatids are bonded together has to lie within a sensitive region.

In this work, the mechanisms for looping of the chromatin fiber and for linking of sister chromatids are effective mechanisms that model the presence of binding proteins such as condensin and cohesin. However, we have to stress that the detailed binding mechanisms of these proteins are still under debate. Therefore we choose a probabilistic model for the effect of binding. Our model parameters p_{loop} and p_{link} effectively describe the binding affinity of fiber segments to each other. This affinity could be altered for example by different protein concentrations. In fact a recent model that explicitly includes diffusing proteins as binding partners for the chromatin fiber found that increased protein concentrations lead to higher number of binding points [205].

A number of studies have shown that genome-wide cohesion between sister chromatids can be established as a reaction to DNA damage by exogenous agents such as irradiation [200, 202]. This damage-induced cohesion could facilitate the homologous recombination repair pathway by tightly holding the parts important for repair together. Here we show that an increase in the number of bonding regions between sister chromatids also results in their intermingling which makes it impossible for each chromatid to condense into a rod-like shaped object. However, it is evident that this condensation is crucial for

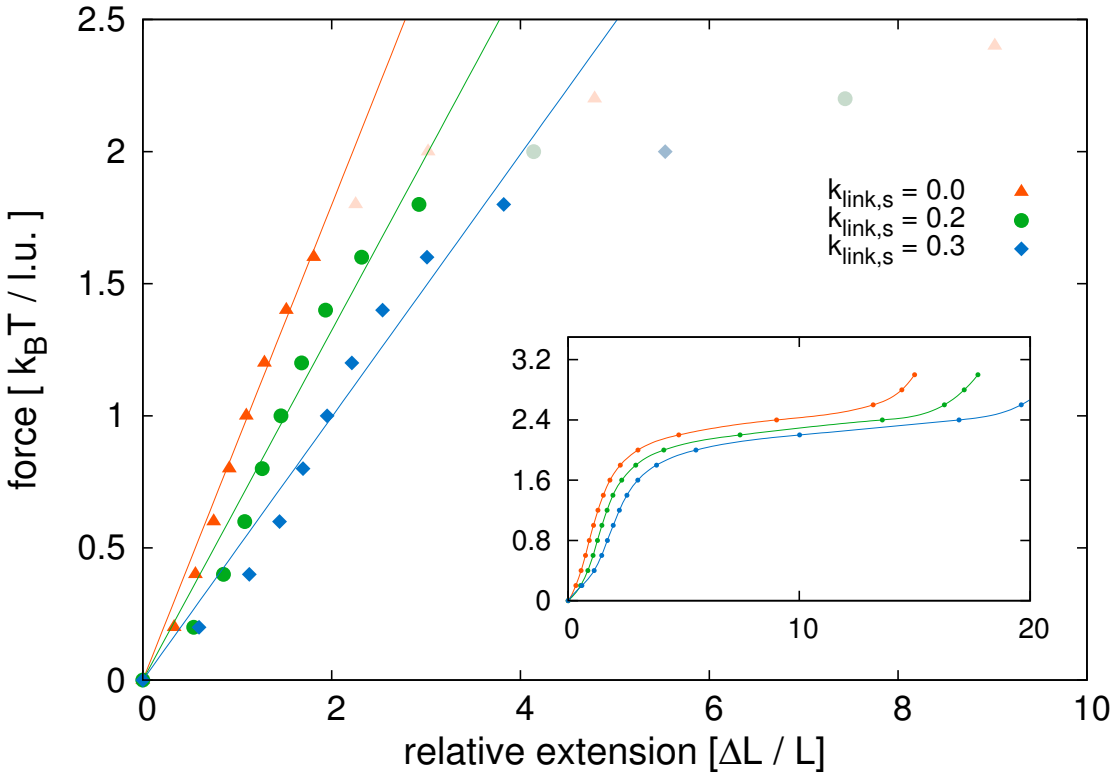

Figure 5.12: Initial region of stress-strain curve. Shown is a close-up view of the linear region of the stress-strain curves. We compare model sister chromatids with different link concentrations. We fitted the initial parts of the curves before the plateau area with linear functions. The dark points are those used for the fits, while the points in light color belong to the force extension curves but were not considered for the linear fits. The results show that non-tethered model chromatids have the highest Young's modulus and the modulus decreases with increasing link concentrations. The values for the Young's modulus are: $Y_1 = 0.90\,k_BT/l.u.$ for $k_{link,s} = 0.0$, $Y_2 = 0.66\,k_BT/l.u.$ for $k_link, s = 0.2$ and $Y_3 = 0.50\,k_BT/l.u.$ for $k_{link,s} = 0.3$.

chromosome segregation in mitosis since intermingled chromatids are hardly distinguishable. We therefore speculate that tight bonding of sisters upon formation of double-strand breaks (DSBs) prior to mitosis could also be a physical mechanism for cell cycle arrest since it inhibits the progression of chromosome condensation. This might also be a reason why one single DSB could trigger the establishment of cohesion in the whole genome.

Our simulations of the behavior of sister chromatid systems upon external stress shows that it is qualitatively the same as for single chromatids. The stress-strain curve shows a initial linear region which is followed by a broad force plateau. In the linear region a spring-like behavior is observed and the force plateau is a decondensation region where the integrity of chromosomes is destroyed by external force. The emergence of force plateaus for large elongations has been observed in many experimental studies before [43]. These experiments included single chromatids that were extracted from eggs [112] and chromosomes consisting of two chromatids extracted directly from cells [96, 117]. In our present work we performed pulling simulations for single chromatids and also for bonded sister chromatid systems. We then compared the behavior of the two systems in order to obtain a better understanding of how experimental results for these could differ from each

other.

Our results show that the required force to reach the plateau region is much higher for bonded sister chromatids than for single chromatids. However, when comparing bonded sister chromatids to non-bonded ones, we observe a decrease of the force plateau. Since the force plateau indicates the region where chromatids are disintegrated by the pulling force, this means that bonded sister chromatids are more easily unfolded by pulling forces. The level of the plateaus decreases with increasing number of bonds between sisters. An explanation for this could be that by being coupled to each other, pulling forces that act on one chromatid are also able to act on the other one. By this dragging effect, a force that is able to elongate one chromatid and thus prevents the formation of loops in this chromatid, could then prohibit the formation of loops in the sister chromatid, too. This mechanism could also be responsible for the decreased slope of the stress-strain curve in the region before the force plateau. Another factor could be that model sister chromatids are aligned in the pulling process. This alignment further facilitates the formation of bonds between them which in turn decreases the possibility of loop formation. Thus we can conclude that the amount of inter-sister cohesion can play a role for the mechanical properties of the chromosome. The differences of the mechanical properties of chromosomes in experimental studies could then be due to different amounts of cohesion between the sisters [43, 96].

Chapter 6

Structural Changes of Chromosomes in Irradiated Cells

References

The content of this chapter was adapted from the following publication

- Y. Zhang and D.W. Heermann, *DNA double-strand breaks: linking gene expression to chromosome morphology and mobility.*. Chromosoma *Epub ahead of print.* doi: 10.1007/s00412-013-0432-y. 2013

Chapter Summary

Ionizing radiation can lead to DNA double-strand breaks (DSBs) which belong to the most dangerous forms of damage to the DNA. Cells possess elaborate repair mechanisms and react in a complex manner to the emergence of DSBs. Experiments have shown that gene expression levels in irradiated cells are changed, and thousands of radiation-responsive genes have been identified. On the other hand, recent studies have shown that gene expression is tightly connected to the three-dimensional organization of the genome. In this work, we analyzed the chromatin organization in the cell nuclei before and after exposure to ionizing radiation with an expression-dependent folding model. Our results indicate that the alteration of the chromosome organization on the scale of a complete chromosome is rather limited despite the expression level change of a large number of genes. We further modelled breaks within sub-compartments of the model chromosomes and showed that entropic changes caused by a break lead to increased mobility of the break sites and help to locate break ends further to the periphery of the sub-compartments. We conclude that the changes in the chromatin structure after irradiation are limited to local scales and demonstrate the importance of entropy for the behaviour of break ends.

6.1 Introduction

Exposure of tissue to ionizing radiation leads to DNA double-strand breaks (DSBs) which belong to the most dangerous forms of damage to the genome. Cells show complex reactions to resolve this kind of damage, including the genome-wide regulation of transcriptional activity. Experiments found large numbers of radiation-responsive genes, which are genes that significantly alter their gene expression activity upon irradiation with ionizing radiation [50, 51, 206–208].

Transcriptional activity along the human genome is anything but homogeneous. Instead, alternating regions of high and low transcription can be found. These regions not only differ in transcriptional activity, but also in other quantities such as gene density and base composition [209]. Classically, actively transcribed regions are referred to as euchromatin, and silent domains as heterochromatin. Studies showed that while heterochromatin is highly compacted, euchromatin takes on a much looser organization organisation [204, 210, 211].

On the other hand, it has become very apparent that entropy plays a major role in the physical organization of the eukaryote genome [31, 32]. Especially, the formation of chromatin loops unfolds entropic conditions that can aid to organize the genome [37, 186]. Furthermore, many experimental studies have documented that genome organization is tightly connected to the looping of the chromatin fibre. Biochemical experiments based on the identification of chromosomal contacts, namely the chromosome conformation capture technology and its variants, showed that functional loops can be found on all scales in eukaryote cells [178–180]. Such physical contacts, for example between promoters and enhancers, could play an important role in gene regulation [14, 212]. On a larger scale, chromosomes in the nucleus occupy distinct chromosome territories that intermingle only little with each other [73]. Additionally, it has been shown that gene expression takes place in transcription factories where different chromosomal parts which can also be from different chromosomes come into physical proximity of each other to be transcribed simultaneously [87, 90].

Many recent works have targeted the physical modelling of chromosome organization in the cell nucleus. Early polymer models including chromatin loops feature random walk polymers with attached loops [213] or a rosette-like kind of organization [93]. Other works examined how unspecific interactions can guide self-organization of the chromosome and co-localization of chromosomal segments [31, 214, 215]. Approaches to reconstruct genome organization from contact maps of chromosome conformation capture experiments let to the suggestion of the fractal globule as a model for interphase chromosomes [180]. The Random Loop and the Dynamic Loop Models suggested dynamic looping of the chromatin fibre and described the globally confined folding of chromosomes without external confinements [183, 184]. A similar approach with explicit binding molecules that facilitate looping of the chromatin fibre yielded similar results [205]. Recently, an expression-dependent folding model was able to show how loops promote compartmentalization of the chromosome into transcriptionally active and inactive regions [216]. The β-globin locus is a very good example for the importance of chromatin loops for activation and inactivation of gene transcription [215, 217].

The exact repair of DNA double-strand breaks is crucial for the cell. Misrepaired or not repaired DSBs can lead to chromosome translocations and subsequently to carcinogenesis [27]. Repair proteins play an important role for the resolution of DSBs. For example, the proteins of the MRN complex, which are usually scattered over the whole nucleus, can accumulate at the site of a DSB within minutes [218]. This is further accompanied

by local chromatin restructuring in the surroundings of the break [125, 219]. Such foci can be made visible by fluorescence microscopy [126]. Post-translational modifications of histone proteins in the surroundings of the DSB are also very important. Especially, the phosphorylation of histone variant H2AX at serine 139 resulting in γ-H2AX is a typical marker for DSBs [49,129] and is believed to cause further rearrangements of the chromatin structure [128,132]. Using fluorescence microscopy, a relaxation of the chromatin structure at these γ-H2AX foci was observed [133].

A very important aspect of the DSB signalling and repair is the mobility of the DSBs. The goal of repair is to mend the correct ends of the broken strand with each other again. Increased mobility of DSBs could help them find each other faster since the roaming area becomes larger. Indeed, enhanced mobility of sites of DSBs was observed in a recent experimental study [136]. Other works showed enhanced DSB mobility in the homologous recombination repair pathway in budding yeast [220, 221]. On the other hand, increased roaming areas of the broken ends also enhance the possibility of two lesions interacting with each other which can lead to misrepair and chromosome aberrations.

The controversy about the repair mechanisms of the highly radio-resistant bacterium *Deinococcus radiodurans* exemplifies how DSB mobility and structure influences their repair. *D. radiodurans* has a remarkable resistance to radiation-induced DNA damage and is able to survive extreme doses in the range of thousands of Gray that produce hundreds of DSBs [222]. There is a long standing controversy about the main repair mechanisms leading to *D. radiodurans* resistance. Levin-Zaidman et al. [223] suggest that DSBs possess very restricted motion due to high compaction of the chromatin fibre and, thus, non-homologous end joining as a possible repair pathway. However, other works advocate homologous repair which would require the chromatin fibre of one genome to penetrate the territory of other genomes in order to find homologue templates [224].

Multiple works have shown that differences exist between the repair of DSBs in heterochromatin and euchromatin in eukaryote cells. Îş-H2AX foci were found significantly less often in highly compacted and silent heterochromatin than in active euchromatin, which could be an indicator for repressed repair in the heterochromatic regions [140,141]. Furthermore, DSBs within the heterochromatic regions were found to be moved to the periphery of the domain in *Drosophila melanogaster* [143]. A very similar kind of behaviour was also observed in human and murine cells, where Îş-H2AX foci bent around heterochromatic regions instead of being located in a linear track after ion irradiation [142].

DSBs do not only form after exposure to ionizing radiation, but rather there are situations when strand breaks are in fact part of the genomic program. Recent works have shown that DSBs could play an important role at the regulation of transcription itself [225,226]. At this, it has been speculated that transient DSBs may be necessary in order to untangle the chromatin fibre and require the enzyme topoisomerase IIβ. Furthermore, such transient DSBs were shown to be also recognized by the cell's repair machinery [225,227]. This can be seen as another hint that recognition of DSBs is dependent on globally effective and stable motifs such as entropy.

In this work, we use an expression-dependent folding model featuring dynamic looping of the chromatin fibre to Chromosoma model interphase chromosomes that were exposed to ionizing radiation. We first show by means of this kind of modelling that no structural changes on the global scale of a complete chromosome can be expected after irradiation with ionizing radiation. We characterize different environments within the model chromosome territory, which arise from compartmentalization of different domains in the chromatin fibre. The expression-dependent model predicts distinct folding compartments

for domains that are highly active, often identified as active euchromatin, and domains that are silent, often identified as heterochromatin. We then model explicit strand breaks within the characterized environments to analyze how entropic changes, due to the arising of free ends, affect the dynamics of the broken parts and what structural changes they cause within the environment.

Our results show that entropic changes caused by the emergence of a break are able to explain enhanced mobility of broken DNA ends in all kinds of environments. Furthermore, we confirm the importance of loops in the chromatin fibre by showing that they efficiently prevent the broken ends from drifting away from each other. Finally, it is found that entropy aids at shifting the location of broken ends towards the periphery of chromosomal sub-domains. Our work, therefore, demonstrates how entropy governs the structure and dynamics of chromosomal domains in the DSB repair.

6.2 Methods

6.2.1 mRNA Microarray Data Preparation

We downloaded mRNA microarray data on immortalized B cells from the Gene Expression Omnibus database with accession number GSE12626 that was published in Smirnov et al. [51]. The cells were exposed to $10\,Gy$ of ionizing radiation and harvested 2 and $6\,h$ later. For the detailed procedures, see the original publication [51]. We used the *Bioconductor* and *Biomart* packages to match the probe IDs in the data to *Ensembl* gene IDs (*www.ensembl.org*) and, in turn, to positions on the genome. By this, we obtained values for $13,502$ sequences on all chromosomes.

For the further study in this work, we focused on chromosome 11 for which we obtained 791 sequences. We calculated moving averages along the genome by averaging over 24 genes upstream and 24 genes downstream for each mapped gene (*MM49*) as in Goetze et al. [211]. We, thus, obtained a genomic profile of the average transcriptional activity in chromosome 11. This profile was then binned into a histogram with $1,350$ bins. All values that contributed to one bin were averaged according to their proportion of the bin. Thus, we were able to assign a value that represents the average transcriptional activity to each 100 kb segment of chromosome 11. This value was subsequently used in the chromatin folding model to determine looping probabilities between different chromosomal parts.

6.2.2 Expression-dependent Dynamic Loop Model

We use a coarse grained approach to model the chromatin fibre in the interphase nucleus. The model fibre is built by repeat units, each representing a statistical segment that consists of $100\,kb$ of DNA and proteins. The coarse grained approaches are justified since the global structure does not depend on the detailed organization inside of each statistical segment. The segments are connected to a flexible chain and with a steric repulsion between each other to realize excluded volume interactions. Human chromosome 11 was, thus, modelled by a chain consisting of $N = 1,350$ monomers.

The loops in the interphase chromosomes play a crucial role for genome function and structure. Their presence creates entropic constraints that are necessary to maintain chromosome structure. In our model, we include chromatin loops by using a dynamic looping mechanism of the model fibre. When two monomers come into physical proximity to each other by diffusion, an additional bond that we refer to as cross-link can be created between them, resulting in a loop. A lifetime of the loop is drawn from a Poisson distribution

with mean value τ. The cross-link dissolves again after this lifetime, and thus, the loop vanishes. By this dynamic mechanism, there is a constant association and dissociation of non-adjacent monomers, resulting in loop creation and dissolution.

In the expression-dependent Dynamic Loop Model [216], we connect the feature of dynamic looping with the observation that chromatin organization is more compact in transcriptionally non-active heterochromatin and more loose in highly active euchromatin. We consider this phenomenological observation by assigning different probabilities for the formation of loops to the fibre segments. Using the gene expression profile gained from the mRNA microarray data as described before, we calculate a number p_i representing the transcriptional activity for each statistical segment of the model chromatin fibre. The probability of two segments i and j to form a cross-link upon random collision is then given by $p_{cross-link} = p_{base} \cdot \frac{1}{2}(p_i + p_j)$, with p_base being a base probability. A more detailed description of the model can be found in Jerabek and Heermann [216].

We sample conformations using gene expression data for untreated human B cells and for human B cells 2 h after exposure to 10 Gy of ionizing radiation (data from Smirnov et al. [51]). The sampling is performed using a lattice Metropolis Monte Carlo method based on the well-established Bond Fluctuation Model [163,164]. The independence of the conformations was ensured by considering the autocorrelation time in the Monte Carlo simulations. In total, around $10,000$ independent conformations were sampled for each setup to get a large enough set that can represent the statistical ensemble. Observables were calculated for each of the conformations first and then averaged over the whole ensemble. Further details on the simulations can be found in previous works [184,203,216].

6.2.3 Simulation of DSBs in Model Chromosomes

We investigated further the behaviour of strand breaks in different environments inside of the chromosome territory. The experimental studies have indicated, for example, increased mobility of sites of DSB compared to "usual" chromatin. To study such kind of behaviours in our model, we first assumed that the time scale of a global reorganization of the chromosomal folding including changes in the loop structure is much higher than the interesting time scales for the DSB repair. This assumption is based on studies that have shown that, for example, γ-H2AX can be found already minutes after the exposure of cells to ionizing radiation, and γ-H2AX and other repair proteins forming such foci are believed to help stabilize the strand breaks.

Therefore, we used the sample conformations gained from the simulation for untreated cells and fixed the loop structure. Thus, the cross-linked monomers were now permanently tethered to each other. We induced one break in each of the conformations by deleting the bond between two adjacent monomers. We carried out different simulations for break positions in different environments, e.g. heterochromatin or euchromatin. We subsequently studied the dynamics of all parts of the chromosomes by means of molecular dynamics simulations. At this, we used a repulsive Weeks-Chandler-Anderson potential to include steric repulsion between the monomers [228]:

$$U_{ij}^{WC} = \begin{cases} 4\epsilon \left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6} + \frac{1}{4} \right], & r_{ij} \leq \sqrt[6]{2}\,\sigma \\ 0, & r_{ij} > \sqrt[6]{2}\,\sigma \end{cases} \tag{6.1}$$

where i and j denote monomers and r_{ij} their spatial separation. In our simulations, σ and ϵ were set to the value 1. FENE potentials were used to connect adjacent monomers to each other:

$$U_{ij}^{FENE} = -\frac{1}{2}Kr_{max}^2 \ln\left\{1-\left(\frac{r}{r_{max}}\right)^2\right\} \tag{6.2}$$

with r_{max} representing the maximal distance between two adjacent monomers. We used the values $K = 30$ and $r_{max} = 1.5$ to avoid bond crossings [229]. Cross-linked monomers were tethered to each other via this potential, too.

The Molecular Dynamics package *ESPResSo* [166] was used to carry out Langevin dynamics simulations of the setup. We used a Langevin thermostat to set the temperature to a value of $k_BT = 1.0\,\epsilon$. The simulation was run in a simulation box with periodic boundary conditions, so the model chromosome could move freely without any spatial constraints. Integration was performed for a time in which typically the mean squared displacement of the centre of mass was more than ten times the average radius of gyration. The monomer coordinates were recorded in equidistant time intervals and later used to calculate the mean squared displacement of each monomer. By performing this simulation for each of the independent conformations gained from the Dynamic Loop Model, we get the behaviour of strand breaks in the whole conformation space.

To further rule out that results are different for broken ends within sub-domains that are dynamically restructuring their loops and contacts, we also simulate strand breaks in subdomains for which the loop structure is not fixed. We, therefore, conduct simulations for DSBs in domains in which loops can further form and dissolve dynamically. The breaks were induced by removing the bond between the two monomers of the model chromatin fibre. At this, we chose the same monomers as in the simulations with fixed loop structure. We then sampled the equilibrium conformations of the broken strands with the same Monte Carlo method that we used to sample the chromosome conformations without breaks. We found that the behaviour of broken ends in dynamically restructuring model chromosomes is qualitatively the same as in chromosomes with fixed loop structures. The results for these simulations can be found in the *Supplementary Information*.

6.3 Results and Discussion

6.3.1 Model Chromosomes with Expression-dependent Interactions

We model a single chromatin fibre by a polymer consisting of N equal repeat units connected to a linear chain including steric repulsion between these monomers. Each repeat unit represents a blob of 100 kb of coiled DNA/proteins. When no further interactions between the repeat units exist, the model chromosomes adopt rather swollen, non-coiled conformations known as self-avoiding walks [149]. To include the effect of loop formation, we employ an expression-dependent looping mechanism for the chromatin fibre as in [216].

For our model, we use mRNA gene expression data on human immortalized B cells from the study by Smirnov et al. [51]. In the study, the cells were exposed to $10\,Gy$ of ionizing radiation, and expression levels were measured at 2 and $6\,h$ after irradiation [51]. We use the data to calculate the average transcriptional activity for each $100\,kb$ segment of chromosome 11, thus gaining a transcription profile for chromosome 11 with $1,350$ values. We then assign these values to the corresponding monomers that represent the same $100\,kb$ segments in the model chromatin fibre.

In the model, monomers diffuse according to stochastic dynamics [184]. Upon collision of two non-adjacent fibre segments, an additional cross-link can be formed between them with a probability that is dependent on the average transcriptional activity of both

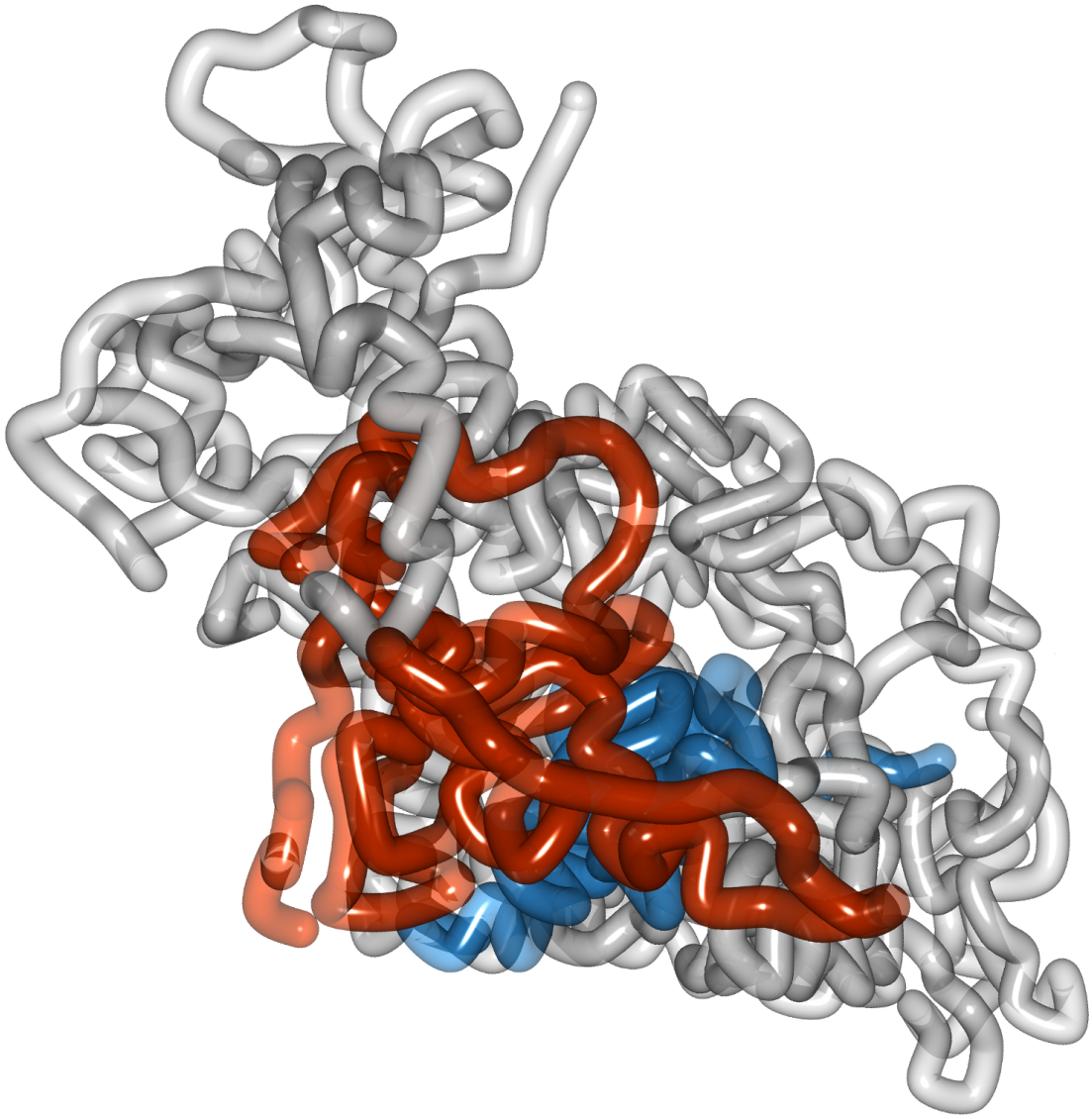

Figure 6.1: Example conformation of chromosome 11. The figure shows a conformation of chromosome 11 generated with the expression-dependent folding model. The grey tube represents the coarse grained model chromatin fibre. An actively transcribed region is coloured in red, and a silent region is coloured in blue. In this example conformation, it becomes apparent how the more active region tends to be oriented further to the outside, while the silent region is closer to the centre of the chromosome territory. Note, however, that this is only one possible conformation and that the described positional preference is an observation based on the study of the whole ensemble of possible conformations, i.e. a population average over cells.

monomers. This expression-dependent looping probability accounts for the phenomenological observation that areas with high transcriptional activity have a different folding behaviour than areas with low transcriptional activity. The different folding behaviours can be explained by different concentrations of loops within the different areas. Highly active regions have smaller loop concentrations than non-active regions [183]. This observation motivates the coupling of the looping probability of a statistical segment in the chromatin fibre to its average transcriptional activity in our expression-dependent folding model. More details of the model can be found in a previous work [216].

The above outlined model is used for the data of untreated cells and cells at $2\,h$ after irradiation from the Smirnov et al. study [51]. Conformations are sampled with a Metropolis Monte Carlo algorithm based on the Bond Fluctuation Model [184, 216].

6.3.2 Global Genome Organization of Post-irradiation Cells does not Differ Significantly

We compare the simulation results for chromosome 11 without irradiation with the results for irradiated chromosome 11. We first investigate if any changes of the global structural properties of the model chromosomes can be found. An example conformation for unirradiated chromosomes is shown in Figure 6.1.

The grey tube represents the model chromatin fibre; the region coloured in red is more active than the region coloured in blue. In the example conformation, it can be seen that the transcriptionally more active region is located further to the periphery of the chromosome territory than the less active region.

In Figure 6.2, the different transcription profiles for irradiated and unirradiated cells are shown on the lower panel. Additionally, the mean chromatin density and the mean squared distance of each chromatin segment to the centre of the mass of the chromosome are displayed in the upper panels. The general tendency that can be observed is that transcriptionally active regions, often referred to as euchromatin, have, on average, a low density due to a low number of loops. These regions tend to be located further away from the centre of the mass or, in other words, closer towards the surface of the chromosome territory. In contrast, silent areas of the chromosome, often identified as heterochromatin, are much more compact and, therefore, dense. These areas are located closer to the centre of the chromosome territory due to entropy.

The results for the density and the positioning of chromosomal sub-domains show that the physical structure of our model chromosomes is highly dependent on the transcription profile. This is due to the fact that the looping probability of a segment and, thus, the mean loop concentration in its vicinity is directly dependent on the transcriptional activity of this segment. In turn, more loops in a region means more contacts between non-adjacent chromatin segments and results in a higher chromatin concentration. As the lower panel in Figure 6.2 shows, the transcription profile for cells prior to irradiation and $2\,h$ after irradiation is widely similar despite the expression change of a large number of single genes. The coarse grained modelling has an averaging effect, where single genes with altered expression levels do not have a great impact on the average expression level of a whole statistical segment.

This is also confirmed by our results for the relationship between the physical distance and the genomic distance, and the distribution of contact probabilities. These results are shown in Figure 6.3. The mean squared displacement of the model chromosome segments depending on their genomic separations show the characteristic levelling off for genomic distances above $10\,Mb$ as has been observed in the fluorescent in situ hybridization (FISH)

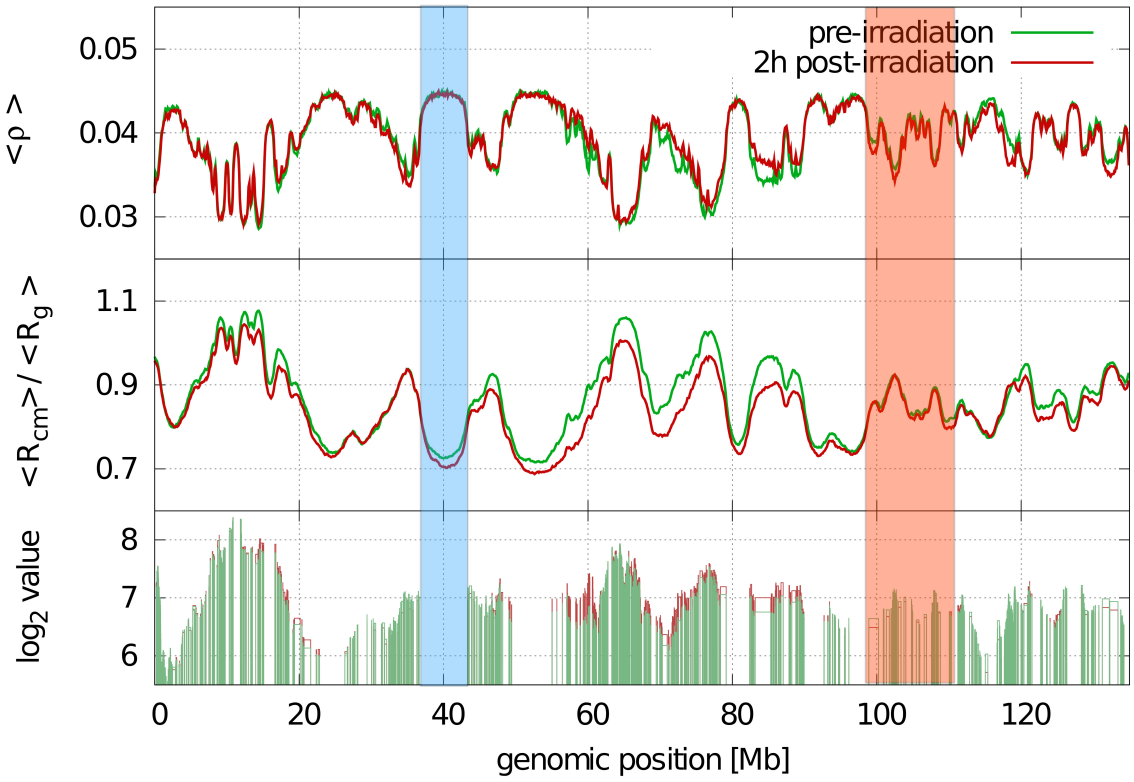

Figure 6.2: Expression profile and physical structure of model chromosomes. In this figure, the relationship between the physical structure of the model chromosome and the expression profile can be seen. In the bottom panel, the average transcriptional profile computed with the sliding window algorithm for chromosome 11 is showed. The top panel shows the mean density in a sphere of radius 1.5ĬČ around each statistical segment. The results demonstrate that in transcriptionally active regions, the density is low, while less active regions are more compacted. In the centre panel, the average distance of each statistical segment to the centre of mass of the chromosome territory is displayed. It clearly shows that active regions tend to be positioned further to the periphery of the territory, while inactive regions are compact and located to the centre of the territory.

experiments and explained with the Random Loop Model before [183]. In our study, we find that the genomic distance at which the levelling off starts is very close for chromosomes before and after exposure to ionizing radiation. Moreover, the level of the plateau for pre- and post-irradiation model chromosomes is approximately the same, apart from some smaller deviations at large genomic separations. We furthermore examined the distribution of the contact probabilities between non-adjacent statistical segments. Here, too, we observe that irradiated and not irradiated model chromosomes have the same behaviour, as shown in the inset of Figure 6.3. In the region between approx. $500\,kb$ and $2\,Mb$, the contact probability obeys a power law with exponent $\gamma = $ âĹŠ1, while for larger genomic distances, the distribution becomes more flat.

Our results show that irradiation of the cell does not significantly alter genomic organization on the scale of a complete chromosome. The density and relative position profiles of loci on chromosome 11 stay very similar after exposure to ionizing radiation. This is directly a consequence of the fact that the transcription profiles of not irradiated and irradiated cells are very similar. We, thus, conclude that even the expression level change of

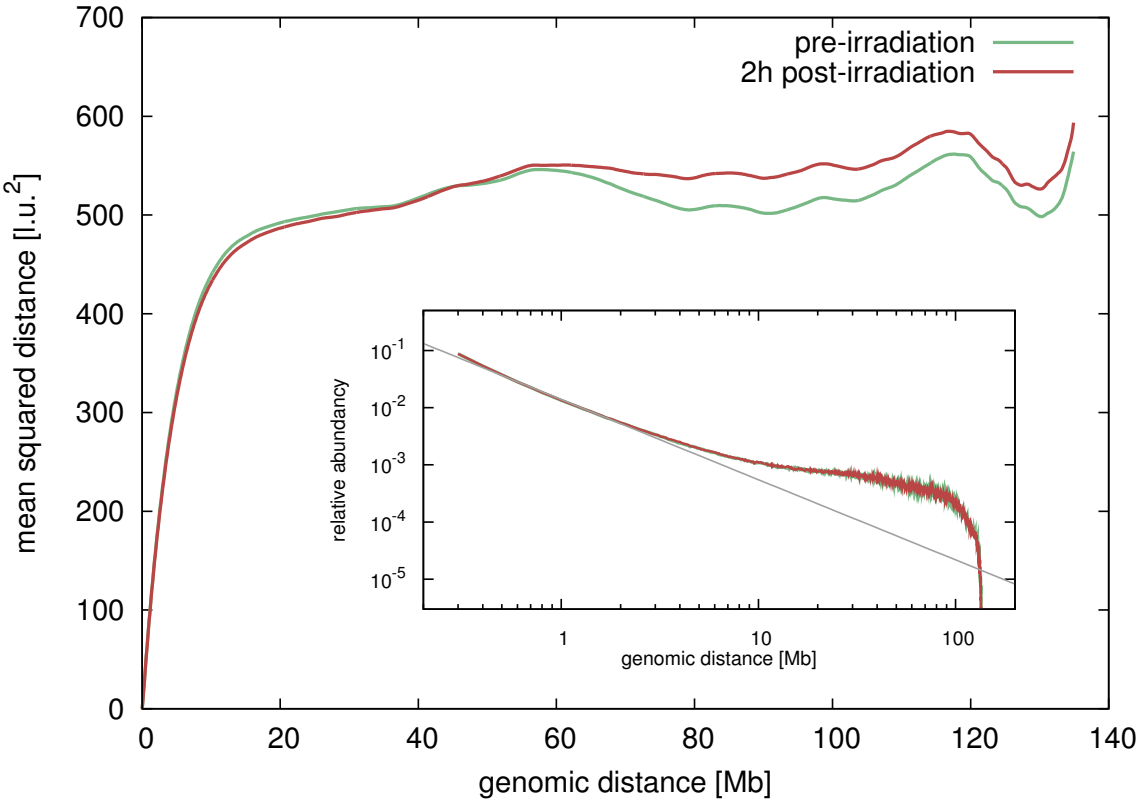

Figure 6.3: Dependency of the mean squared physical distance and contact probabilities to the genomic distance of chromatin fibre segments. The figure shows the mean squared physical distance between model chromosome segments in dependency of their genomic separation from each other. Instead of growing with a power law as expected for ordinary polymer models such as the self-avoiding walk, the curve reaches a plateau from around 10 Mb. At this, model chromosomes prior to irradiation at 2 h after exposure to radiation show a very similar behaviour. The start of the plateau regions are at the same genomic separation distances and the plateaus are at the same heights, i.e. the same physical distances, apart from some smaller deviations at long genomic separations. The inset shows the contact probability of two segments against their genomic distance. A fit to the region between 500 kb and 2 Mb gives a power law with exponential y = -1. The curves for preirradiation and post-irradiation model chromosomes are in agreement.

a large number of single genes does not significantly change the overall genomic structure in the nucleus. It should be noted that this conclusion is based on the average behaviour of the chromosome, since it is necessary to study a large sample of possible conformations in the simulations, just as in experimental studies. Certainly, general cell-to-cell variability as well as the individual reaction of the chromosome in a single specific cell can be different.

We would like to stress that we do not wish to suggest that there are no changes in the physical structure in chromosomes after irradiation. Rather, we state that there are no dramatic structural changes on a coarse grained level. In particular, we do not observe any large-scale conformational reorganizations, which would become visible through altered density and position profiles. We, thus, conclude that large-scale rearrangements of the chromosomal organization are rare and unlikely, and structural changes that help, for example with signalling and repair of DSBs, have to be found on smaller scales. We suggest that our predictions could be experimentally tested by direct measurements of the physical structure in irradiated cells. For instance, FISH microscopy images could be used to determine the dependency between the physical distance between loci and their genomic distance. It could then be verified if this dependency changes in cells that were exposed to ionizing radiation.

6.3.3 Double-strand Breaks Show Increased Mobility due to Reduced Entropic Constraints

We subsequently concentrate on the entropic changes that a double-strand break causes in its local environment. To assess how a double-strand break can affect the local environment, we first investigate how the mobility changes when a double-strand break occurs. For this, we measure the mean squared displacement of different chromosomal parts. We first analyze the mobility of each model chromosome segment of the unirradiated chromosome 11.

Figures 6.4, 6.5 and 6.6 show the results for the mean squared displacement with time for two regions, one in silent chromatin and one in a region with high average transcriptional activity. The results demonstrate how dependent the mobility of a section is on the transcriptional activity. Clearly, the higher number of loops and thus more topological constraints cause a more restricted movement of segments that are located in more dense environments. The steric repulsion in areas with high density also contributes to this. Thus, chromatin in silent heterochromatic regions has a much lower mobility.

We then take the sample of conformations for the unirradiated chromosome 11 and cut the model chromatin fibres at the same position in each of the conformations. For subsequent analysis of the dynamical behaviour of the chromatin fibre especially at the site of the strand break, we conduct molecular dynamics simulations for the conformations. We assume implicitly that the time between the occurrence of the strand break until the arrival of repair proteins and initiation of repair is much shorter than the time for a global conformational change of a chromosome. Therefore, we fix the loop structure in the chromosome during the molecular dynamics simulation and monitor the mean squared displacement of all segments including the broken ends. We conduct simulations for two different break positions, with the first being located in highly active euchromatin and the second silent heterochromatin. We simulate strand breaks in all conformations that were previously sampled with the Dynamic Loop Model. Thus, we consider the behaviour of broken ends in environments with all different possible kinds of loop structures.

The fixed loop structure is motivated by the fact that repair proteins were shown to

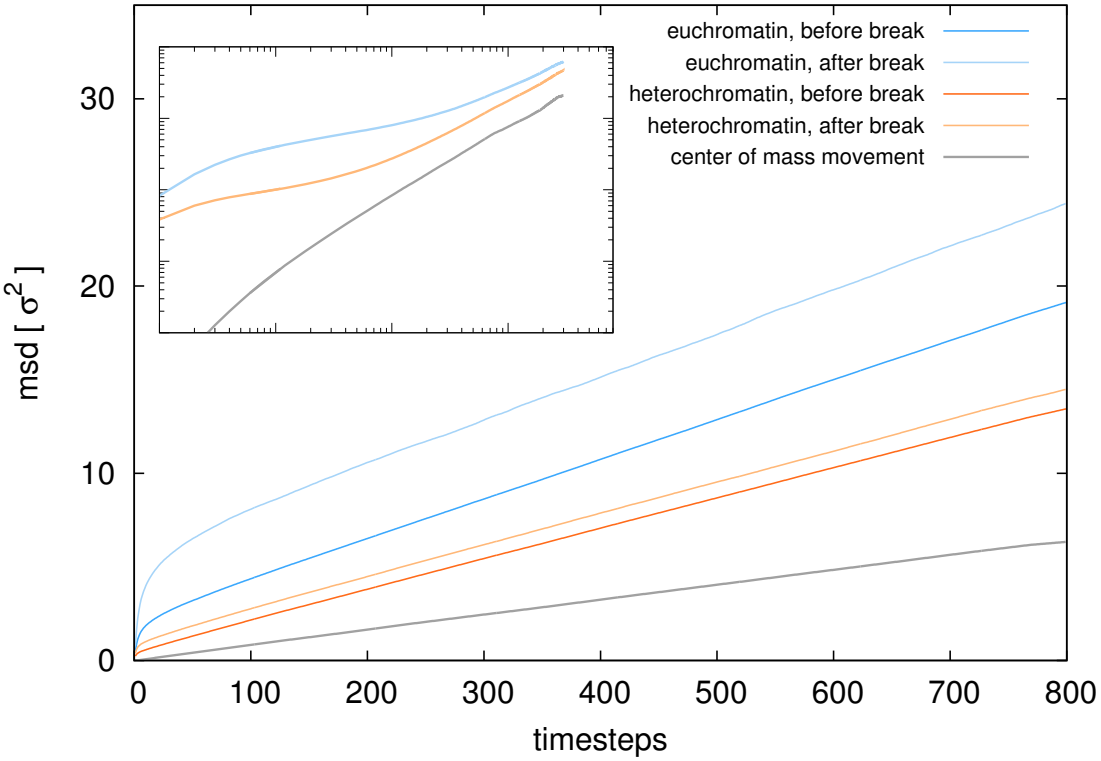

Figure 6.4: Dynamics of different chromosomal regions before and after a break. The mean squared displacements of model chromatin fibre segments before and after a break was induced in human chromosome 11. The chromosome was modelled with the expression-dependent Dynamic Loop Model. A break was induced between the two monomers which were chosen before according to their mean number of loops. First of all, we observed that the mobility of chromatin in the loosely organized euchromatin is higher than the mobility of heterochromatin. However, the mobility increases when the break is induced in both regions. *inset:* The mean squared displacements of the chromosomal segments prior to the strand break and additionally the mean squared displacement of the centre of mass of the chromosome. Different regimes can be seen for the movement of both chromosome segments. From around 800 time steps onward, the movement of the segments is governed solely by the centre of mass movement.

accumulate rapidly at the site of DSBs [218]. On the other hand, large-scale rearrangements of the genome are more likely happening on the scale of hours to days and might even have to involve progression of the cell cycle [146]. DSBs in mammalian cells were also reported to possess positional stability in the nucleus [134]. Furthermore, eukaryote cells possess the ability to arrest the cell cycle and stall it in response to DSBs in order to prevent, e.g. erroneous replication [230,231]. This could also be an indication that no further large-scale rearrangements of the chromosomal loop structure occur during the DSB repair process. Therefore, we assume a time scale separation between the occurrence of a global reorganization of the loop structure and the times that are relevant to the repair of DSBs. In the time scale that is interesting for repair, we expect only small changes in the loop structure that have only little influence on the behaviour of DSBs. We verify this assumption by conducting additional simulations of breaks in chromosomes that are able to reorganize their loop structure. Our results here are similar to the results for strand breaks in fixed loop structures. The simulation results for these verification simulations

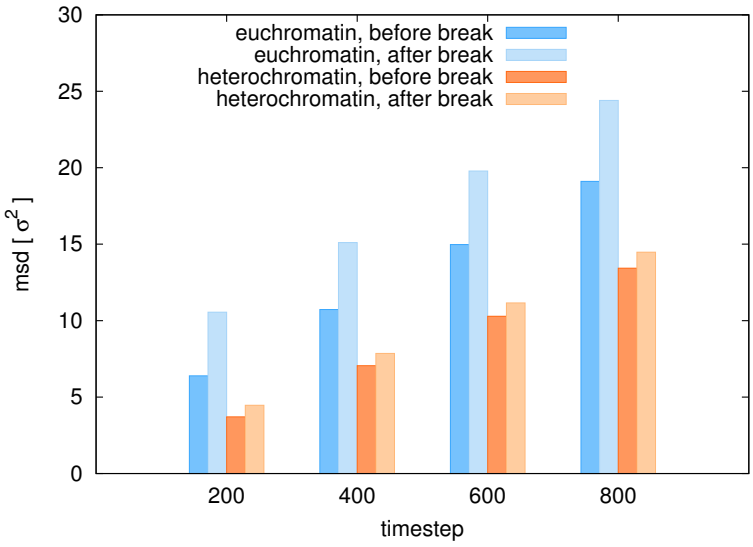

Figure 6.5: Mobility of different domains. To illustrate the mobility increase of the break ends in both kinds of environments, the mean squared displacements at time steps 400 and 800 are shown here. The increase of mean squared displacement for euchromatin is 41 % at 400 time steps and 21 % at 800 time steps. For heterochromatin, the increase is 11 % at 400 and 8 % at 800 time steps. The stronger increase in mobility for the euchromatic region can be explained by the overall lower density of the environment which lets the broken ends take full advantage of their gained entropic freedom.

can be found in the *Supplementary Material*.

Figure 6.4 and 6.5 show a comparison of the mean squared displacements of the segments in the two different regions and before and after the double-strand break was induced. Panel C shows the mean squared displacement with time of chromosomal segments in different areas on a log–log scale and additionally the mean squared displacement of the centre of the mass of the model chromosome. We observe different domains of movement. Clearly, from around 800 time steps, the movement of the sub-domains is governed only by the movement of the whole model chromosome. Therefore, the interesting domain of movement is the region before 800 time steps. Although the mobility of segments in different regions is dependent on the average transcriptional activity, clearly, in all three regions, the mobility of the broken ends increases. The mobility increase is highest in the transcriptionally very active euchromatic regions where the chromatin is more loosely organized. In the compact and silent heterochromatic regions, the increased mobility of broken strand ends is less pronounced. The increase of the mean squared displacement for the euchromatic region is 41 % after 400 time steps and 21 % after 800 which is much higher compared to 11 % after 400 time steps and 8 % after 800 time steps for the heterochromatic region.

We have thus shown that the increase of mobility of the sites of DSB can be explained by the degree of freedom that a broken strand end has compared to the unbroken site. This removal of entropic constraints is sufficient to cause faster movement and a larger roaming area of the break site consistent with experimental findings [208]. Furthermore, increased movement can be found in all kinds of environments, more compact and dense heterochromatin as well as loosely organized euchromatin. Here, however, we can see that the degree of change is different in different pre-break regions. The breaks that occur

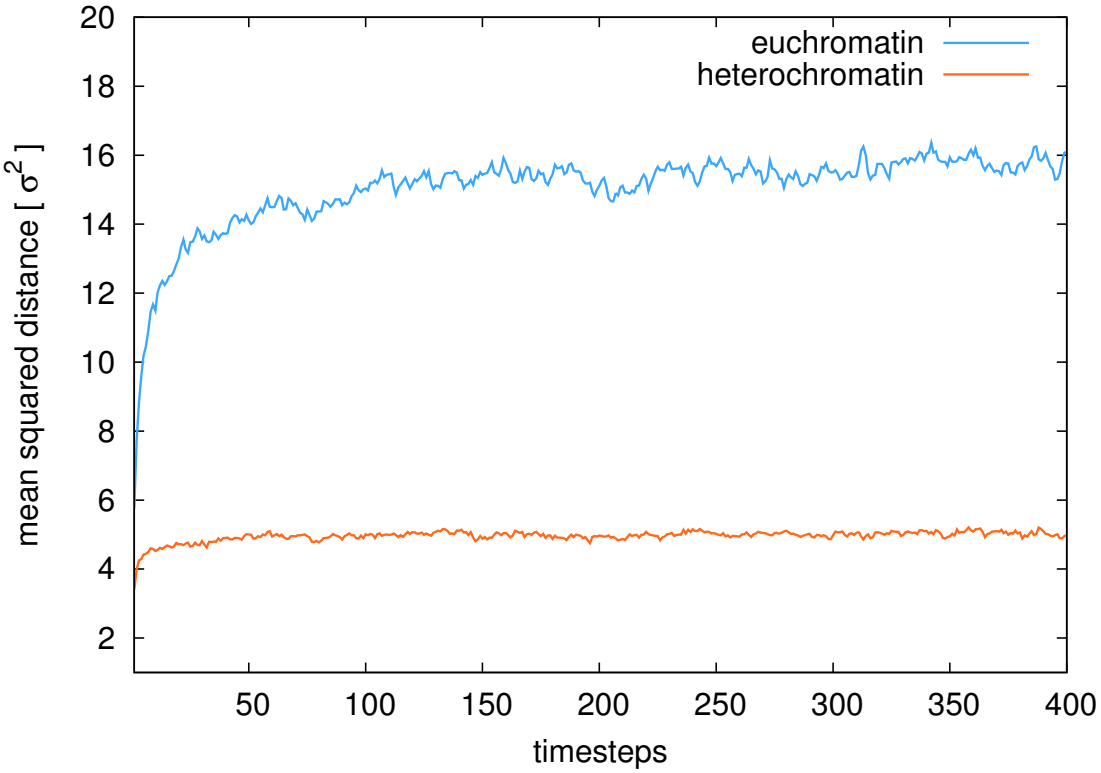

Figure 6.6: Separation of broken ends. Our results confirm that loops can also play a very important role in the context of double-strand breaks. Here, we see that the separation of broken ends is restricted due to the tethering of the model fibre. For comparison, the mean radius of gyration has a value of $\langle R_g^2 \rangle = 38.4$, and the mean squared separation of connected adjacent segments is 0.95. Most importantly, in heterochromatin, the broken ends are held very close to each other by the compact environment and the high number of tethering positions. In euchromatin, the break ends can drift away from each other further, but still well below the mean radius of gyration. This shows again how differently breaks are treated in different chromosomal domains. The fact that break end separation is very limited in very compact regions could cause them to be less well recognized by the cell and, thus, explains why repair foci form less often in heterochromatin than in euchromatin.

in compact heterochromatin do not change the mobility to the same degree as breaks in more loose euchromatin. Therefore, the removal of the entropic constraint by cutting the model fibre has a larger effect in the less dense region than that in the very compact region. This could already be an indication that the entropic changes produced by the breakage of a strand are less pronounced in dense region which in turn could mean that the detection rate of the DSB in heterochromatin is lower. Indeed, experimental findings show that foci of phosphorylated histone variants *H2AX* are found significantly less often in heterochromatin than in euchromatin, indicating different treatment of DSB in these different environments [140, 141]. We suggest that the differential entropic changes may play a significant role in this differential treatment.

For the repair of DNA double-strand breaks, it is important to study how far broken ends can drift away from each other. The success of repairing the strands by rejoining the correct ends critically depends upon the ends to stay in close proximity of each other. Therefore, we analyze the average distance between break ends in our model chromosome

with induced breaks. The result can be seen in Figure 6.6. We observe that the mean squared distance between the broken ends increases for small times but quickly reaches a plateau region where the average distance does not increase any more. The level of this plateau is well below the radius of gyration. We can, thus, conclude that the presence of loops helps prevent break ends from drifting away from each other and, thus, makes it easier for repair proteins to guide the ends to each other and to successfully repair the break. In genomic regions with many loops, this could be of great importance for correct repair by non-homologous end joining which depends on direct mending of the broken ends. Such a behaviour was suggested for DSB repair in D. radiodurans, where the non-homologous end joining pathway could be facilitated by a compact structure that prevents broken ends from drifting too far away from each other [223].

To further highlight the importance of loops within the chromatin fibre in the context of double-strand breaks, we also conduct simulations of a model chromatin fibre without any loops. As in the other simulations, we induce a break and observe the further dynamic behaviour of the model chromosome. Here, the two break fragments rapidly drift away from each other and do not stay together. Such a behaviour would certainly jeopardize any possibility of repair. We show that the loops, on the other hand, can anchor the break fragments to each other and furthermore help keep the broken ends in close proximity to each other. Therefore, they contribute vitally in a key difficulty of DSB repair.

6.3.4 Entropy Helps Break End Relocation to the Periphery of Chromosomal Domains

Recent experimental findings state that DSB in heterochromatin are transported to the surface of the domain where repair can take place more easily than in the tightly compacted region [142, 232]. Therefore, we also analyze the positioning of artificial strand breaks in our model chromosomes. Here, we also induced breaks at different areas in the ensemble of conformations and continued simulations with fixed loop structure. In Figure 6.7, two example conformations are shown, one before and the other after a break was induced. The domain surrounding the break is coloured in red and blue, and the rest of the chromosome is coloured in grey. The lower panels show a detailed view of the break domain, with red being upstream of the break, and green downstream of it. Prior to the break and immediately after, the DSB region is located more to the interior of the domain. Entropy alone is sufficient to quickly drive the broken ends to the periphery of the domain.

Figure 6.8 shows the mean squared distance of each statistical segment of the model chromosome to its centre of mass before and after the break was induced. The distance to the centre of mass depends strongly on the number of loops that the statistical segment has on average. However, independent of this, DSBs clearly have an increased distance to the centre of mass than the unbroken segment, which means that they move further to the outside of the particular domains. This movement is stronger in transcriptionally active regions which are less dense than inactive regions. Therefore, a relaxation of the structures in dense heterochromatic domains could help exert this transport even more, as has been proposed before [142]. Indeed, a $3C$ experiment indicated a decrease in chromosomal contacts in the surroundings of site-specific DSBs [233] which, in turn, could be associated with an opening of the organization of the domain.

We have, thus, shown that the entropic freedom for the DSB region not only enhances the mobility of the region, which could become important, e.g. for homologous repair, but also helps relocate breaks to the periphery of the domains where accessibility to repair proteins is much better. The fact that entropy, as a globally effective motif, plays a major

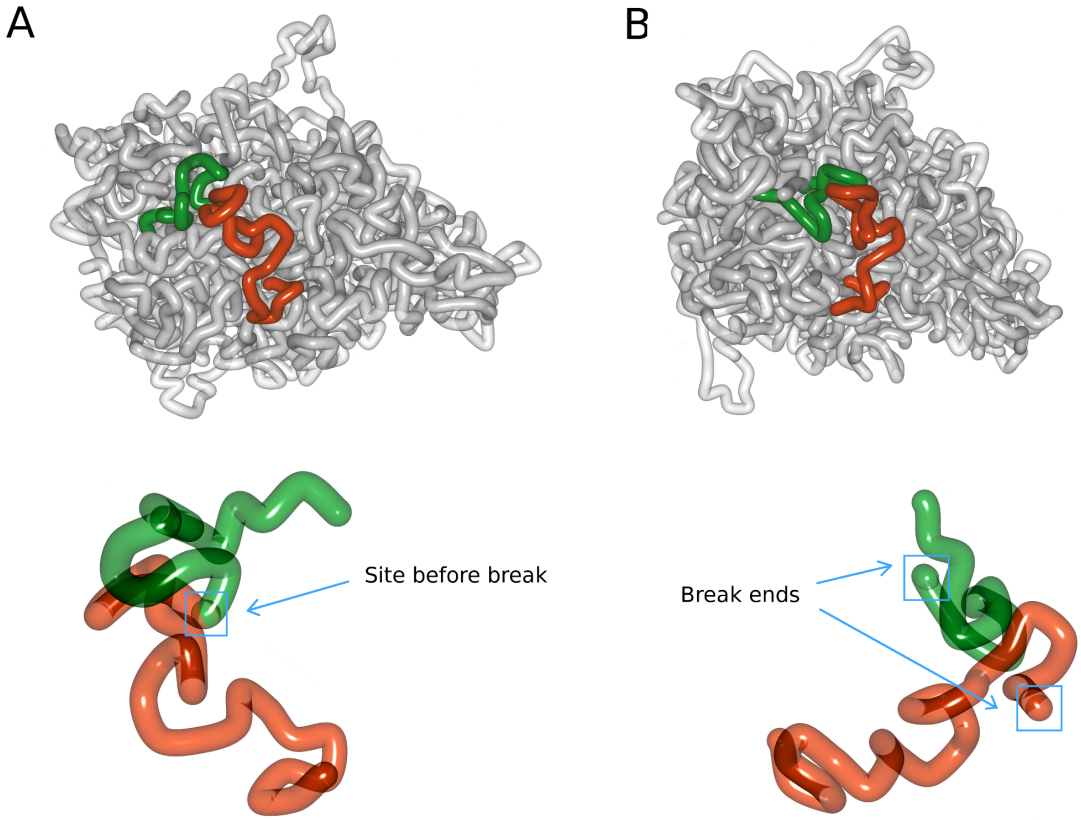

Figure 6.7: Chromosome sub-domain before and after a break was induced. The grey tube represents the chromatin fibre. The part of the fibre that belongs to the sub-domain is coloured in red and green. The fibre section upstream to the break site is coloured in green while the section downstream is coloured in red. The break site before and after the break is marked with blue boxes. Note that the sub-domain is rotated in the close-up view. **a** Before the break, the site is located closer to the centre of mass of the chromosome sub-domain. Additionally, we observe that the region upstream and downstream of the break site intermingle only little with each other. **b** The conformation at 2,000 time steps after the break was induced. Firstly, the chromosome loops prevent the sub-domains from drifting apart from each other and limit the separation between the break ends. Most importantly, the break ends are now located more towards the periphery of the sub-domain.

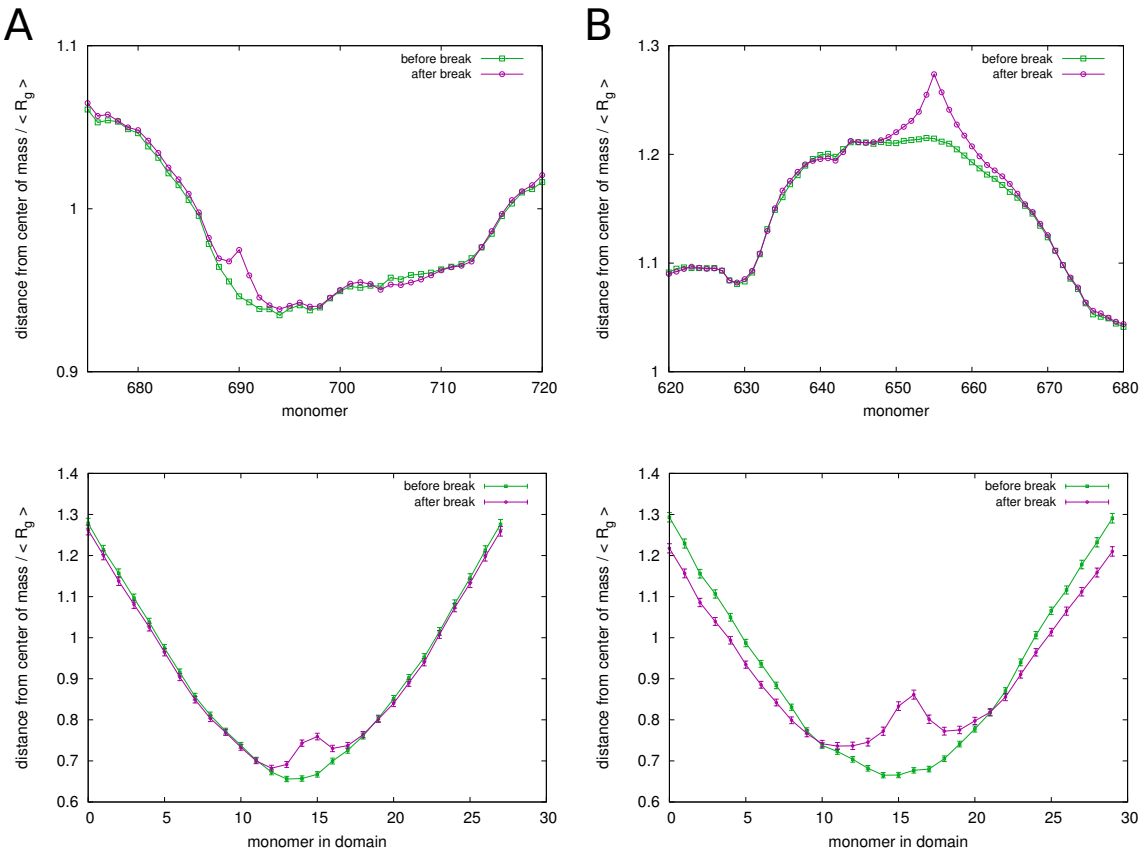

Figure 6.8: Distance of breaks in model chromosomes. The figure shows the distance of parts of the model chromosomes to the centre of mass for human chromosome 11. The green curves are for the average distances of the untreated chromosomes, the purple curves are for chromosomes with a model break after 100 time steps. **a** Distances for a rather compact domain with large number of loops and low overall gene expression activity. In the panel above, the distances of the statistical segments to the centre of mass of the chromosome territory are shown. The panel below shows the distance of the segments to the centre of mass of a domain of approximately $3\,Mb$ around the break. **b** Distances for a rather loose domain with small number of loops and high overall gene expression activity. In both cases, it was observed that a break leads the broken ends to a motion that is directed outwards of their domain.

role in the process of DSB recognition and repair shows that the repair mechanism within the cell nucleus is, in a sense, stable. The formation of a break naturally triggers a reaction of the systemâĂŤbased on its entropic behaviour - that tries to counteract the break by moving the loose ends outside of tight domains where they are more accessible to repair proteins. Seen in the context that the repair machinery reacts to all kinds of DSBs, those induced by exogenous agents such as radiation but also programmed, transient DSBs, it becomes even more evident that passive processes governed by entropy should play a key role in the treatment of DSBs.

6.4 Conclusion

In this work, we employed an expression-dependent folding model to investigate how chromosome structure changes in cells that were exposed to ionizing radiation. As model

input, we used the gene expression data of human immortalized B cells from Smirnov et al. [51]. We calculated the mean chromatin densities and loci locations in the chromosome territories for chromosome 11 in untreated cells and cells at $2\,h$ after exposure to $10\,Gy$ ionizing radiation. Our results show that no significant structural alterations on the scale of the complete chromosome can be found despite the expression level change of many single genes. This lack of overall organizational difference can be explained by the averaging out of the effects caused by a single gene among a large number of other genes. Therefore, we conclude that structural changes are rather limited to local scales.

Using the expression-dependent folding model, we identified different environments within the chromosome territories. Transcriptionally active regions were located further to the surface of the territory and had a much smaller density than inactive regions. To study the behaviour of the strand breaks in these different regions, we induced breaks in these two different kinds of environments while fixing the loop structure of the whole chromosome. We compared the mean squared displacements of the break sites before and after the break was induced. In both kinds of environments, an increase of the mobility of the break site was observed, which is in agreement with experimental findings of Krawczyk et al. [136]. We further showed that the gained entropic freedom by the break ends leads to a stronger mobility increase in the transcriptionally active region. Additionally, our results indicate that broken ends move further to the periphery of their surrounding domains without any external forces. This effect is stronger in less dense domains. Therefore, it could be possible that an opening of the tight organization in the heterochromatic regions can already be sufficient to exert the entropic forces that help transport DSBs to the periphery of the domains. We, thus, conclude that entropy plays a major role for the behaviour of DSBs in the cell nucleus, and we suggest that it forms a basis for cell response to DSBs.

6.5 Supplementary Information

Simulation of Broken Ends in Dynamically Restructuring Chromosomes

For the simulation of explicit strand breaks in chromosomes we first sampled equilibrium conformations using the expression-dependent Dynamic Loop Model. In each of the sampled conformations we then first fixed the loop structure, which means that intra-chromosomal cross-links were no longer able to form or dissolve. We then cut the model chromatin fiber at one specific position and monitored the behaviour of the broken ends by means of Molecular Dynamics simulations. The fixation of the loop structure was motivated by the assumption that the time for a reorganisation of the loop structure is much larger than the time that is interesting for DSB repair.

We further also performed simulations of strand breaks in chromosomes for which the loop structure is not fixed but rather subject to dynamic formation and dissociation of loops. For this we used a very similar procedure as before. We used the same equilibrium conformations and cut the model chromatin fiber at the same positions as in the prior simulations. We then proceeded the Monte Carlo simulations with the dynamical and expression-dependent looping mechanism, i.e. we sampled equilibrium conformations for the âĂIJcut fiberâĂİ. We analysed the behaviour of strand breaks by calculating mean properties for the equilibrium conformations. Our results show that we obtain qualitatively the same results as in the case of chromosomes

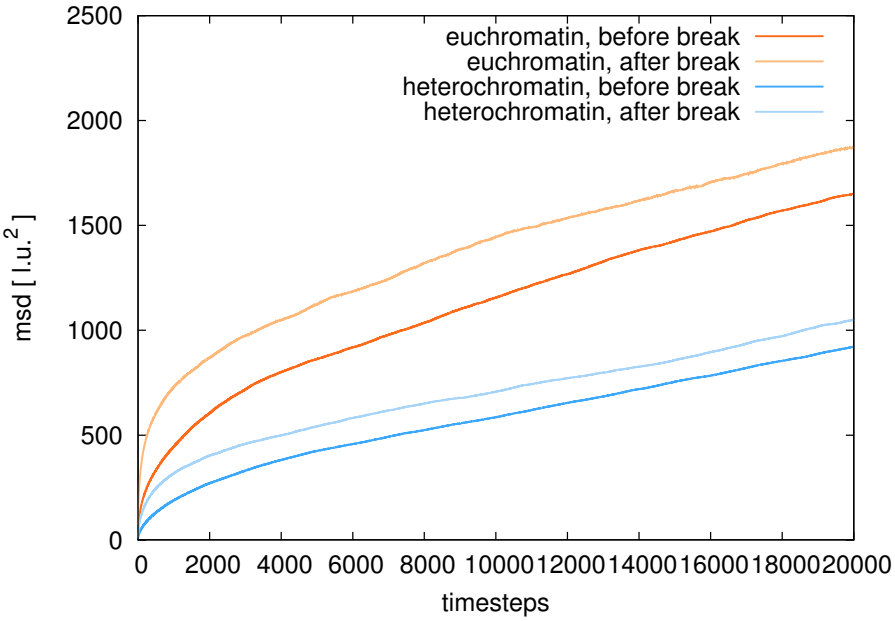

Figure 6.9: Mean squared displacement of chromosome segments before and after break. The red curve shows the mean squared displacement of a segment in active euchromatin while the blue curve shows a segment in an inactive euchromatic region. In both regions the broken ends have an increased mobility due to the gained entropic freedom. The results are qualitatively equal to the results for strand breaks in chromosomes with fixed loop structure.

Dynamics of Broken Ends

In Fig. 6.9 we can show the mean squared displacement of model fiber segments before and after the induced break in two different areas of the chromosome. Our results show that the increased entropic freedom of broken ends result in a higher mobility. This is in agreement with our findings for the simulations of strand breaks in model chromosomes that have a fixed loop structure.

Positioning of Broken Ends in Chromosome Segments

To analyse the positioning of broken ends within their domains we calculated the mean distance of all statistical segments of the model fiber to the center of mass of the chromosome. We then compare the distances at the break position before and after the induced break. Fig. 6.10 and Fig. 6.11 show the results for breaks in heterochromatin and euchromatin respectively. In model chromosomes which can dynamically reorganize their loop structure we also observe that broken ends locate further to the periphery of the chromosome territory and also of their subdomain.

We further compare the time scales on which the broken ends move outside of their domains to the mean lifetime of loops. In Fig. 6.12 we show the mean distance of the broken ends to the center of mass of their domain agains the timesteps after the break was induced. As the simulation shows broken ends quickly move to the periphery of the domains within only a small fraction of the mean lifetime of loops in the chromosome.

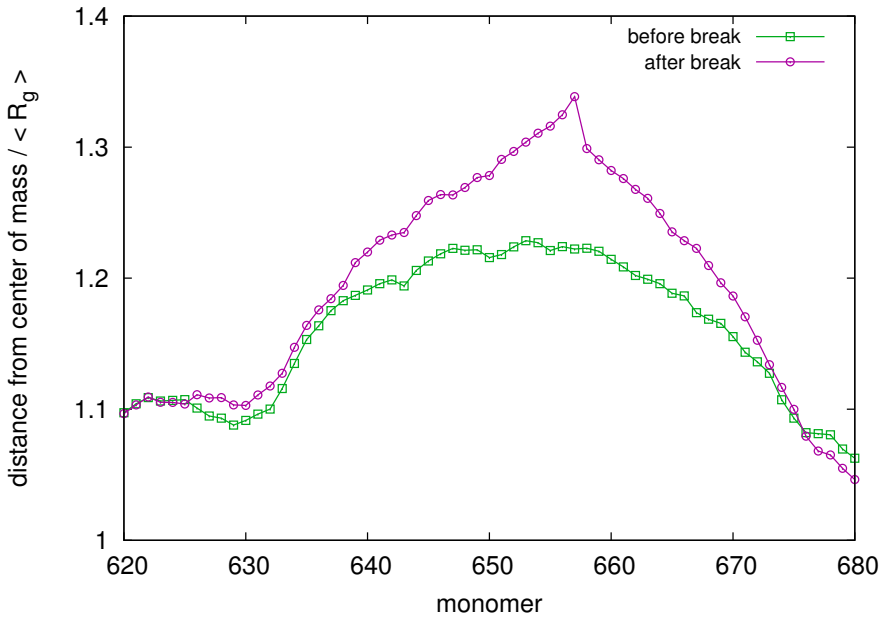

Figure 6.10: Distance of fiber segments from the center of mass of model chromosomes in a heterochromatic area. After the break, break ends and adjacent segments move further to the outside of the domain.

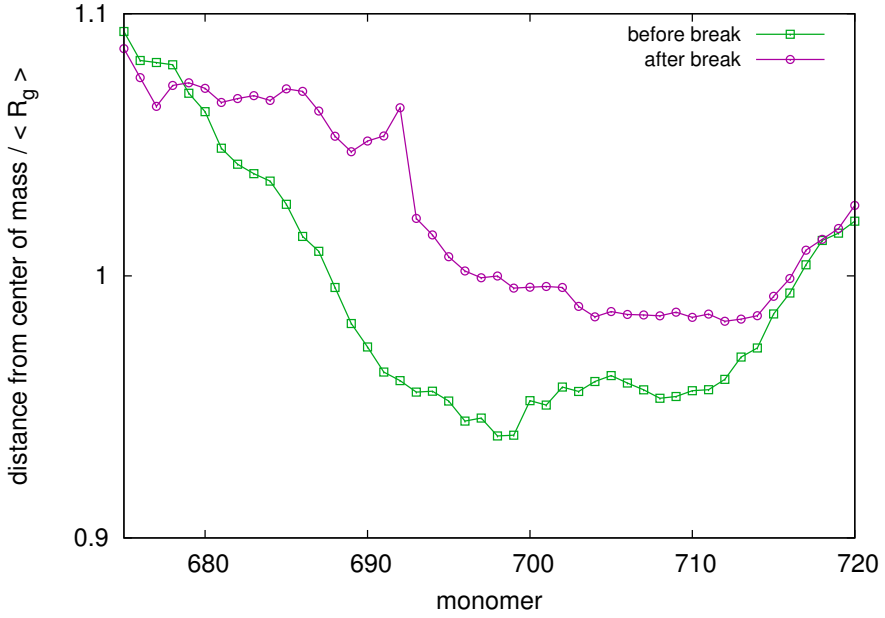

Figure 6.11: Distance of fiber segments from the center of mass of model chromosomes in a heterochromatic area. In heterochromatin break ends are also repelled from the center of the domain and move to the outside due to their entropic freedom.

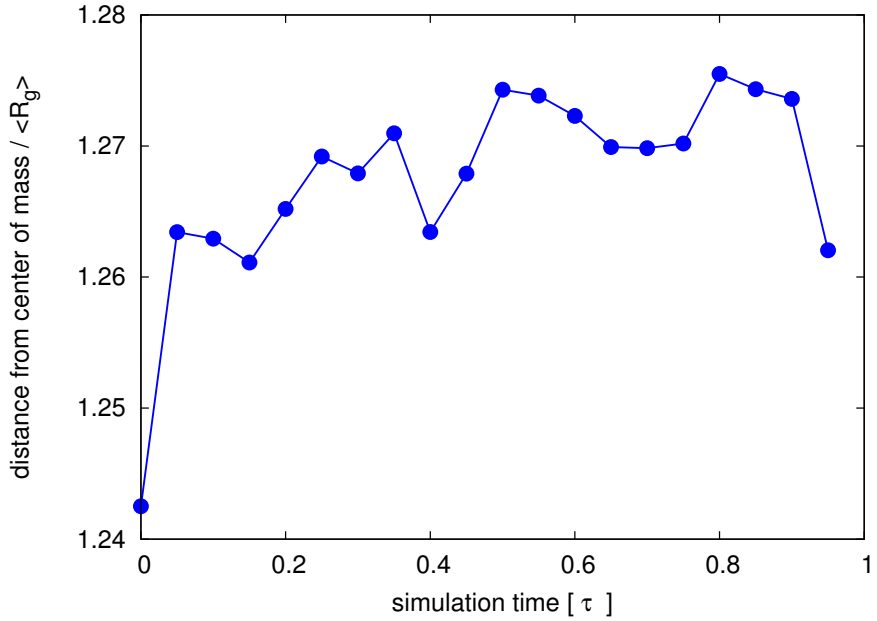

Figure 6.12: Distance of break end to center of mass against simulation time. The simulation time is in units of the mean loop lifetime τ and the distances are in units of the gyration radius R_g. Broken ends rapidly move further to the periphery of their domains on a time scale that is much smaller than the average life time.

Chapter 7

Entropy Facilitates Double-Strand Break Recognition

References

The content of this chapter was adapted from the following publication

- Y. Zhang, C. Schindler and D.W. Heermann, *A Role for Entropy in DNA Double-Strand Break Recognition and Repair*. in preparation.

Chapter Summary

In biology, the term entropy is often understood as a measure of disorder - a restrictive interpretation that can even be misleading. Recently it has become clearer and clearer that entropy, contrary to conventional wisdom, can help to order and guide biological processes in living cells. DNA double-strand breaks (DSBs) are among the most dangerous lesions and efficient damage detection and repair is essential for organism viability. However, what remains unknown is the precise mechanism of targeting the site of damage within billions of intact nucleotides and a crowded nuclear environment, a process which is often referred to as recruitment or signaling. Here we show that the change in entropy associated with inflicting a DSB facilitates the recruitment of damage sensor proteins. By means of computational modeling we found that higher mobility and local chromatin structure accelerate protein association at DSB ends. We compared the effect of different chromatin architectures on protein dynamics and concentrations in the vicinity of DSBs, and related these results to experiments on repair in heterochromatin. Our results demonstrate how entropy contributes to a more efficient damage detection. In conclusion, we identify entropy as the physical basis for DNA double-strand break signaling.

7.1 Introduction

Double-strand breaks (*DSB*s) arise when two single-strand breaks form in close proximity or due to failures in the replication machinery. Physiological processes and ionizing radiation are a major source of *DSB*s [234], which can be generated anywhere within the genome. Although DSBs are quite rare compared with other forms of *DNA* damage, they are difficult to repair and extremely toxic [25]. It is therefore of utmost importance for cells to have intact repair mechanisms and damage sensors to rapidly detect and mend DNA double-strand breaks. If DSBs are not correctly repaired they can lead to cell death or to genome rearrangements that can eventually even result in cancer [27].

DSBs induce a wide-ranged damage response that beyond detection and repair may also lead to cell-cycle arrest and apoptosis [235]. The first stages of the repair involves the recognition of sites of DSBs. External factors such as ionizing radiation can produce DSBs at any position on the genome. Especially radiation with low linear energy transfer (LET), such as gamma radiation usually have a homogeneous dose deposition profile throughout the cell nucleus [236]. Thus, the cell must be able to recognize DSBs at all genomic regions. It is very likely that such a recognition process is dependent on entropic and structural changes that a break of the DNA strand and the chromatin fiber causes in its surroundings. In this work we study what changes of the environment of the break site are induced by the breakage of the fiber and how this relates to the early recognition of DSBs. Our results indicate that entropy can play an important role for the recognition of DSBs.

Many works target the different stages of the repair of DSBs. After the damage has been recognized by the cell, early repair proteins accumulate in the surroundings and attach to the site of DSBs. Among the proteins that arrive early at the break sites in mammalian cells is for example the MRN complex [122, 123]. The binding of the MRN complex can then trigger a cascade of subsequent repair processes involving recruitment of other proteins and chromatin modifications [125,218,219]. Especially the phosphorylation of the histone variant H2AX at serine 139 to γ-H2AX has become an indicator of DSBs. γ-H2AX can spread to up to $1\,Mb$ of the surrounding chromatin and can be made visible using fluorescence microscopy [49, 128, 129]. High levels of γ-H2AX can also be found at Telomere regions indicating the similarity of broken ends and Telomeres [139]. A close relationship between chromatin structure and DSB repair has been noted: from the nucleosome level [237, 238] to local chromatin (de-)condensation [133, 239] and differences between heterochromatin and euchromatin [142].

Eukaryote cells have generally two repair pathways: homologous recombination (HR) and non-homologous end-joining (NHEJ). Non-homologous end-joining is usually the repair pathway of choice in case when no homologous template is available and broken ends are mended directly with each other [119]. The Ku complex was reported to be involved in the early stages of NHEJ and is believed to form a ring that can thread onto broken ends of the strand [131]. In contrast, the HR repair pathway uses the sister chromatid as a template for the repair. Different studies have observed increased mobility of chromatin domains containing DSBs which is supposedly necessary for the search of a homologue template [220]. The duality of both repair pathways is also exemplified in the controversy about DSB repair in the highly radioresistant bacterium *Deinococcus Radiodurans* [223, 224].

Before cells can follow their repair programs, DSBs have to be recognized as such. DSBs which are for example created by ionizing radiation can in principle occur anywhere along the genome and there is, to our knowledge, no evidence for pre-designated break

points. Therefore, the recognition of DSBs and the recruitment of early repair proteins to the site of DSBs have to be based on physical principles rather than sequence search processes.

In the context of target sequence search for proteins on DNA, the model known as *facilitated diffusion*, featuring sliding of proteins on the DNA fiber combined with conventional 3-D diffusion has been proposed and extensively studied [240–242]. Recent works have shown that target search by facilitated diffusion is strongly influenced by DNA / chromatin conformation [243]. Another important factor for these kind of search processes is molecular crowding inside of the cell nucleus [244, 245]. Crowding by other proteins in the nucleus but also by the chromatin fiber itself can efficiently shield the target site from searching proteins. Interestingly, depending on their strength, excluded volume interactions can also have a positive effect on the search time [246]. Hansen et al [247] examined the influence of polymer fluctuations on binding using a mean-field ansatz to show enhanced association rates for ideal chains. Optimal 1D target search along the polymer chain considering the folded polymer structure is investigated e.g. by Lomholt et al [248].

However, target search in the context of recruitment of damage sensor proteins to DSBs in the early stages of the damage response has, to our knowledge, not been the focus of the attention so far. In this work we analyze the behaviour of broken ends of the chromatin fiber and their recognition by repair proteins. We hypothesize that the change in entropy associated with inflicting a DSB influences the target search of damage sensors and the local chromatin structure. We compare the effect of different chromatin architectures on the search process for DSBs. We first model the chromatin fiber with a simple polymer model representing the immediate surrounding of the DSB. The ends of the chain can then be interpreted as one part of the DSB.

In recent years, it has become more and more evident, that looping of the chromatin fiber is a general motif for the organization of the genome. The *chromosome conformation capture* technology (3C) and its variants (4C, 5C, HiC) have shown the presence of loops on all scales in eukaryote cells [178–180]. Different modeling attempts have shown that loops can create entropic conditions that facilitate the compartmentalization of the chromosome into sub-domains [31, 184]. Loops are believed to be created by binding proteins. Both, models that include the effect of binding proteins implicitely and explicitely are able to describe the folding of the chromatin fiber in agreement with experimental observations [183, 184, 205].

These studies have shown that the number of loops and binding proteins respectively have a profound influence on the local organization of the chromatin fiber. In areas with high numbers of protein-mediated contacts of the fiber, the chromatin density is increased, while low contact numbers result in low chromatin density regions [183]. The former regions could be identified with silent heterochromatin while the latter areas with actively transcriped euchromatin. Different densities in the environment of DSBs could also play an important role for the recognition dynamics of the DSB. Just as molecular crowding effects by surrounding proteins, the chromatin fiber itself could provide efficient shielding of the target site [144].

To model the behaviour of DSBs in different environments we then employ a more realistic model for chromosome subdomains which features dynamic looping of the fiber. Strand breaks within the chromosome sub-domain are modeled by disconnected chains in this model. Additionally, we represent repair proteins by diffusing particles. We use computer simulations to calculate entropic and dynamic properties for the ensemble of possible fiber conformations and protein configurations. The search time denotes the

simulation time until a protein binds to a specified target site on the chromatin fiber. We use the average search time as a measure for the detectability of a chromatin segments by surrounding diffusing proteins.

Our results show that the entropic freedom that the broken fiber segments gain in turn leads to dynamical and structural conditions that increase their detectability for diffusing repair proteins. Furthermore, we demonstrate that broken ends within chromosome subdomains are relocated to the periphery of the domain driven by the entropic changes. We suggest that early DSB recognition can be in a sense viewed as a stable process where the disturbance of the system - the DSB - leads to conditions that helps with the removal of the disturbance.

7.2 Methods

7.2.1 Models for Chromatin Fiber and Proteins

We use a coarse grained approach to model the chromatin fiber. The fiber consists of a chain with N monomers each representing a coiled unit of DNA / chromatin. The detailed folding within each monomer is insignificant for the large scale folding of the coarse grained fiber and can thus be neglected [41]. To prevent the different monomers from overlapping with each other we include a steric repulsion between them leading to excluded volume interactions.

In this work we explore the different properties and search times for DSBs with different polymer models. First we use a simple polymer model with no additional interactions between the monomers other than excluded volume. In this simple model the double-strand break is considered to behave essentially like a free polymer end, which is also supported by the similarity between Telomeres and DSBs. Since the displacement of the entire chromosome is likely occuring on a much larger time scale, we neglect the center of mass motion of the chromomsome. Therefore, we employ a polymer model, where one end of the DNA is fixed to represent the not included large part of chromatin bulk and to remove simple diffusion of the polymer as a whole.

Our model chromatin fiber is immersed in a diluted solution of model proteins or ligands. Each protein is represented by a simple, freely-diffusing monomer. They can form bonds with DNA monomers with a binding probability p_b, which is specified for each monomer of the chain individually. The binding probability can be set to define a unique binding site (target site) on the DNA. This is used to compare different target site positions along the model chromatin fiber.

For the more realistic model of a chromosome subdomain we further add a dynamic looping mechanism for the chromatin fiber. This looping mechanism accounts for the experimentally observed intra-chromosomal contacts [17]. Non-adjacent monomers are able to form an additional bond with probability p when they come into close proximity of each other by diffusion, thus creating loops. A lifetime t drawn from a Poisson distribution with mean value τ is assigned to each newly created loop. After expiration of the lifetime, the loop dissolves again. This dynamic looping mechanism represents an effective incorporation of binding proteins in the solution around the chromatin fiber [184,205]. By varying the parameters p and τ we obtain different mean numbers n_{loop} of loops within the fiber. The concentration of loops is denoted by $k_{loop} = \frac{n_{loop}}{N}$ and is a measure for the compactness of the model subdomain. Higher loop concentrations are associated with high compaction due to the large number of intra-fiber contacts. In this work we model

a heterchromatic domain with a value for the mean loop concentration $k_{loop} = 0.83$. The loop concentration for our euchromatin subdomain is $k_{loop} = 0.47$.

7.2.2 Monte Carlo Simulations

For the implementation of our model we choose the well-established lattice model for polymers, the Bond Fluctuation Model (BFM). Excluded volume interactions between monomers are included by forbidding any lattice site to be occupied by two monomers. Adjacent monomers of the polymer chains can be connected to each other by 108 different bonds of varying length [164]. Nearest neighbor relations can be easily determined in lattice models, which is a very useful characteristics for the study of binding processes.

We sample equilibrium conformations for our model chromatin fiber with Metropolis Monte Carlo simulations. In each Monte Carlo (MC) step a local move is suggested for each of the monomers of the chain and accepted if the rules for the bond vectors in the BFM are not violated. This ensures excluded volume interactions between the monomers and preservation of the topological constraints by denying the possibility of bond crossings [164].

After each Monte Carlo move a random monomer is selected and adjacent lattice positions are checked for potential binding partners, which can be any other monomer that is not already bound to the selected monomer. A random monomer is chosen among these potential binding partners and an additional bond between the two monomers is established with the probability p. Thus, detailed balance is ensured in the entire simulation process. If the bond is created, a lifetime t is drawn from a poisson distribution with average value τ and assigned to this bond. After that, loop bonds for which the lifetime has expired are deleted.

Since we only apply local moves in each MC step, subsequent conformations are highly correlated. This correlation decays with simulation time. We calculate the integrated autocorrelation time τ_{int} to determine the simulation time until two conformations are independent from each other. The autocorrelation time also determines the equilibration time from the starting conditions. We first calculate the autocorrelation function

$$C_{XX}(t) = \frac{1}{N\sigma^2} \sum_{k=1}^{N-t} \left(X_k - \langle X \rangle\right) \left(X_{k+t} - \langle X \rangle\right) \tag{7.1}$$

where k denotes the simulation time in MC steps, X_k the observable at time step k, $\langle X \rangle$ the expectation value of X, which can be estimated by the mean value over all conformations $\overline{X}$ and σ^2 the variance of X. We then use the windowing procedure [161] to determine τ_{int}

$$\tau_{int} = \frac{1}{2} \sum_{t=1}^{t_{max}} C_{XX}(t) \tag{7.2}$$

We use $10\tau_{int}$ as equilibration time from the starting conditions and consider two conformations to be independent after $5\tau_{int}$.

7.2.3 Search Time Simulations

In the simulation of search times we first equilibrate the chromatin fiber. We then randomly place a given number of model proteins in the simulation box and again equilibrate the mixture of model chromatin fiber and proteins. We then run the simulation until any

of the proteins attach at the defined target site. The attachment mechanism is very simliar to the looping mechanism of the chromatin fiber. In each Monte Carlo step, all possible neighbouring lattice sites of the target monomer are checked for potential proteins that can bind to the target. One of the potential proteins is then randomly chosen and can bind to the target with a probability p_{bind}. In the case of a succesful binding to the target the simulation is terminated and the total simulation time saved. Typically, a minimum of 40,000 independent simulations are carried out to obtain a good sample for the distribution of the search times for one target site. The search time for a single target site is exponentially distributed with an average value of τ_{search}. We estimate the average value by calculating the mean value of all independent simulations.

7.2.4 Simulation of DSBs in Chromosome Sub-domains with Fixed Loop Structure

To compare the environments around different segments of the model chromatin fiber we additionally perform Molecular Dynamics (MD) simulations of model chromosome sub-domains in solution of model proteins. We chose the MD method here to avoid possible effects by the cubic lattice of the BFM. For the MD simulations we first take equilibrated polymer conformations which we obtained from the MC simulations with our model for the chromatin fiber. Typically we sampled a minimum of 5,000 independent conformations with MC simulations for any parameter setting. Due to the dynamic looping mechanism of our model for the chromatin fiber, each conformation has a number of non-adjacent monomers which are connected to each other by additional bonds. We assume that the time scale on which the recognition and recruitment of repair proteins to the site of a DSB is much smaller than the time scale for the reorganization of the loop structure in the sub-domain. Therefore, we assume that the loop structure does not change significantly during this time. Consequently we fix the loop structure the MD simulations meaning that the existing bonds between non-adjacent monomers are permanent in the MD simulations. On the other hand, we allow no further loop formations throughout the MD simulation.

Proteins are modeled as hard spheres and placed at random positions in the simulation box. We realize excluded volume interactions between monomers with a truncated Lenard Jones potential known as Weeks-Chandler-Anderson potential [228]:

$$U_{ij}^{WC} = \begin{cases} 4\epsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6} + \frac{1}{4}\right], & r_{ij} \leq \sqrt[6]{2}\,\sigma \\ 0, & r_{ij} > \sqrt[6]{2}\,\sigma \end{cases} \tag{7.3}$$

where i and j denote monomer numbers and r_{ij} their spatial separation. In our simulations, σ and ϵ were set to the value 1. Both, bonds between adjacent chain monomers and loop bonds are modeled by FENE potentials

$$U_{ij}^{FENE} = -\frac{1}{2}Kr_{max}^2 \ln\left\{1 - \left(\frac{r}{r_{max}}\right)^2\right\} \tag{7.4}$$

with r_{max} representing the maximal distance between two adjacent monomers. We used the values $K = 30$ and $r_{max} = 1.5$ to avoid bond crossings [229]. We chose the same r_{max} between all model proteins and all chain monomers giving all particles in the simulation the same radial dimensions.

To characterize the environment of different chain segments we evaluate the density distribution of surrounding proteins in the vicinity of each fiber segment. For this we calculate the density of model proteins in a sphere with radius r_{sphere} around each monomer

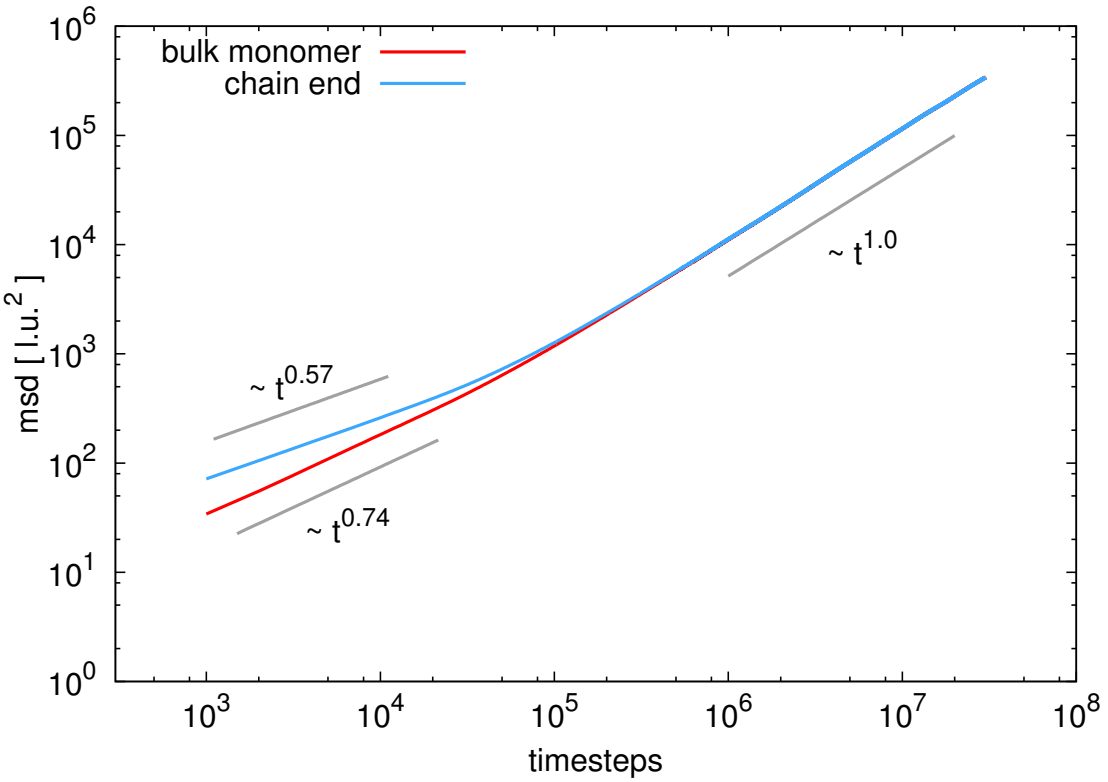

Figure 7.1: Dynamics of different chromosome segments. The figure shows the mean squared displacement of the free chain end and a bulk segment. On small timescales the segments show subdiffusional behaviour $MSD \sim t^{\alpha}$ with scaling exponents $\alpha < 1.0$. We observe that the scaling exponent is smaller for the chain end than for the bulk segment. However, the proportionality constant for the free end is larger and results in total in a higher MSD of the end compared to the bulk segment. On large scales the movement of the fiber segments is governed by the movement of the center of mass and therefore shows diffusional motion with scaling exponent $\alpha = 1$

of the chain. The density distribution follows a Gaussian distribution. To obtain the radial denstiy profile we calculate the density of model proteins in spherical shells with inner radius r and width d. We compare the density distributions around bulk monomers, broken ends, natural ends and far away from the chain. The radial density distribution is small in the direct vicinity of a monomer due to the excluded volume interactions between the chromatin fiber and the proteins. It gradually increases for larger distances until it reaches the value of the average protein concentration far away from the chain.

7.3 Results and Discussion

7.3.1 Entropic Freedom Results in Higher Mobility of Fiber Ends

We first examine the mobility of each segment of the chromatin fiber. We do this by measuring the mean square displacement of the fiber segments with time and average over a large number of independent simulations. Simulations are performed for the simple polymer model as well as the more realistic model for a chromosome sub-domain. We compare mean square displacements for all fiber segments, including natural ends and induced break ends.

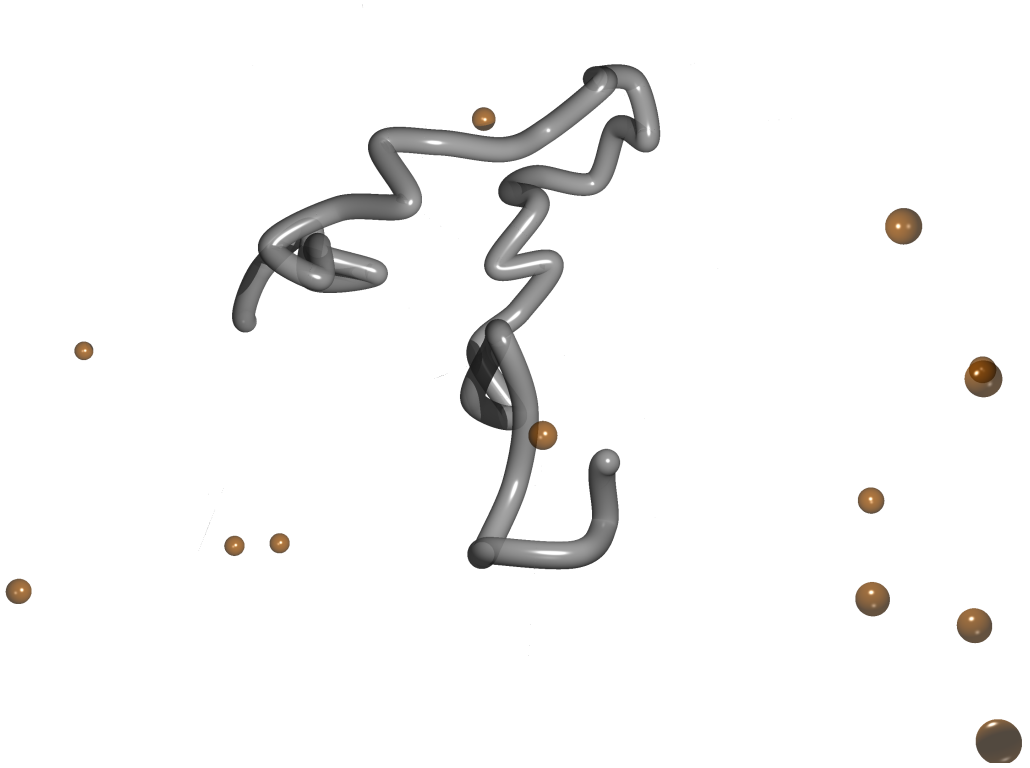

Figure 7.2: The chromatin fiber in the direct vicinity of the DSB is modeled by a self-avoiding walk polymer. To model the close proximity of a DSB we neglect interactions and model the segment of the chromatin fiber with a self-avoiding walk polymer. One end of the polymer is fixed and can be viewed as attached to the bulk of the chromosome. The free end represents one end of the DSB.

The first characterization of the different entropic conditions at the site of DSBs is the change of mobility of the end monomer in our model fiber. To analyse the difference between the mobilities of broken ends of the chromatin fiber and unbroken fiber segments we calculate the mean squared displacement (MSD) of each segment of the model fiber. In contrast to bulk monomers, chain ends are only permanently connected to one neighbour. Therefore, they have a higher entropic freedom which results in a higher mobility. We observe that the chain end has the highestest mean squared displacement at all simulation time steps and the monomer that is at the center of the polymer has the smallest MSD. In Figure 7.1 we show the MSD of monomers on long time scales. For small simulation times, we see the typical subdiffusional scaling of the MSD as a power law t^{α} with $\alpha < 1.0$. However, we observe that different chain segments also have different scaling exponnts. The free end has the smallest exponent of $\alpha = 0.57$ while other fiber segments have larger exponents. On the other hand, the proportionality constant of the MSD is the highest for the free end. This results in the fact that the free end has the highest MSD of all segments which means that its mobility is higher than the mobility of other segments. At large simulation times, the MSD for single chains crosses over to a $alpha = 1.0$ behaviour since the movement of all monomers is then dominated by the movement of the polymer center of mass.

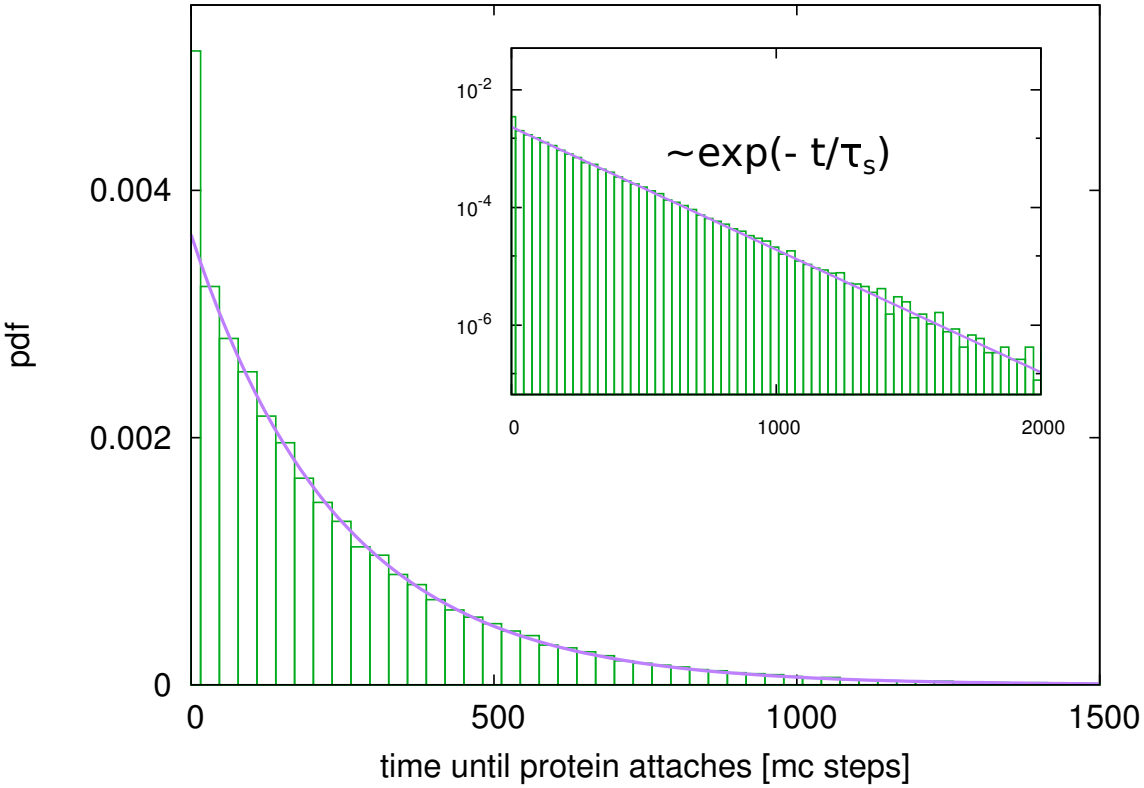

Figure 7.3: Distribution of search times. The figure shows the distribution of search times that proteins need to find a specific monomers on the chromatin chain. The search times are distributed according to an exponential distribution. The characteristical time scale in the distribution is the expectation value which we denote as mean search time $\langle \tau_s \rangle$

7.3.2 Target Search Time of Repair Proteins is Significantly Decreased for Sites of DSBs

The main goal of this work is to examine the binding of repair proteins to DSBs. Since ionizing radiation can produce DSBs at any site along the whole genome, the sensing of DSBs is most likely based on entropic changes that arise due to the emergence of a new end rather than a sequence search process. Therefore we analyze the times required for proteins to find sites of DSBs based on diffusional motion. Proteins are represented as single monomers and placed randomly in the simulation box. Model chromatin fiber and model proteins move according to stochastic dynamics and we measure the time until the first protein binds to the fiber end. We compare these results to simulation results for the first binding time with other segments of the chromatin fiber. In particular, we conduct simulations where one monomer of the polymer is fixed in space, thus having no mobility.

We first perform simulations with self-avoiding walks as a simple polymer model for the chromatin fiber. We fix one end of the fiber in order to model its attachment to the bulk of the chromosome. The free end of the monomer chain represents one end of a break in the chromatin fiber. The proteins are intially placed at random positions in the simulation box at a given concentration c. In the first setup we assume no excluded volume interactions between the chromatin fiber and the proteins and therefore proteins can also move through the fiber. This approximates the case that the size of the proteins are considerably smaller or larger than the chromatin fiber and therefore do not feel the

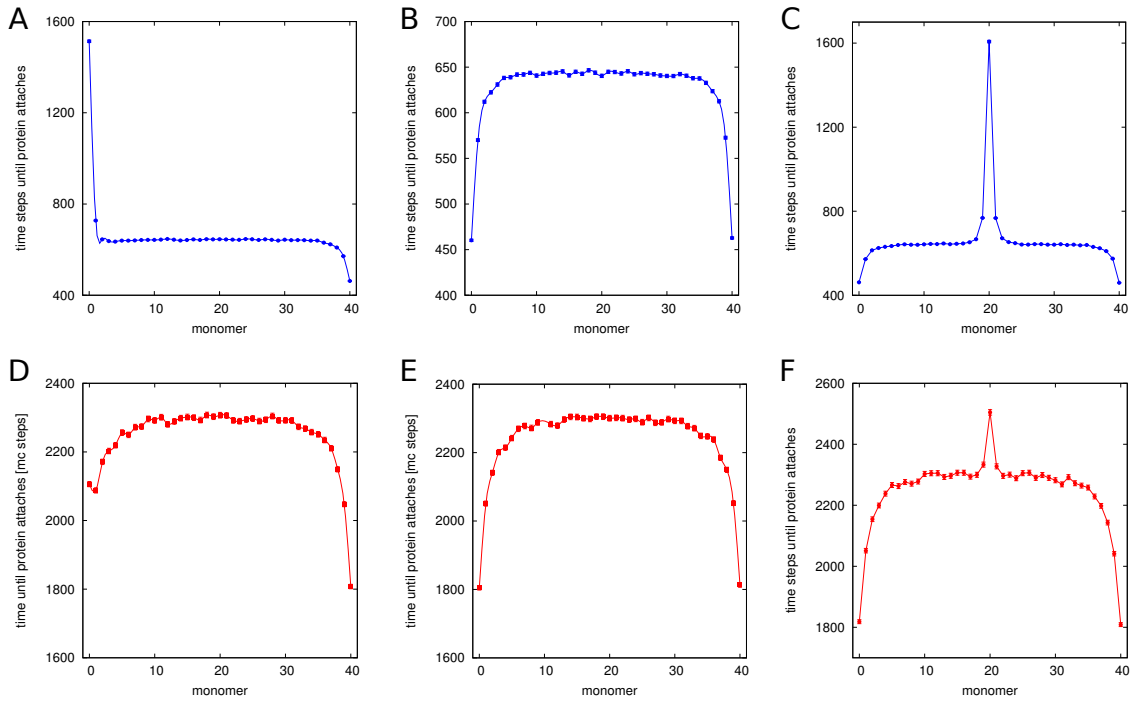

Figure 7.4: Profiles for the mean search time for different linear polymers with $N = 41$ monomers. A. The first monomer is fixed in space while the rest of the chain can diffuse freely in space. The profile of the mean search time shows that the central monomers of the fiber can be found faster by diffusing proteins than the fixed end that does not move. Target search is even further accelerated for the free end of the polymer. **B.** In this system both ends of the chain are free resulting also in a symmetric search time profile. *C.* For chains with fixed center monomer, mean search times are highest at the center and smallest for the free ends. *D.* We then analyzed how excluded volume interactions between model chromatin fiber and proteins can influence the

chromatin presence. In the second setup we include a steric repulsion between the model chromatin fiber and the model proteins which prevents proteins and chromatin to occupy the same space. In both scenarios the different fiber segments possess a steric repulsion with each other and the proteins also possess the same steric repulsion between each other.

For each simulation run we define a chromatin segment as the target segment for the proteins. Proteins are then able bind to this target segment if they come into the proximity of it by diffusion. For this we introduce a binding probability p_{bind} for the protein-chromatin binding affinity at the target site on the fiber. We define the search time τ_s as the time in simulation steps until the first protein binds to the target site. For each target site we perform a number of simulations and calculate the average search time for this target. Figure 7.3 shows the distribution of search times for a target site on a chain in the scenario of no excluded volume interactions between proteins and chromatin fiber. The search times follow an exponential distribution and the characteristical time scale is the average search time $\langle \tau_s \rangle$. The average search time then represents a measure for how well and how fast a specific target on the chromatin fiber can be found by repair proteins.

We calculate the average search time for all segments of our model polymer fiber and focus particularly on the comparison between the the chain end which represents one part of a DSB and the other parts of the fiber. In Figure 7.4 we show results for the simple polymer model in both scenarios - with and without steric repulsion between the

chromatin fiber and repair proteins. Figure 7.4A shows average search times for a polymer without steric repulsion between proteins and for which one end is fixed in space while the other is free. The average search time of the fixed end is clearly increased compared to all other segments which is due to its low mobility and thus inability to roam the space. On the other hand, the free end is more quickly found by repair proteins since it has the highest mobility. To compare additionally conduct the same kind of simulations with a completely free chain and a chain for which the segment in the center is fixed in space. The results are shown in Figure 7.4B and Figure 7.4C respectively. The average search times for the different segments show the according characteristics that we also observed for our model of a chromatin segment that is in the vicinity of a DSB. In the case of no excluded volume interactions between chromatin fiber and proteins the average search time is entirely dominated by the mobility of the target sites. The search for a broken end is facilitated due to the entropic freedom that the end gains.

The situation becomes different when we include a steric repulsion between the repair proteins and the chromatin fiber. Beside the mobility of the target site, the accessibility of the target site to the repair proteins becomes an important factor. We can clearly see this for the results of the average search times. In Figure 7.4D we show that the average search time of the fixed end is in fact decreased compared to the average search times for other bulk segments of the fiber. The decrease of the search time for the free end compared to the bulk segments is also more pronounced than in the case of no excluded volume interactions. This can be explained by the better accessibility of the chain ends which facilitates the search process. Despite smaller mobility of the fixed end, it can be more easily detected and bound to by repair proteins than bulk monomers, because bulk monomers are more strongly shielded from binding due to their more crowded environment caused by the other surrounding fiber segments. The situation between the models with and without excluded volume is qualitatively similar in the case of the completely free fiber and the fiber that is fixed at the center (Figure 7.4E,F). However, in the case of the center-fixed fiber we observe that the average search time difference between the fixed segment and other bulk segments is much smaller in the case of included steric repulsion. This can be seen as another indicator that the accessibility of the target site has a higher impact on the search process than its mobility. In terms of accessibility, all segments including the fixed one are equal to each other. The increased mobility of a segment only plays a small role and thus only changes the average search time by a small amount.

We conclude that the relative mobility between proteins and their target sites is the crucial factor in the search time. Unmoving sites need more time until proteins can attach to them by random diffusion while more mobile polymer ends are attached to after significantly less time. Our results thus show that the formation of a DSBs naturally helps repair proteins to find these sites and attach to them already by the changed entropy in the vicinity of the break.

7.3.3 Search Times in More Realistic Chromosome Subdomains

Our comparison of the search times for simple polymers with and without steric repulsion between the fiber and proteins showed us that the accessibility of the target region is of great importance for the recognition time of the break. Therefore it is very interesting to analyse the search time of DSBs in a more realistic model of chromosome sub-domains. In this model we capture the coiling of chromatin by a dynamic looping mechanism of our model chromatin fiber. Fiber segments which come close to each other by diffusion are able to form temporary bonds with each other mimicing the effect of binding proteins.

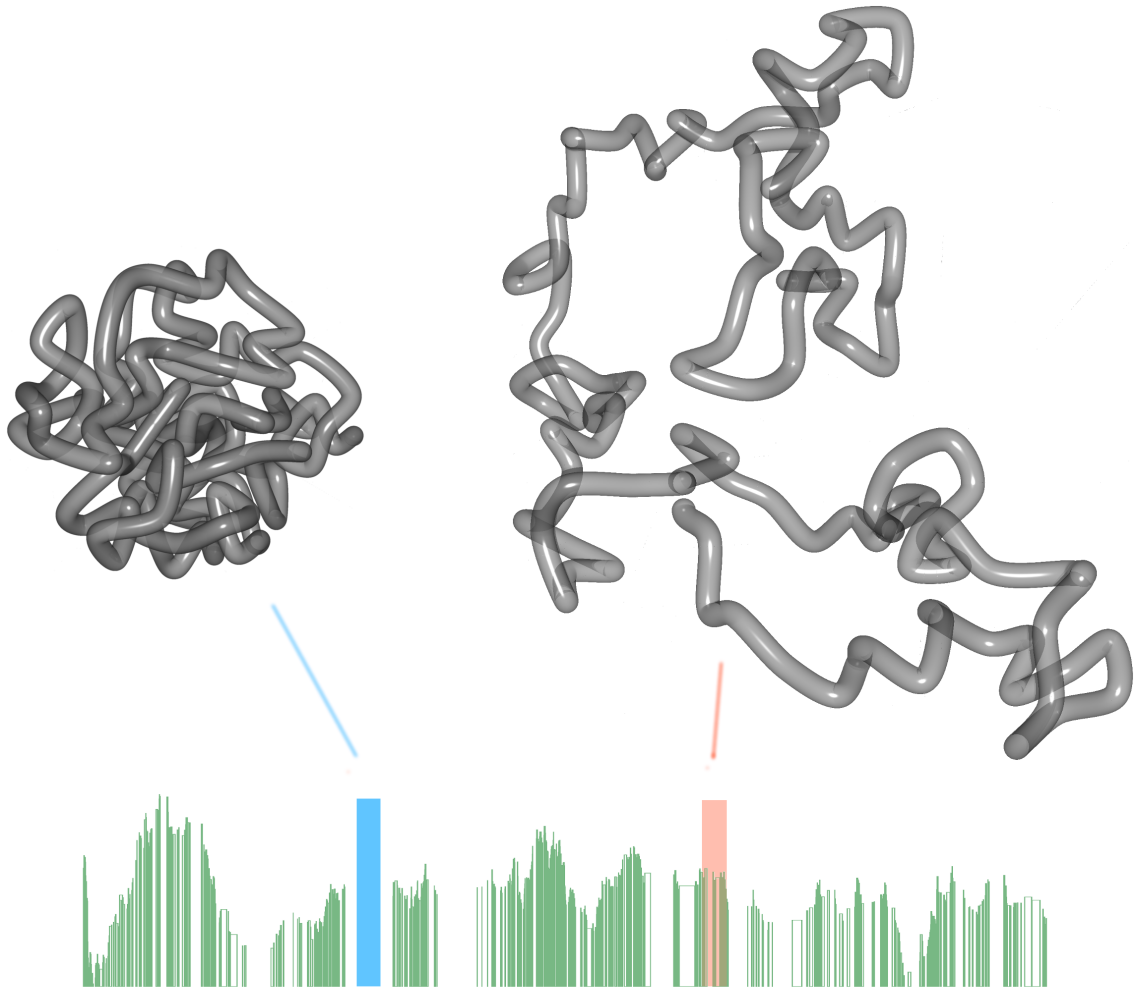

Figure 7.5: Dynamic Loop Model for chromatin subcompartment. We model the folding of the chromatin fiber in the subcompartments with the Dynamic Loop Model [184]. In this model, diffusion-driven cross-links within the model chromatin fiber and thus chromatin loops can be formed. By controlling the average concentration of loops within the subdomain we can controll the degree of compaction of the domain. Large mean loop concentrations lead to compact domains which can model silent heterochromatin. Domains with a small mean loop concentration are less compact and can model active euchromatic regions. An example conformation for a compact heterochromatic domain can be seen on the left and a more loose model conformation for an active euchromatic region is shown on the right.

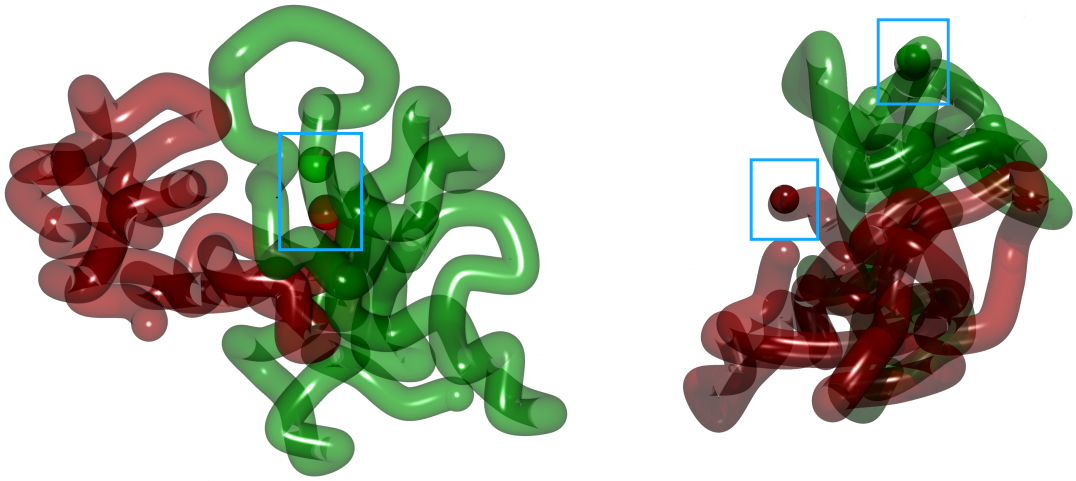

Figure 7.6: Visualization of a sub-domain before and after break is induced. The figure shows chromosome sub-domain obtained with our folding model. The chromatin fiber section upstream to the break is colored in green and the section downstream is colored in red. The two sections do not intermingle due to the formation of domains by the dynamic looping mechanism. On the left we show the conformation before the break is induced. The area that is later to be broken is located at the inside of the domain. On the right we show the conformation after induction of the break and equilibration. The break ends move to the periphery of the domain due to their entropic freedom.

The probability of bond formation is given as a model parameter p which we call looping probability. Loop lifetimes are drawn from a poisson distribution with mean value τ which is also a model parameter. We perform Monte Carlo simulations to sample equilibrium conformations of the model chromatin fiber with given parameters (p, τ). By adjusting the looping probability we can control the average number of loops within our chromosome subdomain. We define the loop concentration as the number of loops over the number of monomers in the fiber $k = \frac{n_{loop}}{N}$. High loop concentrations lead to very compact conformations while low loop concentrations result in more open conformations. We associate the compact conformations with heterochromatin, since these regions are usually highly condensed and the more open conformations with active euchromatin. Figure 7.5 shows example conformations for both types of conformations. On the left we show a very compact conformations which could for example be found in silent heterochromatic regions of the genome. On the right we have a much more open and accessible conformation which can for example be found in actively transcribed regions.

To simulate the behaviour of a DSB in our model sub-domains we create a break of the fiber by deleting the permanent bond between two central monomers of the fiber. We assume that the time scales on which proteins recognize the DSB and on which repair occurs is much smaller than the time scale needed for the complete rearrangement of the loop structure of the subdomain. Indeed repair proteins are observed to arrive at the site of DSBs within minutes [45]. Large rearrangements of the chromosome on the other hand are expected to happen on the timescale of the cell cycle [146]. We recognize this time scale separation in our model by fixing the loop structure after induction of the break. In Figure 7.6 we show an example conformation of sub-domain before and after the break of the fiber was induced. The break regions is marked with a blue box.

We first monitor the dynamics of the broken ends and compare them to the other bulk

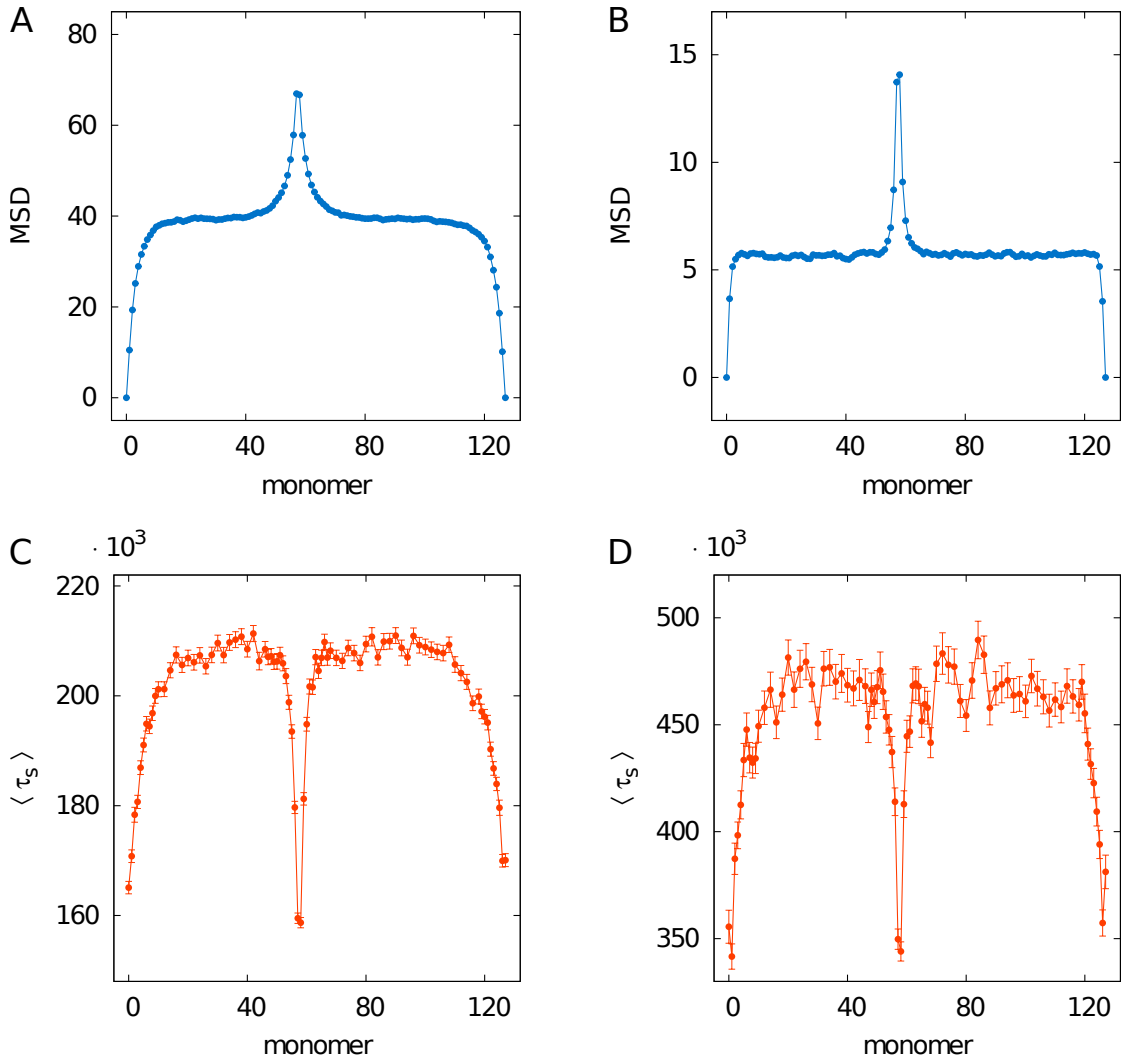

Figure 7.7: Mean square displacement for euchromatic and heterochromatic chromosome subdomains. A. The panel shows the MSD of segments in an euchromatin domain. For the simulation we fixed the ends of the domain to model the fact that the domain is attached to the rest of the chromosome. Therefore the ends have no mobility while the break ends have the highest mobilities. **B.** The profile of MSD for the heterochromatin domain is similar to the profile of the **C.** The figure shows that the average search time in a euchromatic chromosome domain. Average search time for the monomers were determined by running simulations with the monomer specified as target site for repair proteins. Our results show that the average search time at the site of DSBs is considerably decreased compared to bulk chromatin. **D.** In heterochromatin the average search times for bulk chromatin is generally higher than the search times for targets in euchromatin due to the high density of the domain. Nevertheless, the entropic freedom of a broken end gains makes it more being recognized by diffusing proteins more quickly.

monomers. We observe that in this case the mobility of the broken ends is also increased compared to the other fiber segments. This can again be explained by the entropic freedom that is created by the breakage of the fiber. Break ends are thus able to roam the space faster which facilitates their recognition by surrounding repair proteins. However, in this realistic model the aspect of accessibility plays an even more important role. As in the case of the simple polymer model we then randomly place a given number of model repair proteins in the simulation box until we reach a predefined concentration. We define a segment of the model chromatin fiber as the target area and measure the time until any of the proteins first attaches to the target. We peform a large number of these simulations and measure the mean search time for different segments of the fiber.

Figure 7.7 shows the mean search times for a highly condensed heterochromatic domain and an open euchromatic domain respectively. We see that the average search time for targets in the heterochromatic region is in general by a factor of around 4 higher than average search times in the euchromatic domains. We see that the accessibility of euchromatin means, that DSBs in these areas are much more easily recognized than in condensed heterochromatin. This could be an explanation why signs of the DSB repair process, for example the formation of γ-H2AX foci is strongly suppressed in heterochromatin [140,141]. This is also demonstrated by the difference in average search time between the natural ends of the domain and the interior of the domain. Since the natural ends of the domain are much better accessible to proteins they can much more easily be attached to by proteins. The average search time for natural ends in the euchromatin domain is decreased by around 7%. On the other hand the average search time for the natural ends in the case of the heterochromatin domain is decreased by around 30% compared to the interior of the domain.

Most importantly however, our results show that the average search times are decreased at the broken ends. In euchromatin, the mean search time for the broken ends is even lower than the average search time for the natural ends of the domain. The value at the broken ends is 19% smaller than the average search time for bulk chromatin and also 13% smaller than the search time at the natural ends of the chromosome. Our results show that recognition of DSBs in open regions of the genome is strongly facilitated due to the increased entropic freedom of the broken ends. In heterochromatin, the average search time for the broken ends is not lower than the average search time of the natural ends. However, the value for the broken ends and the natural ends is approx. 27% smaller than the search time of bulk chromatin.

We can thus conclude that the breakage of the chromatin fiber in turn facilitates the recognition of the break by diffusing proteins. Recognition is in general easier in more open euchromatic regions than in strongly condensed heterochromatic regions. The dense surrounding of the break in heterochromatin impedes the binding of proteins to the break. This is in agreement with the suggestion that molecular crowding has a strong influence on the target search process [245]. Firstly, it strongly decreases the accessibility of the break region to proteins by hindering their approaching. Additionally, the high density decreases the mobility of both the chromatin fiber including break ends and repair proteins. Therefore, the attachment of a repair protein to the broken ends in a heterochromatic region takes much more time than the recognition of a break in an open euchromatic region. We believe that this might be a reason for the suppression of γ-H2AX formation in heterochromatic regions. Nevertheless, we show that the search time for breaks in heterochromatin is still facilitated compared to bulk chromatin. We thus can conclude that the entropy based recognition of DSBs can be viewed as a stable process. The break,

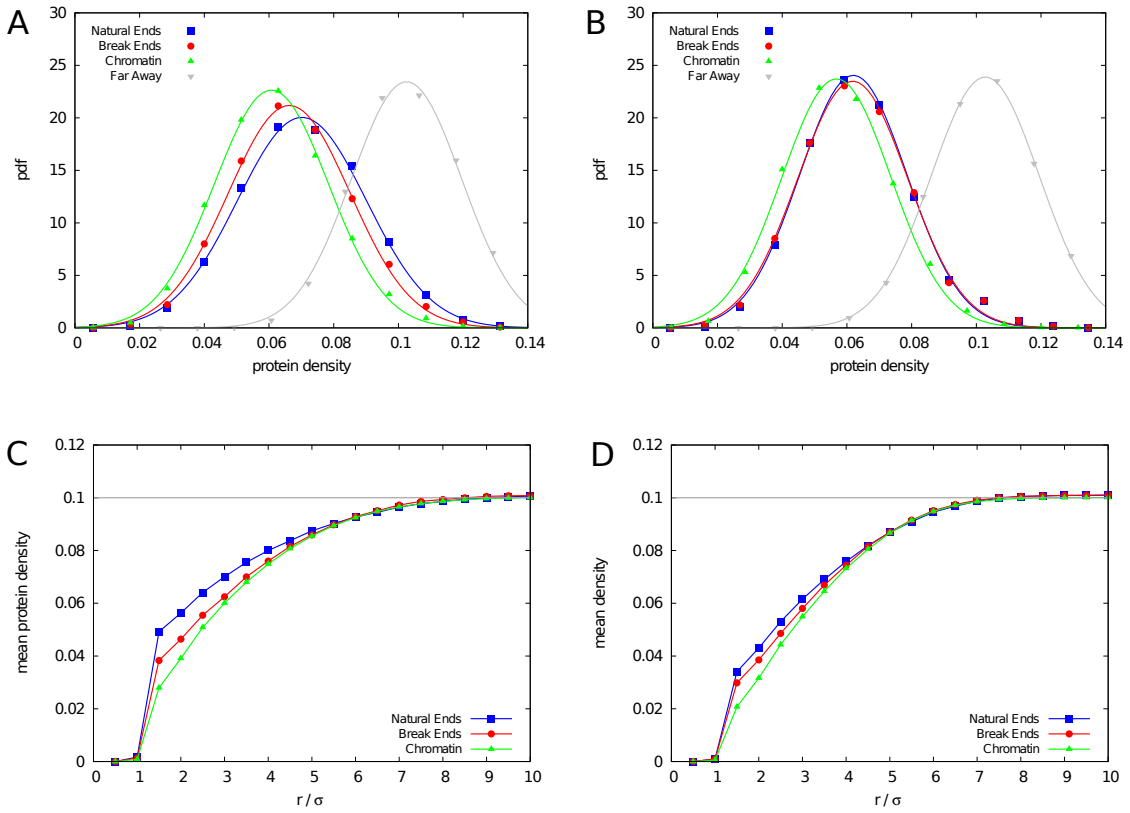

Figure 7.8: Density distribution of surrounding proteins in vicinity of different chromatin segments. A. The figure shows the density distribution of model proteins around different parts of a euchromatic model chromosome sub-domain. The mean density in the surrounding of the natural ends is the highest, followed by the density around the broken ends which is in turn higher than the density around the bulk chromatin. Generally, the density is smaller than far away from the chromosome sub-domain due to the excluded volume interactions between proteins and the chromatin fiber. **B.** The figure shows the protein density distributions for a heterochromatic model chromosome subdomain. In this case the density around the natural ends is the same as the break ends. As for euchromatin, the density is higher than around the bulk chromatin. Panels A and B show that surrounding proteins accumulate in the vicinity of broken ends. **C.** This panel shows the radial distribution of the mean density of surrounding model proteins around different segments of the euchromatic sub-domain. Due to the excluded volume interactions between model proteins and the model chromatin fiber the density of proteins is small close to the fiber. Far away from the fiber the mean density reaches a value of 0.1. The graph shows that for all distances the mean protein density around the broken ends is higher than around the bulk chromatin segments. **D.** This panel shows the radial distribution of the mean protein density for the heterochromatic sub-domain.

as a disturbance of the system, automatically leads to entropic conditions that facilitates its recognition and repair to restore the conditions before the occurence of the disturbance.

7.3.4 Repair Proteins Accumulate in the Vicinity of DSBs

In the previous section we found that the changed entropy of broken ends facilitates their detection by diffusing repair proteins in all kinds of chromosome domains. Therefore an interesting question is how the environment of a segment of the chromatin fiber generally changes if this segment is affected by a strand break. To address this question we first

analyse how the concentration of proteins in the surrounding of the break side changes after breakage of the fiber.

We first take the conformations for heterochromatin and euchromatin sampled with our dynamic folding model. In each conformation we induce a break at the same genomic positions in the center of the fiber. We place the polymer in a solution of diffusing proteins at a given total concentration c. We include steric repulsion between the chain monomers and the proteins but no other interactions. Model chromatin fiber and model proteins then move according to stochastic dynamics and we sample equilibrium conformations of this system consisting of polymer chain and proteins.

To analyse how the environment of the broken ends changes we calculate the density distribution of both proteins and chromatin in the surrounding of each segment of the fiber. Figure 7.8 shows the resulting distributions for the euchromatic chromosome subdomain with a surrounding total protein concentration of $c = 0.1$. We compare the density distributions around the natural ends of the domain, a segment in the bulk region of the domain, the broken ends of the fiber and far away from the chromatin fiber. Firstly, the protein concentration near any part of the fiber is decreased demonstrating the decreased accessibility of the domain for proteins due to the presence of chromatin. Our results show that the protein concentration is the lowest in the vicinity of the bulk chromatin region and highest in the surroundings of the natural ends. More importantly, compared to the bulk region, the protein concentration in the region close to the break is increased. Therefore, there is a clear change of the environment of the break. The entropic changes caused by a break of the fiber leads to conditions where diffusing proteins accumulate in the surrounding of the break site. One could speculate that this automatic process that is only dependent on physical principles is a mechanism for the signalling of DSBs.

7.3.5 Break Ends are Relocated to the Periphery of the Chromatin Subdomains

Our results show that the dynamics of broken ends in chromosome subdomains is increased. At the same time the entropic changes caused by a break of the fiber results in an accumulation of surrounding proteins in the vicinity of the break ends. Both effects contribute to the recognizability of a DSB. The result is that the time until a repair protein binds to the site of the DSB is decreased compared to undamaged sites.

To give further insight into this process of recognition of DSBs we examine the positioning of DSBs in their chromosome sub-domain. For this we calculate the mean square distance of each segment of our model fiber to the center of mass of the model fiber. Figures 7.9 and 7.10 show the results for euchromatic and heterochromatic domains, respectively. We observe that the distance to the center of mass is highest for the natural ends and the bulk region is naturally closer to the center of mass of the domain. This means that segments that are in the middle of the domain are physically oriented further to the center of the domain. In the heterochromatin domain the plateau region for the bulk chromatin is much broader compared to the euchromatin domain. While for the euchromatin the plateau is rather short and large portions of the fiber towards the natural ends are physically oriented further to the outside of the region, in heterochromatin we see a broad plateau. A large part of the fiber has therefore the same average distance to the center of mass. This effect is caused by the high compaction of the heterochromatin subdomain. This compaction leads to globule-like configurations that are very spherical. Therefore, the fiber is strongly intermingled with itself and thus nearly all segments in the bulk have the same average radial distance to the center of mass of the domain.

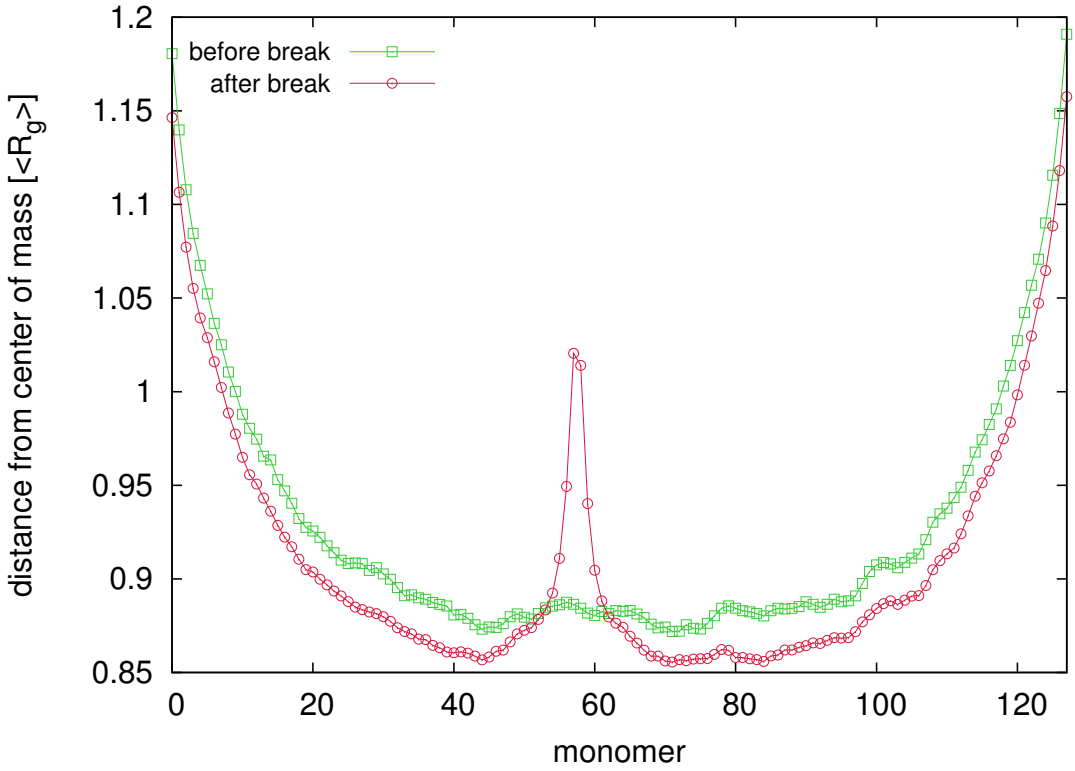

Figure 7.9: Distance of chromatin segments from the center of mass of the euchromatin subdomain. This figure shows the distance of each monomer in the chromatin fiber to the center of mass of the euchromatic subdomain. The green curve are the distances before the induced break and the purple curve are the distances after the induced break. In the unbroken state, the natural ends are located further to the outside of the domain and the inner parts of the fiber are located closer to the center of mass. There are long transition parts of the fiber from the inside to the outside. After break induction the break ends locate further to the outside of the domain but are on average still not as far outside as the natural ends.

As in the previous simulations we induce a break in each of the conformations. We assume that the time scale on which DSB repair is happening is much shorter than the time scale for a complete rearrangement of the loop structure of the domain. The time scale for global movements and conformational changes of a chromosome are in the range of hours to days and may require cell cycle activities [146].

Due to this assumed time scale separation we fix the loop structure of the conformation after inducing the break. The results for the damaged chromosome domains are also shown in Figure 7.9 and 7.10. The mean distance to the center of mass for the broken ends is now clearly much higher than for the remaining bulk segments. This means that despite the unchanged loop structure in the domain, broken ends are drifting to the outside of the domain due to the altered entropic conditions. This effect is more pronounced in the heterochromatin domain than in the euchromatin domain. Multiple experiments have shown that fluorescent foci seem to be repelled from heterochromatin domains [142]. Our results indicate that the entropic conditions that a break of the fiber in a highly condensed domain cause could at least contribute to the relocation of the break ends to the periphery of the domains, where the chromatin density is smaller. This relocation could in turn facilitate the recognition of the breaks due to the decreased shielding effect of the other

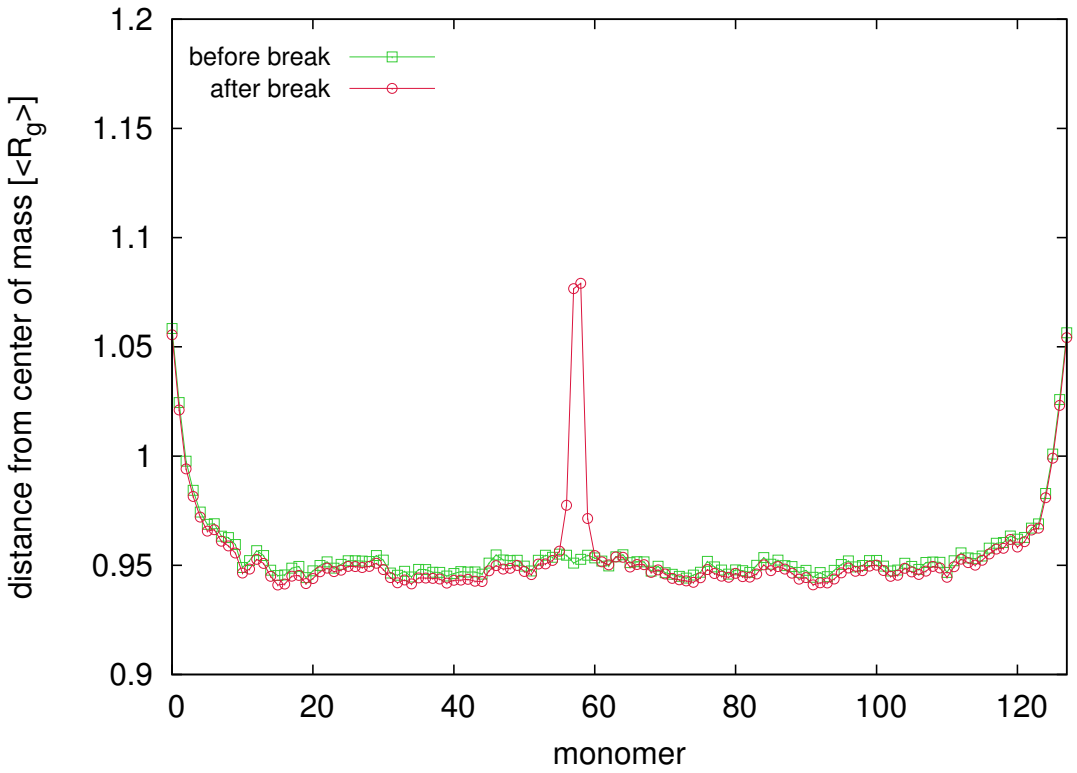

Figure 7.10: Distance of chromatin segments from the center of mass of heterochromatin subdomain. In the case of the heterochromatic chromosome domain the natural ends have also a larger distance to the center of mass. Here, we do not observe long transition parts. Instead, nearly the whole bulk of the domain has the same average positioning in the domain territory. This is a result of the high compaction of the domain, which takes globule-like conformations. The effect of repulsion of the broken ends from the inside of the territory to the outside is even stronger than in the euchromatic phase. Broken ends have on average an even larger distance to the natural ends of the fiber.

fiber segments.

7.4 Conclusion

In this work we used a coarse grained approach to model chromosome sub-domains with different chromatin densities. We then analyzed the behaviour of strand breaks in our model sub-domains. We represented repair proteins by diffusing particles and measured the mean search time τ_s for the proteins to find a pre-defined target segment on the fiber. We found that the mean search time for the broken ends of the fiber is decreased compared to the rest of the model chromosome. This decreased search time is facilitated by the reduced entropic constraints of broken fiber ends. The gained entropic freedom of the broken ends leads on the one hand to a higher mobility of the break ends and in turn to decreased mean search times. On the other hand, we found that broken chromatin ends quickly diffuse out of the center of their domain to the surface. The chromatin fiber, just as other proteins cause a molecular crowding effect that can efficiently shield the target area from searching proteins. We find that although the density of surrounding proteins outside of the domain is enhanced, the total density of proteins and chromatin is

decreased. The molecular crowding effect for chromatin segments that are located at the surface of the domain is thus smaller than for segments which are located in the center of the domain. Therefore, the passive transport out of the domain also facilitates the search process for broken chromatin ends. At the same time, chromatin loops limit the distance that the broken ends can gain from each other and thus prevent them from driftig away from each other. We suggest that the detection of DSBs can be seen, in a way, as a stable mechanism. In this picture, the strand break can be viewed as a disturbance to the system. This disturbance itself changes the conditions of the system in such a way that the automatic reaction of the system helps with the removal of the disturbance. Therefore, we believe that the entropic organization of the chromatin fiber plays a very important role for early DSB detection.

Chapter 8

Structural Changes and Healing of Irradiated Cells

References

The content of this chapter was adapted from the following publication

- Y. Zhang*, G. Máté*, S. Hillebrandt, P. Müller, M. Hausmann and D.W. Heermann, *Measuring Structural Changes in Chromatin Induced by Ionizing Radiation.* in preparation.
  * shared first authorship due to equal contribution

Chapter Summary

Ionizing radiation (IR) can induce many types of damages to the genome. Amongst the most toxic ones are DNA double-strand breaks (DSBs). The cells damage response to the emergence of radiation induced DSBs involves the fast recruitment of repair proteins and local rearrangements of chromatin on different levels. Highly compacted heterochromatic regions were observed to be refractory to the formation of radiation-induced foci. Instead, areas containing double-strand breaks seem to be located to the periphery of heterochromatic domains. Local decondensation of heterochromatin was suggested as a mechanism that facilitates this transport process. In this work, we analyzed localization microscopy images by means of statistical physics and graph theory to provide a quantitative description of structural changes induced by IR. Stably transfected HeLa cells expressing GFP labelled histone 2B (H2B) were exposed to different doses of IR. Simultaneously, fluorscently labelled, specific antibodies were used to mark heterochromatic regions. Afterwards, these markers were localized by spectral position determination microscopy with high precision. We then used statistical physics methods and a graph theoretical approach to analyze the structure of the marker positions. Our results show that HC regions undergo a relaxation immediately after exposure to IR while the overall chromatin structure does not change. We further demonstrate that at later times after irradiation, these alterations become less pronounced.

8.1 Introduction

Double strand breaks (DSBs) are probably the most harmful damages the DNA fibre may suffer as they can lead to the rearrangement of the genome [249]. DSBs are caused by ionizing radiation, oxidizing agents, replication errors and certain metabolic products of the cell [54]. As a response to the DSB, the repair process can take different pathways, the chosen one mostly depending on the phase of the cell cycle. It has been observed experimentally, that as a reaction to the presence of a DSB, the chromatin opens up in the region affected by a DSB [133]. This might be viewed as a mechanism which facilitates the repair.

Since DSBs are the causes of many pathologies, it is crucial to understand the repair mechanisms and the occuring structural reorganization. In this project, we will investigate the effects of ionizing radiation on the structure of the chromatin. Although biological implications are very important, in the following we emphasize the analysis approaches we used to gain an insight in the occurring structural changes.

8.2 Methods

8.2.1 Experimental Data

Experiments were performed on HeLa cell cultures. Cells were transfected to stably express green fluorescence proteins (GFPs) on histone H2B. Furthermore heterochromatic regions were marked with antibodies tagging the histones 4 (H4) that are methylated at Lysine 20. This specific histone modification is indicative for heterochromatic areas. Images were captured by Spectral Position Determination Microscopy (SPDM), a relatively recently developed super resolution technique. As a result of the experiment, the most probable location of the fluorophores is recorded as pairs of x, y coordinates in a two dimensional plane. The two dimensional points represent the projection of a roughly 600 nm thick slice from the nucleus.

Cells were irradiated with doses of 0.5, 2 and $4\,Gy$. After irradiations cells were fixated either at $30\,min$ or after allowing $48\,h$ of repair time. Furthermore, a control group of untreated cells was also imaged.

8.2.2 Segmentation and Masking of Images

In the first step of the analysis we perform a segmentation of the image in order to detect regions of interest. In this process we exclude the areas that have a considerably lower than average number of localized fluorophores. This includes the area that is on the outside of the cell nucleus but also areas inside the nucleus that only have a small chromatin content, e.g. at the site of nucleoli.

In the image, the fluorophores or antibodies represent a set of points in the two-dimensional plane. To perform the segmentation we calculate a density distribution for the fluorophores on the entire image by applying a gaussian kernel density estimation [250]. This means that for each point on the image we set a two-dimensional gaussian probability distribution with the coordinates of the point as its mean value. We choose radially symmetric gaussian distribution by setting the bandwidth matrix to a multiple of the unit matrix. The multiplier is a parameter in the model which defines a correspondence between the standard normal distribution and the physical scales in the experiment. We set it to $40\,nm$. The sum of all the gaussian probability distributions is then normalized

yielding a probability distribution of the position of the points from the image. Since the values of the probability distribution are proportional to the density of points, the obtained probability distribution in fact represents the spatial density distribution in the nucleus.

Using the density distribution we calculate a mask that accepts all areas with a density that is above 25% relative to the lowest value and blanks out all other areas. In order to make sure that also the low density regions at the border of the cell nucleus are not included, we further erode the masked area by another $150\,nm$. Thus we efficiently exclude areas that are outside of the cell nucleus and areas which have a low marker density indicating sites of nucleoli.

8.2.3 Calculated Measures

In order to analyse the spatial arrangement of the points, we calculate different measures which may characterize the structure. Besides standard measures like the radial distribution function [145], we construct a spatial graph describing the neighbouring relations and analyse the structure of this graph.

Radial Distribution Function

The radial distribution function $g(\mathbf{r}_1, \mathbf{r}_2)$ is a measure of structure for many particle systems such as liquids. It is the probability distribution function that two particles are at the positions $\mathbf{r}_1$ and $\mathbf{r}_2$. Since absolute positions are not important, the correlation function is in fact only dependent on the directed distance between the two positions $\mathbf{r} = \mathbf{r}_{12} = \mathbf{r}_2 - \mathbf{r}_1$. Thus the radial distribution function can be simply written as $g(\mathbf{r})$ and is given by

$$g(\mathbf{r}) = \frac{1}{N} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} \delta(\mathbf{r} - \mathbf{r}_{ij}) \right\rangle \tag{8.1}$$

with summations over all N particles of the system and $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$. The probability to find a second particle in the directed distance $\mathbf{r}$ of the first particle is $g(\mathbf{r})d\mathbf{r}$. Assuming an isotropic system, the radial distribution function only depends on the undirected distance $r = |\mathbf{r}|$ and the probability to find a second particle in the distance r of the first particle is given by $4\pi r g(r) dr$.

For each image the radial distribution function is determined by calculating all pair distances between points inside of the masked regions of interest. Due to the finite system size, we cannot use all points for the calculation of $g(r)$ since points that are close to the border of the masked image do not have the correct surrounding. We therefore first reduce the masked image isotropically from the border by a distance d_{shrink} thus yielding a second masked image that is considerably smaller than the first one. All points in the second mask have then equal environments for distances up to d_{shrink}. For each of the points in the second mask we then calculate the distances to all other points that lie in the original mask. The normalized distribution of these distances then give the radial pair distribution function $g(r)$ for the image.

We then average over the complete set of images to obtain the average $g(r)$ for each experimental setup. The standard deviation of the mean value is then used as a measure for the uncertainty of the radial distribution function $g(r)$ at each distance r.

The Spatial Graph of the Neighbourhoods

Another possibility to gain insight in the structure presented by the localization data is to build a graph on the skeleton defined by the localized points. We then investigate this graph by means of graph theoretical methods and conclude the topological or relational properties of the underlying structure.

Graphs are abstract mathematical objects designed to capture interdependencies of certain entities [251]. Entities are represented by nodes (or vertices) while the connections are encoded by edges between nodes. A graph is usually denoted by $G(V, E)$, where V is the set of nodes and $E = \{(a, b)|a \in V, b \in V\}$ is the set of edges. Graphs have been used to study a variety of structures and phenomena [252–256], most of these studies relying on the mathematical field of graph theory.

We are mainly interested in observing local properties of the geometric arrangements formed by the points. Therefore, we choose the graph building procedure so that the resulting graphs emphasize the local relations among the fluorophores. We build the nearest-neighbor graph (NNG) in which each node is connected with its first-order neighbours.

There are many possible ways to build the NNG for a given set of points. The most widely used approach is the construction of the Delaunay triangulation [257]. However, for a better understanding, let us give a more intuitive description regarding the construction of the Delaunay triangulation. We start again by introducing the dual of this triangulation, the Voronoi tessellation [258]. By definition, the Voronoi tessellation of a set of points S is a tessellation in which each Voronoi cell V_i corresponding to the site S_i consists of all the points of the space closer to S_i than any other site. The faces of the Voronoi diagram consist of all the points in the space that are equidistant to sites corresponding to touching Voronoi cells. Thus the Voronoi cell V_i defines the space dominated by the site S_i. The Voronoi tessellation can be transformed into the Delaunay triangulation by connecting the sites of the neighbouring Voronoi cells. Two points will be considered neighbours if they are directly connected by an edge of the Delaunay triangulation or equivalently, if their corresponding Voronoi cells are touching. In Figure 8.1 we illustrate the Voronoi tessellation and the Delaunay triangulation for a given set of points.

We construct the NNG defined by the Delaunay triangulation for each experiment and calculate three properties of the obtained graphs: the degree distribution $h(d)$, the rescaled probability density $f(r)$ of the edge lengths and the conditional probability $p(r|r')$ of the edge lengths.

The degree of a node is defined as the number of connections of the node. The degree distribution is the probability distribution of the degrees over the whole graph. Since the degrees are integer numbers, $h(d)$ is a discrete distribution. It is in fact the histogram of the degrees of the nodes normalized such that the values in the histogram bins sum up to unity. That is, if $\bar{h}(d)$ is the frequency of the degree d then the degree distribution is given as

$$h(d) = \frac{\bar{h}(d)}{\sum_{d'=0}^{d_{max}} \bar{h}(d')}, \tag{8.2}$$

where d_{max} is the maximal degree in the graph.

Let us denote the probability distribution of the length of the edges in a given graph $G(E, V)$ obtained by the Delaunay triangulation of a point-set corresponding to one of the localization images by $\bar{f}(r)$. For different experiments, we expect to get different $\bar{f}(r)$ distributions. These differences may stem either from different underlying structures or

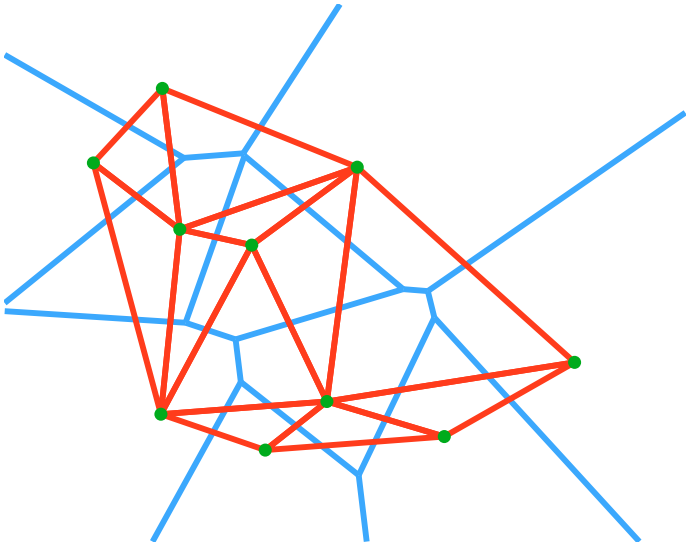

Figure 8.1: The Delaunay triangulation and its dual, the Voronoi tessellation for a random set of points. The blue lines are the segments of the Voronoi tessellation, the red ones are the edges of the Delaunay graph (triangulation)

from different experimental conditions as these can vary from experiment to experiment (slightly different concentration of stains or different microscope gain). Provided that different experimental conditions have a liner effect on the density ρ of marked sites, variations in $\bar{f}(r)$ can be eliminated. This is achieved by rescaling the distributions to a reference density ρ_0. For this, we apply the following procedure: Since we are calculating the $\bar{f}(r)$ distributions, we are, in fact, treating the edge lengths in a probabilistic manner, that is, r is considered a random variable. Thus, the rescaling of the distributions to a reference density ρ_0 is in fact a probability transformation over r. To see this, let $d_{ij} = d[(x_i, y_i), (x_j, y_j)]$ denote the length of the edge between vertexes i and j. If we multiply the coordinates (x, y) of all the points with an positive real number $\alpha = \rho_0/\rho$, the following relation holds:

$$d[(\alpha x_i, \alpha y_i), (\alpha x_j, \alpha y_j)] = \alpha d[(x_i, y_i), (x_j, y_j)], \tag{8.3}$$

On the other hand, this multiplication corresponds to a uniform dilation (or contraction) of the system, that is, a uniform scaling of the density. Therefore, to calculate the scaled probability density, we have to apply the probability transformation defined by

$$t(r) = \alpha r. \tag{8.4}$$

Applying basic probability theory, we obtain the rescaled probability density

$$f(r) = \frac{1}{\alpha}\bar{f}(\frac{1}{\alpha}r). \tag{8.5}$$

We can either define a reference density ρ_0 and calculate the corresponding transformations for the different experiments, or, fit the probability densities to each other, using α as a fitting parameter.

Certain local properties are averaged out both by the $g(r)$ and the $f(r)$. For instance, non-regular density variations are not captured by these measures. Another example is

a situations when certain regular structures appears multiple times but the size of the structures varies.

In order to detect these situations, we calculate the conditional probability $p(r|r')$ of the edge lengths of the nearest neighbour graph. This conditional probability is the probability of finding an edge with length r attached to a node which for sure has an edge with length r'. The conditional probability can numerically be represented in a matrix structure P where each row corresponds to a condition for a given interval over r' and each column represents an interval over r. The matrix entry P_{ij} corresponding to a given row i and a column j will give the probability of finding an edge with a length between r_j and $r_j + dr$ attached to a node which has at least one edge with a length between r'_i and $r'_i + dr$. In case the localization points are arranged according to a specific structure, the P matrix will indicate a preferential attachment. For instance, in a long chain which has constant link-lengths over larger domains, but different domains have different link lengths, the P matrix will be almost diagonal. Figure 8.2 shows the P matrix for two randomly generated data-set, one set containing points with coordinates distributed uniformly while the other set has additional Gaussian clusters. As the figure illustrates, uniformly distributed points will produce a rather uniform matrix, while point-sets with clusters will have a more emphasised diagonal.

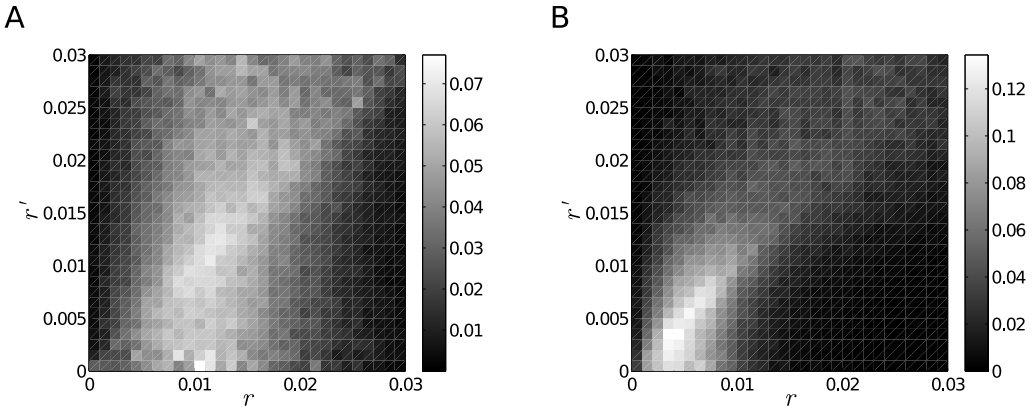

Figure 8.2: Conditional Probabilities of Edge Lengths for Random Data. The figures illustrate how the $p(r|r')$ conditional probability looks for randomly generated point positions. **A**. Conditional probability of points with coordinates generated according to a uniform distribution. **B**. Conditional probability of points with coordinates generated according to a mixture of uniform distribution and clusters of Gaussian distributions. In the latter example we observe an emphasised diagonal which is the result of the Gaussian clusters. Tightly packed points tend to produce short edges, while points from the edges of the clusters mostly have longer edges.

8.3 Results and Discussion

8.3.1 Exposure to Ionizing Radiation Cause Local Changes

We first perform a segmentation of each of the images to cut out those areas in the image that do not belong to the cell nucleus. We also neglect areas inside the nucleus that have a

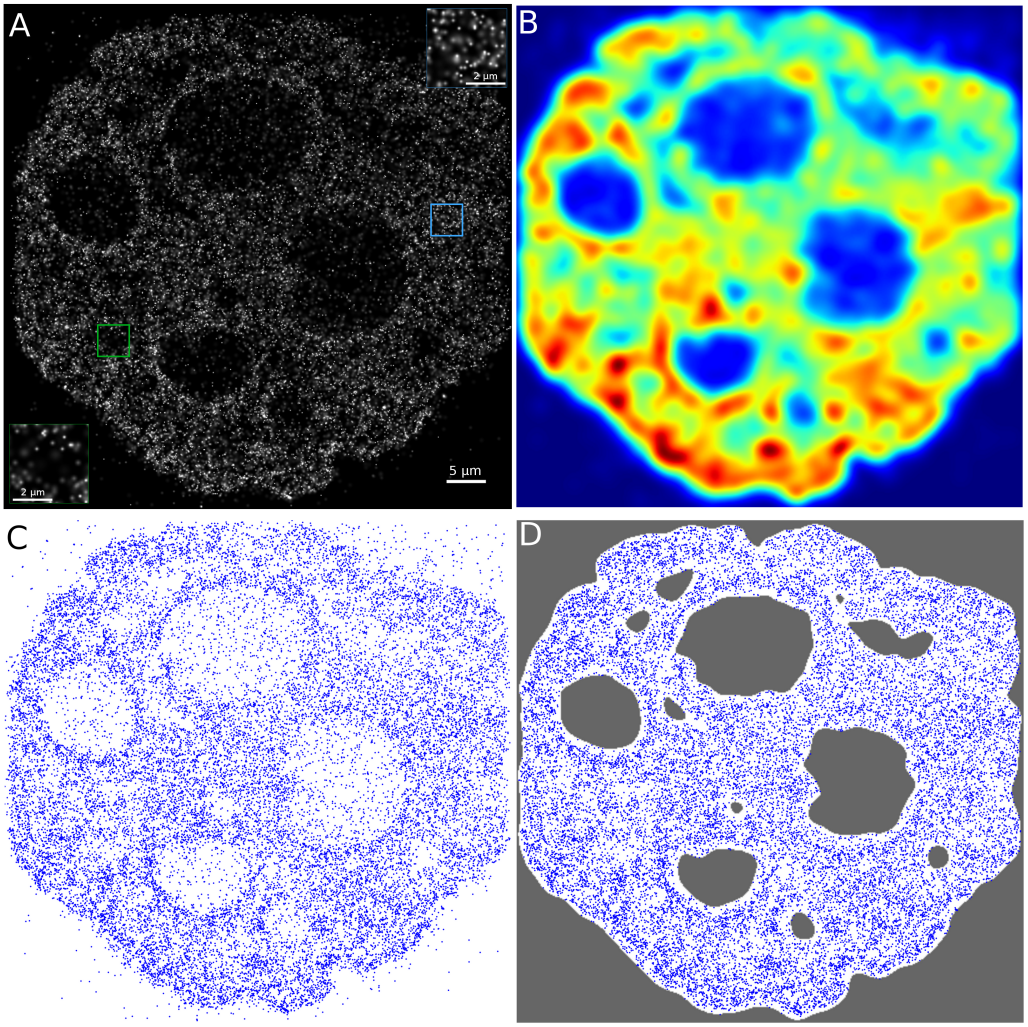

Figure 8.3: Localization Microscopy Images. **A**. The green inset shows an area with a low average density, which could correspond to euchromatin and the blue area corresponds to an area with high point density, possibly belonging to heterochromatin. **B.** This panel shows the calculated density distribution of the localized markers using a gaussian kernel density estimation with a uniform gaussian kernels. **C.** The figure shows the localized points of the image. From the points, areas with very low point density possibly corresponding to nucleoli can clearly be made out visually. **D.** Shown is the segmented image where only the area of interest is kept for the analysis. The segmentation was based on the intensity distribution with areas below an intensity threshold were discarded for analysis.

very low density of markers since these areas are indicative for nucleoli. The segmentation is based on the calculation of the density distribution of markers in the image. We only take the areas with a density of more than 25% of the largest density into consideration. Figure 8.3 shows the image with the localized H2B histones and their density distribution and the final segmented image.

To assess how the overall organization of chromatin changes after cells are exposed to ionizing radiation compared to untreated cells. We calculate the radial distribution function $g(r)$ for localized H2B histones in untreated cells and cells that were exposed to ionizing radiation. The radial distribution function is a measure for the positional

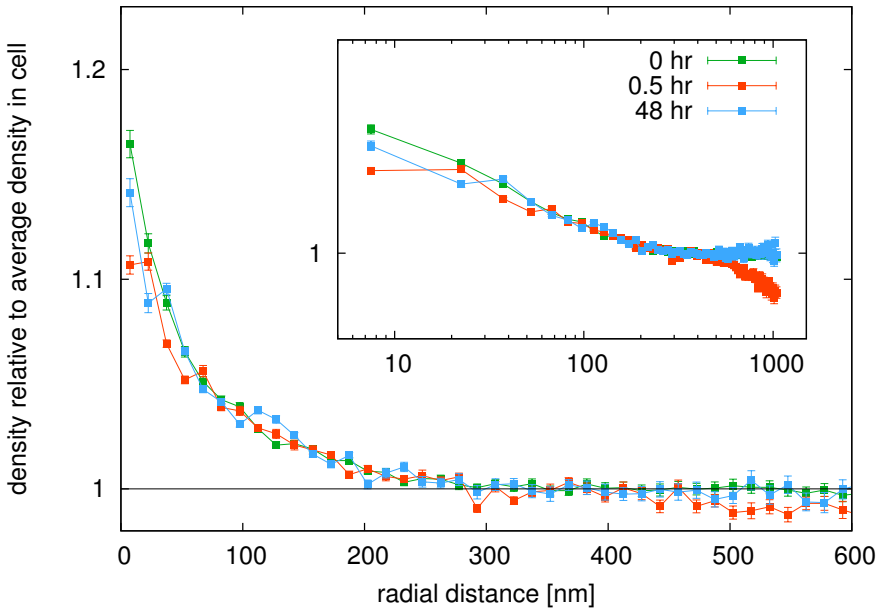

Figure 8.4: Results at different times after 0.5 Gy radiation dose. The radial pair correlation function $g(r)$ shows that correlations between the positions of labeled H2B histones exist up to a distance of roughly $300\,nm$. Marker and thus chromatin densities in the surroundings of each marker is elevated compared to the average density of markers in the cell nucleus. Above $300\,nm$ however, histone positions are uncorrelated and the marked histones can be viewed as being positioned randomly relative to each other. Furthermore, the radial pair correlation function for cells exposed to $0.5\,Gy$ γ-irradiation is the same as for untreaded cells, regardless of the time passed after irradiation.

correlation of points in a many-particle system. At this, $2\pi r g(r)dr$ denotes the probability of finding two points separated by a distance between r and $r + dr$ for infinitesimal dr.

Results for the radial distribution function of H2B markers in untreated and irradiated cells are shown in Figure 8.4. We analyzed images of untreated cells and cells fixed $30\,min$ after they were exposed to $0.5\,Gy$, $2\,Gy$ and $4\,Gy$ of γ-irradiation. For all setups we can see deviations of $g(r)$ from unity only for distances up to $300\,nm$. It thus shows that structured organisation of chromatin in the cell nucleus is only apparent up to distances of roughly $300\,nm$. Below this critical distance, locations of labeled histones H2B are visibly correlated. However, on distances larger than $300\,nm$ the radial correlation function drops to unity and the locations of histones can be viewed as being randomly distributed relative to each other. The radial distribution functions for the H2B histones are lying on top of each other and no deviations can be seen for cells that were exposed to ionizing radiation.

In order to assess the local positional correlations that are apparent up to a length of $300\,nm$ we pursued a graph theoretical approach. We first calculated the nearest neighbour distance distribution of the localized fluorophores. For this we performed a Delaunay triangulation to obtain *nearest neighbour graphs (NNGs)* for all points in the segmented images and then calculated the length distribution of the edges of the NNGs. Results for H2B markers for untreated cells and cells fixed $30\,min$ after exposure to ionizing radiation are shown in Figure 8.5. We observe that there is a significant difference between them, which can be seen in the inset. However, the distributions belong to the same family since the rescaled distributions $f(r)$ are exactly the same.

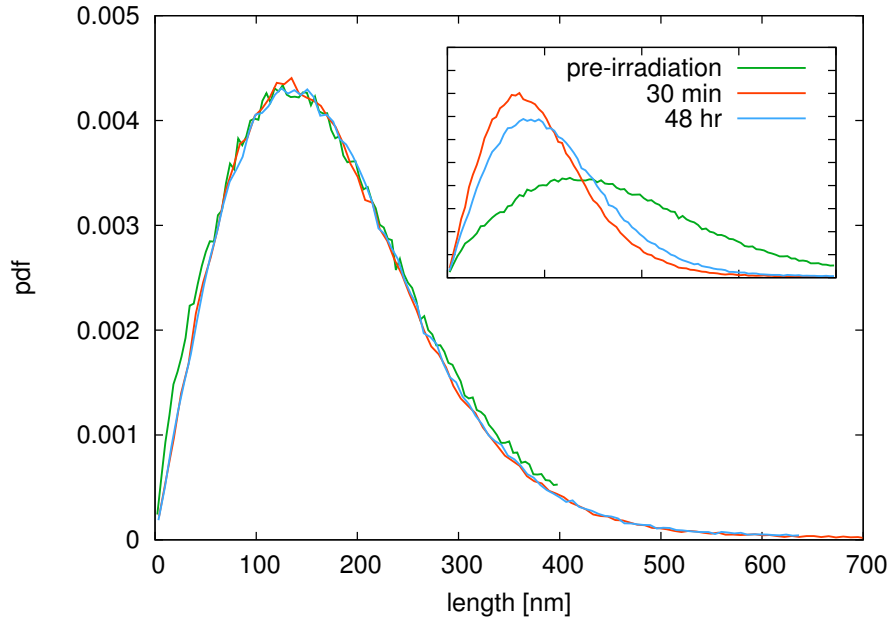

Figure 8.5: Results at different times after 0.5 Gy radiation dose. Shown are the rescaled distribution of the length of edges in a Delaunay triangulation of the H2B marker positions. The inset shows the original distributions for the different images. The differences in the edge length distributions are due to the different marker densities in the images. We therefore performed a rescaling of the distributions with respect to the point density to clear out this effect. The rescaled distributions can be seen in the main panel and show that the distributions belong to the same family.

Therefore we can state that the observed differences may stem either from different experimental conditions such as different overall marker density or a uniform dilation/contraction of the system.

We also calculate the $p(r|r')$ conditional probability for the localized H2B markers. The matrix representations P of the conditional probabilities are plotted in Figure 8.6, the three panels corresponding to pre-irradiation configuration and to configurations fixed 30 minutes and 48 hours post-irradiation respectively. While the plots resemble a mixture of the plots from figure 8.2, they do not exhibit big discrepancies amongst each other. Figure 8.7 presents the differences between the three panels from 8.6. This plot shows that for samples imaged shortly after irradiation the diagonal of the conditional probability is more prominent but for samples recorded after a longer healing time the diagonal recedes. This means that, while upon irradiation the system changes towards a less uniform structure, that is, towards a less homogeneous P matrix, with longer healing time the changes are reverted.

Our results demonstrate that positional correlations of H2B histones are not altered by DNA damage caused by γ-irradiation. As H2B histones are distributed homogeneously along the genome we can conclude that ionizing radiation does not alter the overall organization of the chromatin in the cell nucleus.

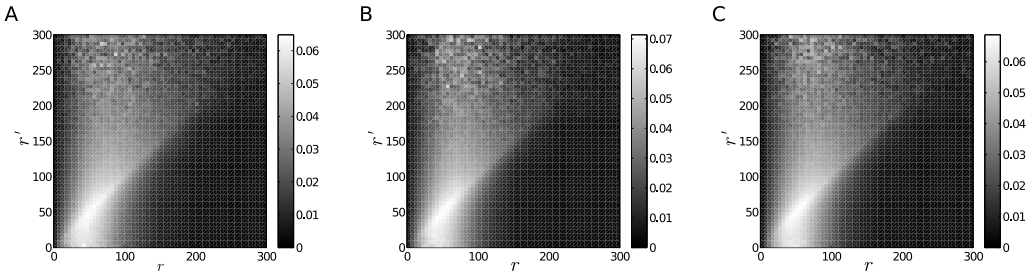

Figure 8.6: Conditional Probability Distribution of the Edge Lengts. **A**. The panel shows the conditional probability distribution $p(r|r')$ for the H2B markers before irradiation. The relatively prominent diagonal indicates locally a varying density. **B**. The panel shows the conditional probability for the H2B markers 30 minutes after irradiation. **C**. The panel shows the conditional probability for the H2B markers $48\,h$ after irradiation.

8.3.2 Heterochromatic Regions Show a Decondensation upon Irradiation

Antibodies for H4K20 made visible with the same fluorescence technique were used as markers for heterochromatic regions. Microscopy for antibodies were performed at the same time with the cells on the stably expressing histone H2B variants. Images were processed and segmented with the same methods described above and in the Methods subsection. In Figure 8.8 we show the density distribution of the heterochromatin antibody markers on panel A and the segmented image on panel B. The density distribution shows clearly how heterochromatic regions can be seen as coarse clumps located within the nucleus. These clumps are visible as bright spots in the density distribution.

The calculated radial distribution function for localized H4K20 antibodies in untreated cells and cells exposed to $0.5\,Gy$ irradiation after different times are shown in Figure 8.9. The correlation function for the antibodies show apparent differences to the correlation function for H2B histones. The high value of up to 20 times the average density that can be found at small distances r reflects the fact that antibodies are clustered together in small spots. The quick decay of $g(r)$ to the average density in the nucleus indicates that the clumps are relatively small confirming the visual impression of the images.

Upon irradiation, the value of the correlation function drops for small distances r. This means that the average density of antibodies in the surrounding of each antibody is lower after exposure to ionizing radiation. It indicates that the overall density becomes smaller in the heterochromatic clumps. We can conclude that heterochromatic regions on average become less compact and the strongly compacted organisation opens up and adopts a more loose structure upon exposure to ionizing radiation. We observe an average drop of $70\,\%$ of the mean antibody density in a sphere with a radius of $30\,nm$ around an antibody for cells irradiated with $0.5\,Gy$ γ-irradiation after $30min$. Therefore, initial repair of DNA double-strand breaks caused by irradiation in heterochromatin seems to require a drastic decrease of the chromatin density and a strong relaxation of the compact organisation of the chromatin fiber here. $48\,h$ after irradiation, this value is only at around $30\,\%$ which indicates that structures seem to have recovered after successful repair of DNA damages.

The conclusions drawn from the radial distribution function are verified by the graph theoretical analysis. In Figure 8.10 the length distribution of the edges in the Delaunay triangulation of the marked H4K20 antibodies is shown. The distribution for unirradiated cells shows a very characteristical peak at small distances centered at around $30\,nm$. This emphasizes that there is a characteristical nearest-neighbour distance for the H4K20

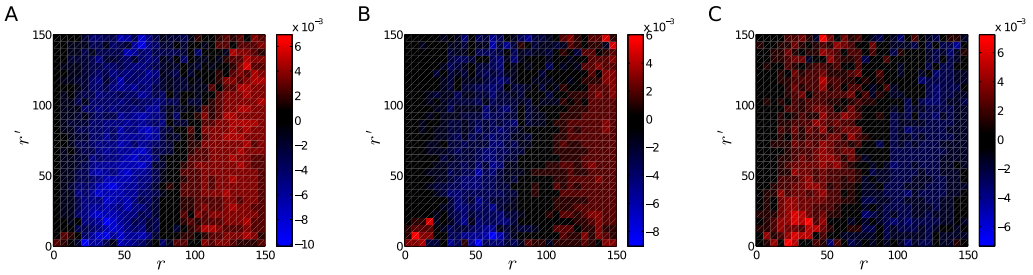

Figure 8.7: Difference in the Conditional Probability Distribution of the Edge Lengts. **A**. The panel shows the difference of the conditional probabilities $p(r|r')$ measured for structures recorded before irradiation and for structures registered 30 min after irradiation. Shades towards red indicate values which are larger in the samples before irradiation, while values towards the shades of blue indicate probabilities which are larger in images registered 30 min after irradiation. The plot indicates slightly increased values along the diagonal after irradiation. This might mean a slightly increased clustering of the points. **B**. The panel shows the differences of the conditional probability before irradiation and 48 h after irradiation. Entries in shades of red are larger in samples recorded before irradiation, while entries in shades of blue are larger in the samples recorded 48 h after irradiation. The trend is similar to that observed in panel **A**, however differences are less prominent. **C**. The panel shows the differences if the conditional probability measured 30 min and 48 h after irradiation respectively. The red shades indicate larger probabilities in samples recorded 30 min after irradiation while blue shades indicate larger probabilities in samples registered 48 h after irradiation. Here we observe a reversed trend compared to panel **A**.

markers thus meaning a preference of them to form clumps.

In cells at 30 min after exposure to ionizing radiation, the sharp peak in the edge length distributions vanishes. Instead, the distribution becomes a uniform distribution up to large lengths. The disappearance of the peak is a clear indicator that the heterochromatic regions are no longer organized as compact clumps in the nucleus. This difference can be seen visually in Figure 8.8C. Compared to the untreated cell nucleus shown in Figure 8.8A we clearly observe that the small bright spots have mostly vanished and the heterochromatic regions in the irradiated cell are much more smeared out. Our graph theoretical analysis therefore verify our observations of the behaviour of the radial distribution function. The heterochromatic regions undergo a decondensation upon exposure to ionizing radiation.

At 48 h after irradiation, the peak in the edge length distribution emerges again. Just as the pair correlation function is again very similar to untreated cells, the edge length distribution has now also the same shape as in the case of untreated cells.

Calculating the conditional probability $p(r|r')$ we see that the corresponding matrix-representations are strongly diagonal (figure 8.13). This is due to the tight clusters of the antibodies marking heterochromatic regions. This structure of the conditional probability matrices support our previous conclusions. Upon irradiation the conditional probability (panel **B.** in figure 8.13) becomes more homogeneous, indicating a more uniform structure. 48 h after irradiation the conditional probability is again diagonal (panel **C.** in the same figure). Note that in this case the colour-map has a wider range and values on the diagonal are in fact very close to values from panel **A.**, except for very small radii.

The differences of the conditional probabilities are plotted in figure 8.14. While the changes upon irradiation barely depend on the condition (the value of r'), we observe that before irradiation smaller radii were more abundant (red colour in panel **A.**). The difference between the structure 30 min after irradiation and the structure we see 48 h

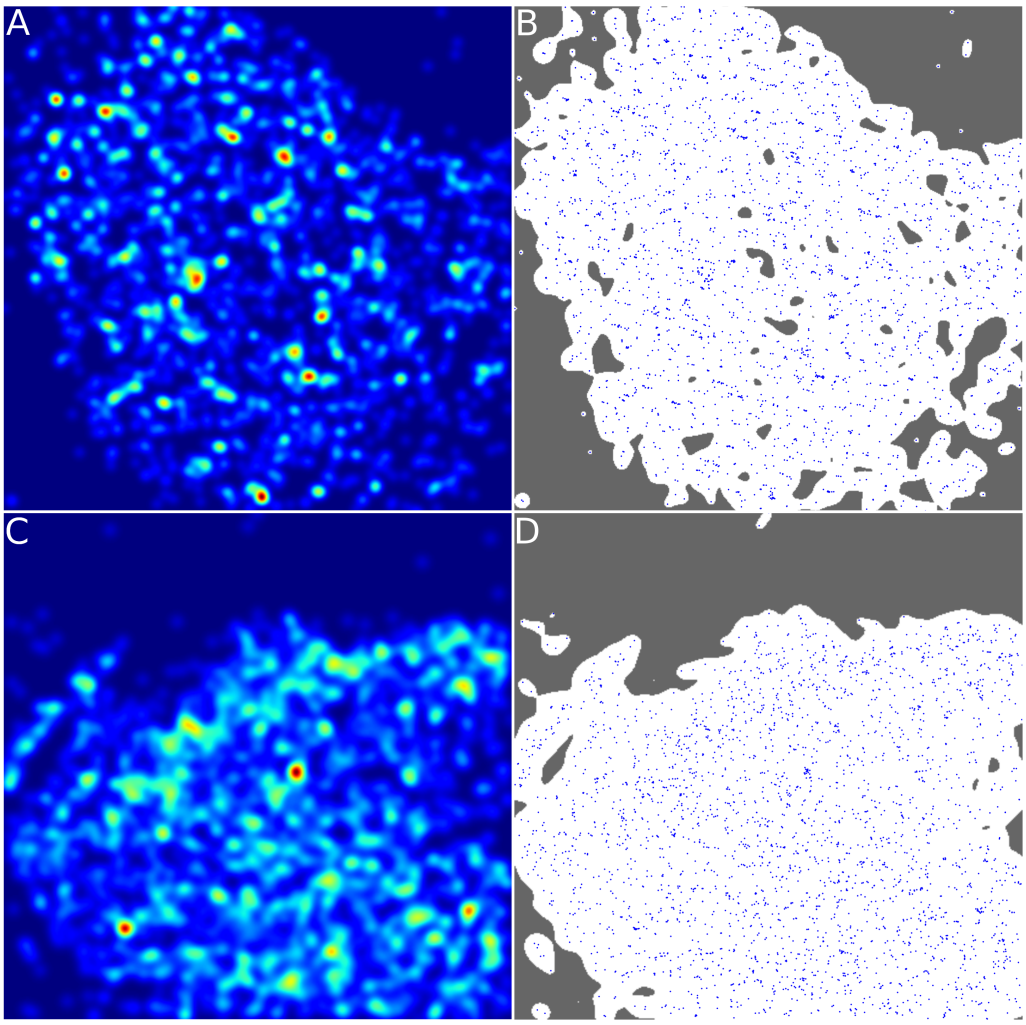

Figure 8.8: Localization Microscopy Images of Heterochromatin Markers. A. Shown is the density distribution of the localized markers in a cell prior to irradiation. Small, bright spots where markers are agglomerated can be seen. This means that heterochromatin is mainly organized in coarse clumps. **B.** Shown is the segmentated image of the not irradiated cell that is used for subsequent analysis of the marker distribution. **C.** The density distribution of a cell at $30\,min$ after irradiation with $0.5\,Gy$ is shown here. Differences between this cell and the not irradiated cell can be made out by visual inspection. We observe that the density has much less agglomerated and bright spots and is instead much more homogeneous. **D.** This effect can also be seen by visual inspection of the heterochromatin markers directly. Marker positions are visibly more spread out and less strongly clumped together. Heterochromatin clearly undergoes structural changes upon irradiation.

after irradiation (in panel **C.**) indicates an almost uniform and unconditioned increase in the probabilities of the short edges, just as in the case of the unconditioned edge length distribution.

Our results here show that heterochromatic domains undergo structural reorganizations after exposure to ionizing radiation. At $30\,min$ after irradiation the previously very compact and densely organized domains open up and adopt a more loose organization. After $48\,h$ the structures heal again and the organization approaches again the initial configuration of the untreated cells.

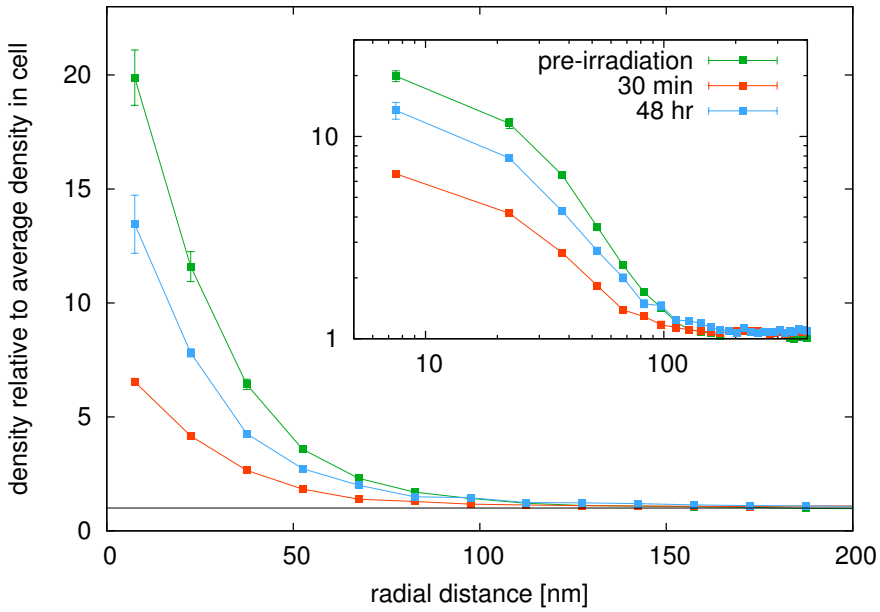

Figure 8.9: Results at different times after 0.5 Gy radiation dose. The figure shows the radial radial distribution function for methylated histone variants H4K20 antibodies representing heterochromatin for unirradiated and irradiated cells. Error bars represent the standard deviation of the mean value after averaging over the sample of cells. The value for g(r) at small distances goes up to around 20, indicating the high marker densities in regions where heterochromatin is located. The rapid drop off of the radial distribution function within a distance of less than $100\,nm$ shows that heterochromatin forms small clumps that are spread throughout the cell nucelus. Upon exposure to $0.5\,Gy$ γ-irradiation, a dramatic change in the correlation function can be observed in cells that were microscoped after $30\,min$. The value at small radial distances drops to around 6, or around 70 % smaller than in unirradiated cells. This indicates that the density in the heterochromatic regions is on average much lower in irradiated cells, requiring that the organisation of the chromatin fiber in these regions has to have loosened compared to before due to DNA damage such as double-strand breaks. In cells measured $48\,h$ after irradiation, the correlation function have recovered again and the value at small r is at around 14, only 30 % less than in unirradiated cells.

8.4 Conclusion

We analysed the effects of the reorganization of chromatin upon exposure to ionizing irradiation. Samples were imaged by Spectral Position Determination Microscopy (SPDM). The samples were subject to a preprocessing step in which regions of interest were detected.

A nearest neighbour graph was built by calculating the Delaunay triangulation of the localization points. The spatial organization of the fluorophores was characterized by different graph theoretic measures. Moreover, the radial distribution function was also calculated.

Comparing the quantities calculated for non-irradiated and irradiated samples, we found that although overall the chromatin may compactify to a certain degree, the local neighbouring properties of the localized points do not change. For instance, the distribution of the edge-lengths belong to the same family for irradiated and non irradiated samples. At the same time, we found that heterochromatic regions, marked separately, decondense upen irradiation. Furthermore, assuring a long enough healing time after irradiation, we observe a recovery of the heterochromatin from the decondensation. Our

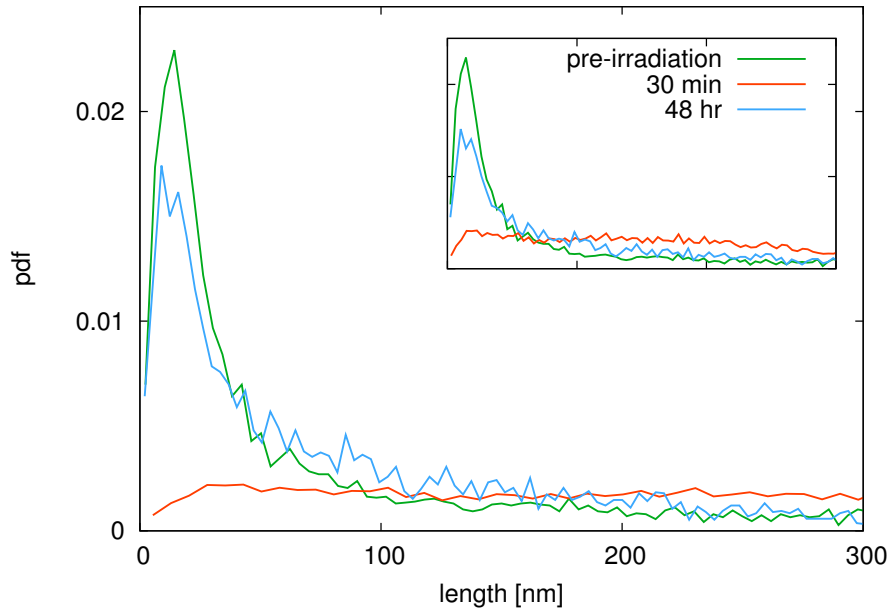

Figure 8.10: Results at different times after 0.5 Gy radiation dose. The distribution of edge lengths in the Delaunay triangulation of the markers confirms these observations. A sharp peak in the distribution at around 30 nm can be seen in untreated cells. In 30 min post-irradiation cells the peak vanishes and a spread distribution can be seen. In 48 h post-irradiation cells however, the peak reappears again but less pronounced than in untreated cells.

findings are in agreement with other experiments observing structural changes caused by the presence of the double strand breaks.

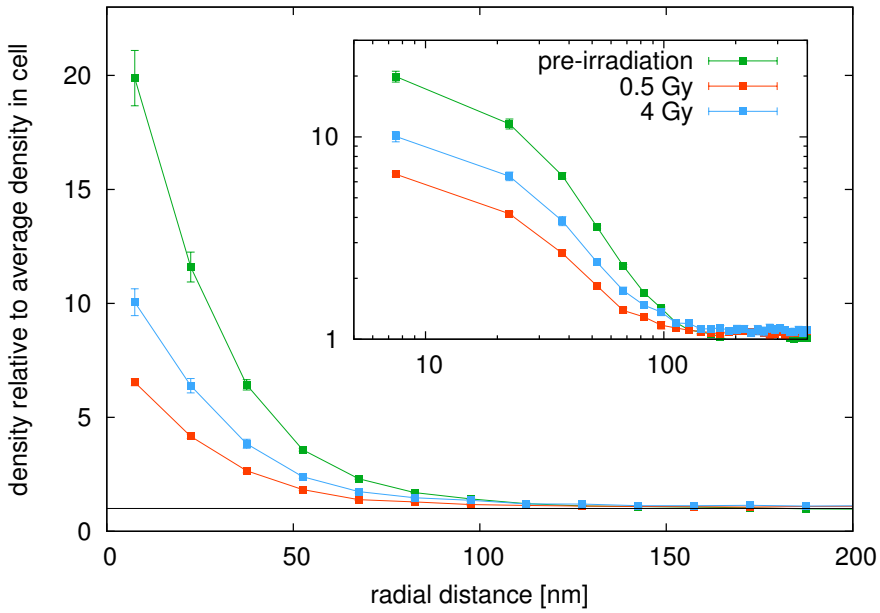

Figure 8.11: $g(r)$ **of heterochromatin markers for different doses** $30\,min$ **post irradiation.**

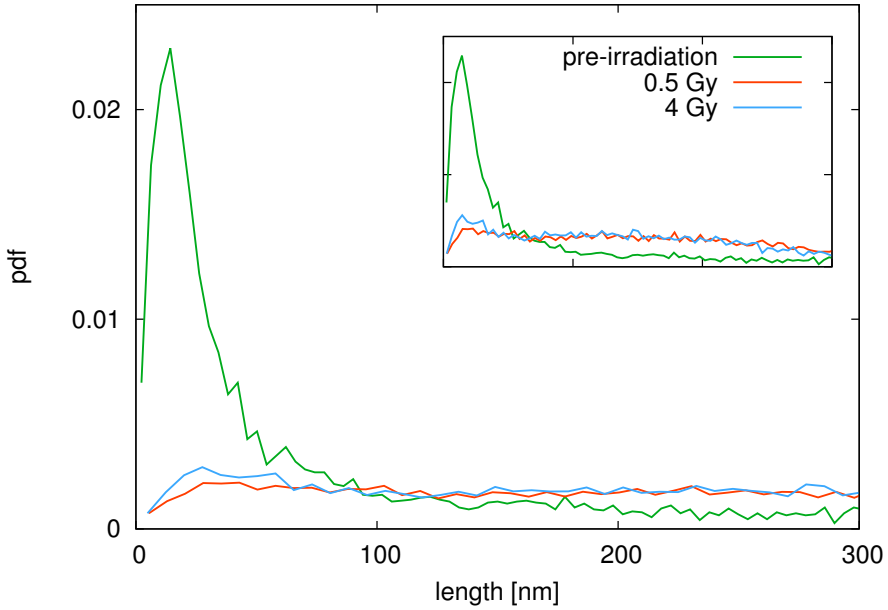

Figure 8.12: Edge length distribution of heterochromatin markers for different doses at $30\,min$ **post irradiation.**

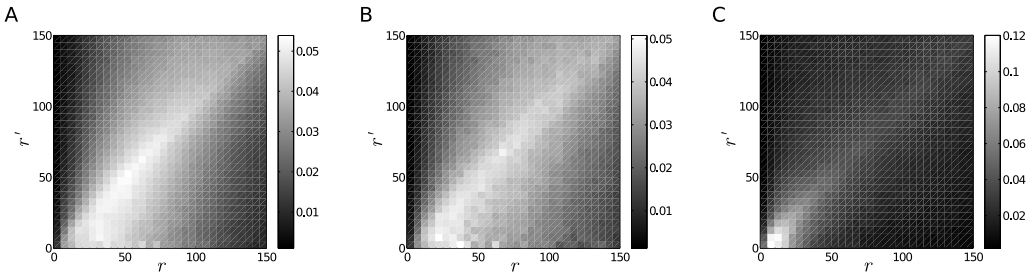

Figure 8.13: Conditional Probability Distribution of the Edge Lengts for Heterochromatin Markers. A. The panel shows the conditional probability distribution $p(r|r')$ calculated for the positions of the antibodies marking heterochromatic regions before irradiation. **B**. The panel shows the conditional probability for the heterochromatin markers $30\,min$ after irradiation. **C**. The panel shows the conditional probability for the heterochromatin markers $48\,h$ after irradiation. In all three cases the diagonal is very emphasised indicating preferential spatial distribution of the edges. This may stem from the clustering of the heterochromatin markers. Note that although shades along the diagonal are darker in panel **C**, the value of the corresponding probabilities are very close to the probabilities along the diagonal of panel **A** except for small radii.

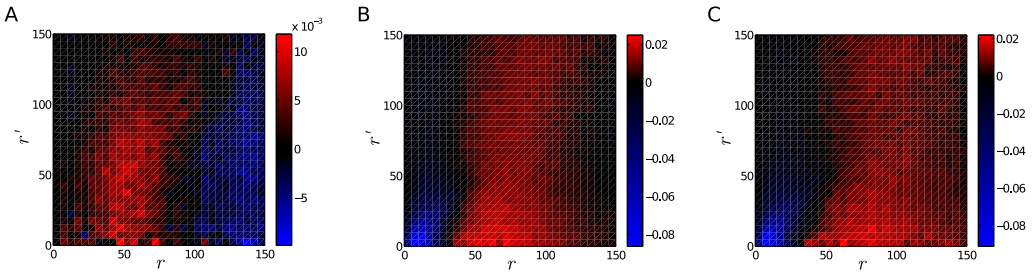

Figure 8.14: Difference in the Conditional Probability Distribution of the Edge Lengts for Heterochromatin Markers. A. The panel shows the difference in the conditional probability distribution $p(r|r')$ before irradiation and $30\,min$ after irradiation. A red shade of the colour-map means that the probability is higher before irradiation while a blue shade means that it is higher after irradiation. The panel indicates a slightly stronger change along the diagonal indicating a more homogeneous system after irradiation. However the change is almost independent of the value of the condition r'. **B**. The panel shows the difference of the conditional probability distribution calculated for samples before and $48\,h$ after irradiation. Shades of red indicate higher probabilities for the samples recorded before irradiation while shades of blue indicate higher probabilities in samples recorded after irradiation. **C**. The panel illustrates the difference of the conditional probability distribution $30\,min$ and $48\,h$ after irradiation. Red shades correspond to higher probabilities $30\,min$ post-irradiation while shades of blue indicate larger values $48\,h$ after irradiation. The difference between structures observed $30\,min$ after irradiation and $48\,h$ after irradiation indicate a reversed trend compared to panel **A**.

Chapter 9

Conclusion and Outlook

9.1 Short Summary of the Results

The primary target of this thesis was to examine different chromosomal systems by means of biophysical modeling. We aimed at developing models that could explain the behaviour of eukaryotic chromosomes under different conditions. We were able to show that physical principles are able to explain certain properties and characteristics of these different systems. The findings could contribute to a better understanding of mitotic chromosome organization and sister chromatid cohesion, the structural changes of chromatin posterior to exposure to irradiation and the behaviour of DNA double-strand breaks in chromosomal domains.

In chapter 4 we developed a model for the folding of the mitotic chromosome. Based on predictions from micromechanical experiments that mitotic chromatin could form a network we suggested size-restricted looping of the chromatin fiber as the major mechanism for chromatin condensation in mitosis. We represented the chromatin fiber with a polymer and incorporated chromatin loops with a dynamic and probabilistic mechanism. This binding mechanism implicitly model the presence of binding proteins such as condensin. One crucial aspect of our model is the restriction of loop sizes in our model chromosomes. The introduction of this quantities is based on the phenomenological observation of banding patterns and the requirement of length-wise condensation in mitotic chromosomes.

We evaluated our model by sampling model conformations with Metropolis Monte Carlo Simulations. Our results showed that our model can indeed result in a length-wise condensation of the chromatin fibre into rod-like objects. Furthermore, we were able to demonstrate that loops of the chromatin fiber can contribute in part to the bending rigidity of single chromatids as observed in Xenopus egg-extracted chromatids. We further analyzed the force-extension behaviour of single model chromatids. Our results for the force-extension curve were qualitatively consistent with experimental findings from micromechanical pulling experiments with an initial linear region followed by a force plateau. We found that the number and size of loops within the chromatin fiber governed the bending rigidity and elasticity of model chromosomes. We can conclude that dynamic looping

of the chromatin fiber is indeed a suitable model for the folding of mitotic chromosomes. Loops not only lead to a length-wise condensation of the fiber but contribute also to their mechanical properties through exertion of entropic forces.

We expanded our model for mitotic chromosomes to also describe sister chromatid cohesion in chapter 5. Cohesion between sister chromatids is of great importance for the correct distribution of them into the two developing daughter cells. Cohesin is the protein that is responsible for sister chromatid cohesion. It forms ring-like structures that could possibly clamp around sister chromatid fibers or tether them via two rings that form a handcuff. We incorporated cohesin activity implicitly in our model through the possibility of dynamic binding and unbinding of sister fibers.

We first explored the phase space of different binding patterns. We were able to show here that if binding of the fibers takes place on the scale of the underlying model fiber, then there must be a limited range for the average number of tethering points between sister fibers to obtain aligned and condensed chromatids. In particular our results demonstrated that a large average number of tethering points efficiently prohibits the condensation of each individual chromatid and leads to intermingled sister fibers. Since reinforced sister chromatid cohesion was reported for chromosomes that were exposed to ionizing radiation, we speculate that this effect could be a physical mechanism that contributes to cell cycle arrest at the entrance to mitosis as observed in irradiated cells. Our investigation of the mechanical properties of the tethered sister chromatid system showed that tether points between sisters increase the elasticity of the fiber which could be due to a dragging effect.

In chapter 6 we targeted the organization of chromatin after exposure to ionizing radiation and the behaviour of DNA double-strand breaks in chromosome sub-domains. We employed an expression-dependent chromatin folding model that couples the probability of intra-chromosomal formation of loops with the local transcriptional activity. For model data input we processed DNA microarray data from cells that were exposed to γ-radiation.

Using the according data we simulated the organization of chromosome 11 in untreated cells and in irradiated cells. The comparison of the simulation results showed that there are no significant global changes in the three-dimensional structure of chromosomes after exposure to radiation. This means that structural changes are limited to local chromosome domains. We therefore then modeled actual breaks of the fiber in different chromosomal areas and analyzed their behaviour. We showed that breakage of the fiber results in higher mobility and transport of the break site to the surface of the domain consistent with experimental observations.

In the third part of the thesis we showed that the overall organization of chromatin does not significantly change after irradiation but that entropic changes caused by particular strand break leads to an altered behaviour of the broken segment. We therefore concentrated on modeling chromosome sub-domains with different densities and induced breaks in the center of the domains.

We then evaluated how entropy could play a role in the signaling of the breaks to diffusing proteins. We first found, in agreement with our previous results that break sites have an increased mobility due to less entropic constraints. We then measured the time until a model repair protein attaches to an break end. The attachment times were generally higher in densely organized heterochromatin than in more open euchromatin sub-domains due to a crowding effect by the chromatin fiber itself. This may be an

explanation why heterochromatic domains often lack the typical formation of radiation-induced foci for DSBs and it demonstrates that the search process does not only depend on protein dynamics but also on the chromatin structure itself. Most significantly, our results show that the increased freedom of broken ends of the fiber facilitates the search process for them by diffusing proteins. We conclude that strand breaks can, in a way, be understood as a stable mechanism: the emergence of the disturbance to the system - the break - automatically, i.e. through entropic changes, leads to conditions that facilitates its recognition and eventual removal.

In the final project of the thesis we analyzed microscopy images of HeLa cell nuclei that were exposed to ionizing radiation. The images were obtained with the spectral precision distance microscopy technique and contains the localized, 2D-projected positions of marked histone proteins and specific antibodies. Using statistical physics methods and a graph theoretical approach we assessed the structural changes that chromatin undergoes after different doses of radiation and after time for repair.

Our results showed that the structure of histone positioning in the nucleus is not altered by irradiation. This indicates that the overall organization of the chromatin fiber does not change significantly after the exposure to ionizing radiation which is in agreement with our previous modeling results. However, our analysis of the positioning of antibodies indicating heterochromatic domains demonstrated that heterochromatic regions undergo a decondensation after irradiation. This possibly indicates that these normally highly compacted regions open up to become accessible for DSB repair proteins. Additionally, after long healing times, these domains recover from their structural changes and adopt configurations that are again similar to untreated cells.

9.2 Outlook

In this work we have presented a model for the structure and three-dimensional folding of mitotic chromosomes and the tethering of sister chromatids. While we explored how the number of loops and the number of tethering points can affect the mechanical properties of chromosomes in mitosis, the exact mechanisms of how these proteins act to fulfill their functions is not finally resolved yet.

The chromosome conformation capture technology has been extensively used to gain insight into the functional organization of interphase chromosomes in many different organisms [16, 18, 180, 259–262]. This exciting technology could also be employed to study the internal organization of mitotic chromosomes. Just like in interphase, HiC experiments could determine the interaction maps for different fiber segments in mitotic chromosomes. Fluorescence in-situ hybridization (FISH) experiments have also contributed to the understanding of interphase structure through the measurement of distances between labeled foci [183]. Since the compaction of mitotic chromosomes is much higher than in interphase, such measurements in mitotic chromosomes would require higher resolutions. Super-resolution methods such as the localization microscopy technique could be a possibility to carry out these kind of experiments. Their results would provide consistent experimental methods for the assessment of interphase and mitotic chromosome structure.

We are convinced that consistent experiments on the organization in interphase and mitosis are a prerequisite for the development of a unified folding model for chromatin in both stages of the cell cycle. HiC experiments have shown that loops of all sizes exist in interphase chromosomes. However, long-range interactions would lead to the collapse of

the fiber into a globular shaped object [263]. Our model for mitotic chromosomes shows that local folding of the chromatin fiber leads to length-wise condensation into rigid and rod-like objects. Therefore, when entering mitosis, long-range loops that possibly have regulatory functions in interphase are lost [33]. At the same time, the abundance of small loops must increase to achieve the tight compaction of mitotic chromosomes. On the other hand, the question of how epigenetic information can be preserved through mitosis is of great significance. Epigenetic information is not stored in the DNA sequence. Besides different kinds of modifications to DNA and associated proteins, chromatin loops could provide a means for the preservation of epigenetic states [264]. Therefore, the transition from interphase to mitosis with the changes in the internal loop structure is a fascinating topic for further research.

Furthermore, a very interesting question in the context of sister chromatid cohesion is the process of chromatid separation after anaphase onset. In anaphase, cohesin is mostly cleaved from sister chromatids allowing the mitotic spindle fibers attached to the kinetochore at the centromere region of the chromosomes to transport sister chromatids to the different daughter cells [265]. The separation dynamics of two sister chromatids that are aligned and maybe even intertwined with each other is an very interesting question. Here, pulling forces through the mitotic spindle could act together with entropic forces and drive the segregation. The segregation of two intertwined long polymers has been analyzed in a recent study [36]. How loops that condense each single chromatid could contribute to the separation dynamics is a highly interesting question that could be targeted in future studies.

We have assessed how the organization of chromatin changes in cells that were exposed to ionizing radiation through biophysical modeling and also by image analysis. Although we showed that global changes of the organization is unlikely, there is abundant evidence, that local changes of chromatin structure in the vicinity of DSBs is very common, for example through chromatin remodelers and the formation of γ-H2AX [45, 49, 135, 266]. The more detailed study of such local changes to chromatin structure would therefore be of great interest. Experiments, for example high-resolution fluorescence microscopy, that are able to measure properties of chromatin in the vicinity of predefined sites of DSBs could give a great contribution. Especially, the measurement of physical distances between known loci would be of great benefit. In combination with polymer models that are able to predict the three-dimensional organization, they could reveal structures and structural changes that are specific for DSBs. Such specific structures could moreover then serve to analyze the process of protein recruitment to the site of DSBs in greater detail.

Our study on the recognition of DSBs does take into account different chromatin architectures. We have assessed the search time for breaks that lie within compact heterochromatic areas and open euchromatic areas. Our results demonstrate that the crowding effect by the chromatin fiber itself has a profound influence on the search time. Dense chromatin regions provide less accessibility to diffusing proteins which is also reflected in slowed recruitment kinetics observed in experiments [210]. However, crowding is not only established through the chromatin architecture, but also through other proteins [245]. The inclusion of binding proteins other than repair proteins in the model could be a way to assess how this influences the search process for DSBs.

One major obstacle for the direct experimental evaluation of our results is the lack of techniques to determine the physical location of DSBs prior to the arrival of early repair proteins. Experiments mainly visualize DSB foci by labeling proteins and damage-

associated proteins after the recognition of the DSB by the cell. Experiments in which the chromatin fiber is cut by nucleases could provide a means to label the specific genomic positions at which the DSBs will occur and time-resolved measurements of the dynamics would be highly beneficial to validate our model results. On the other hand, additionally to the estimation of the search time for early damage recognition proteins, it would be an interesting option to target the recruitment dynamics of subsequent repair proteins by means of biophysical modeling. In combination with the assessment of different chromatin architectures, this could be a direction of future studies that aims at explaining different repair kinetics for chromosome sub-domains.

The analysis of high-resolution images of HeLa cells verified our modeling results that there are no global changes in the organization of interphase chromosomes after exposure to ionizing radiation. The images provided information on the positions of histones H2B and at the same time of antibodies that are specific for heterochromatic regions. This allowed us to show the decondensation of heterochromatin after irradiation. Antibodies for other chromatin regions could be of great interests for the study of damage-induced local structural changes, too. For example, γ-H2AX is commonly accepted as an indicator for DSBs. High-resolution images that can determine the position of antibodies specific to γ-H2AX would make it possible to analyze the detailed structure of chromatin in the surrounding of DSBs.

In conclusion, the understanding of chromosome structure throughout the cell cycle and the precise comprehension of the damage response are big challenges. The work in this thesis was aiming at contributing to this goal. Our approach with biophysical modeling was able to show that physical principles play an essential role in chromatin organization in mitosis and in the damage recognition and response to DNA double-strand breaks. With results from advanced experiments, more detailed and specific models for chromosome domains could be designed to better understand the intricate cellular processes to resolve radiation induced DSBs and contribute to the comprehension of the relationship between these microscopic mechanisms and macroscopic effect of radiation in living tissue. This could support the development of methods in radiotherapy and in diagnostics that utilize ionizing radiation.

Conference/Workshop Participation

I have participated in the following conferences and workshops:

- 4th Workshop on Monte Carlo Methods, January 15, 2010, Heidelberg
- Frühjahrstagung der Deutschen Physikalischen Gesellschaft, March 13 – 18, 2011, Dresden [poster]
- 5th Workshop on Monte Carlo Methods, January 14, 2011, Heidelberg
- Summer School on Probabilistic Graphical Models, September 4 – 16, 2011, Shanghai [poster]
- Topical Workshop on Modelling of Chromosomes, September 9, 2011, Shanghai [talk]
- Doktorandenforum Natur der Studienstiftung des Deutschen Volkes, November 8 – 11, 2012, Potsdam
- Annual Colloquium of the Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences, November 20 – 21, 2012, Erbach im Odenwald [poster]
- Frühjahrstagung der Deutschen Physikalischen Gesellschaft, March 10 – 15, 2013, Regensburg [talk]
- Frühjahrstagung der Deutschen Physikalischen Gesellschaft, March 10 – 15, 2013, Regensburg [2 posters]
- March Meeting of the American Physical Society, March 18 – 22, 2013, Baltimore [talk]
- 6th Workshop on Monte Carlo Methods, April 12 – 13, 2013, Heidelberg [poster]

Bibliography

[1] Daniele Zink, Andrew H Fischer, and Jeffrey A Nickerson. Nuclear structure in cancer cells. *Nat Rev Cancer*, 4(9):677–687, September 2004.

[2] Shihua He, Katherine L Dunn, Paula S Espino, Bojan Drobic, Lin Li, Jenny Yu, Jian-Min Sun, Hou Yu Chen, Susan Pritchard, and James R Davie. Chromatin organization and nuclear microenvironments in cancer cells. *J Cell Biochem*, 104(6):2004–2015, August 2008.

[3] Janice L Farlow and Tatiana Foroud. The genetics of dementia. *Seminars in neurology*, 33(4):417–422, September 2013.

[4] The human genome project. http://web.ornl.gov/sci/techresources/Human_Genome/.

[5] Race to cut whole genome sequencing costs. http://www.genengnews.com/gen-articles/race-to-cut-whole-genome-sequencing-costs/939/. accessed: 2013-12-02.

[6] Lauren Gravitz. TR10: $100 genome. http://www2.technologyreview.com/article/412186/tr10-100-genome/, 2009. accessed: 2013-12-02.

[7] Katherine Harmon. Genome sequencing for the rest of us: Scientific american. http://www.scientificamerican.com/article.cfm?id=personal-genome-sequencing, 2010. accessed: 2013-12-02.

[8] Kevin Struhl and Eran Segal. Determinants of nucleosome positioning. *Nature structural & molecular biology*, 20(3):267–273, March 2013.

[9] J Bednar, R A Horowitz, S A Grigoryev, L M Carruthers, J C Hansen, A J Koster, and C L Woodcock. Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proceedings of the National Academy of Sciences of the United States of America*, 95(24):14173–14178, November 1998.

[10] Philipp M Diesinger and Dieter W Heermann. Two-angle model and phase diagram for chromatin. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 74(3 Pt 1):031904, September 2006.

[11] Philip J J Robinson and Daniela Rhodes. Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Current opinion in structural biology*, 16(3):336–343, June 2006.

[12] D.J. Tremethick. Higher-order structures of chromatin: The elusive 30 nm fiber. *Cell*, 128(4):651–654, 2007.

[13] K Maeshima, S Hihara, and H Takata. New insight into the mitotic chromosome structure: irregular folding of nucleosome fibers without 30-nm chromatin structure. *Cold Spring Harbor symposia on quantitative biology*, 75:439–444, 2010.

[14] Job Dekker. Gene regulation in the third dimension. *Science (New York, N.Y.)*, 319(5871):1793–1794, March 2008.

[15] Stephan Kadauke and Gerd A Blobel. Chromatin loops in gene regulation. *Biochimica et biophysica acta*, 1789(1):17–25, January 2009.

[16] Kuljeet Singh Sandhu, Guoliang Li, Wing-Kin Sung, and Yijun Ruan. Chromatin interaction networks and higher order architectures of eukaryotic genomes. *Journal of Cellular Biochemistry*, 112(9):2218–2221, September 2011.

[17] Job Dekker, Marc A Marti-Renom, and Leonid A Mirny. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nature reviews. Genetics*, 14(6):390–403, June 2013.

[18] Alejandro Rodriguez and Pernilla Bjerling. The links between chromatin spatial organization and biological function. *Biochemical Society transactions*, 41(6):1634–1639, December 2013.

[19] Elena Fedorova and Daniele Zink. Nuclear genome organization: common themes and individual patterns. *Current opinion in genetics & development*, 19(2):166–171, April 2009.

[20] Bas van Steensel and Job Dekker. Genomics tools for unraveling chromosome architecture. *Nature biotechnology*, 28(10):1089–1095, October 2010.

[21] Wendy A Bickmore and Bas van Steensel. Genome architecture: domain organization of interphase chromosomes. *Cell*, 152(6):1270–1284, March 2013.

[22] Neidhard Paweletz. Walther flemming: pioneer of mitosis research. *Nature Reviews Molecular Cell Biology*, 2(1):72–75, January 2001.

[23] A. S. Belmont. Mitotic chromosome structure and condensation. *Current Opinion in Cell Biology*, 18(6):632 – 638, 2006. Cell division, growth and death / Cell differentiation.

[24] Raquel A Oliveira and Kim Nasmyth. Getting through anaphase: splitting the sisters and beyond. *Biochemical Society Transactions*, 38(6):1639–1644, December 2010.

[25] Kum K. Khanna and Stephen P. Jackson. DNA double-strand breaks: signaling, repair and the cancer connection. *Nature Genetics*, 27(3):247–254, 2001.

[26] Kai Rothkamm and Markus Löbrich. Misrepair of radiation-induced DNA double-strand breaks and its relevance for tumorigenesis and cancer treatment (review). *International journal of oncology*, 21(2):433–440, August 2002.

[27] Kevin D Mills, David O Ferguson, and Frederick W Alt. The role of DNA breaks in genomic instability and tumorigenesis. *Immunol Rev*, 194:77–95, August 2003.

[28] Suckjoon Jun and Bela Mulder. Entropy-driven spatial organization of highly confined polymers: Lessons for the bacterial chromosome. *Proceedings of the National Academy of Sciences*, 103(33):12388–12393, 2006.

[29] Davide Marenduzzo, Cristian Micheletti, and Peter R Cook. Entropy-driven genome organization. *Biophysical journal*, 90(10):3712–3721, May 2006.

[30] Mario Nicodemi and Antonella Prisco. Thermodynamic pathways to genome spatial organization in the cell nucleus. *Biophysical journal*, 96(6):2168–2177, March 2009.

[31] P. R. Cook and D. Marenduzzo. Entropic organization of interphase chromosomes. *J Cell Biol*, 186(6):825–834, September 2009.

[32] Kieran Finan, Peter R Cook, and Davide Marenduzzo. Non-specific (entropic) forces as major determinants of the structure of mammalian chromosomes. *Chromosome research: an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology*, 19(1):53–61, January 2011.

[33] Dieter W Heermann. Physical nuclear organization: loops and entropy. *Current opinion in cell biology*, 23(3):332–337, June 2011.

[34] Martin A White, John K Eykelenboom, Manuel A Lopez-Vernaza, Emily Wilson, and David R F Leach. Non-random segregation of sister chromosomes in escherichia coli. *Nature*, 455(7217):1248–1250, October 2008.

[35] Suckjoon Jun and Andrew Wright. Entropy as the driver of chromosome segregation. *Nature reviews. Microbiology*, 8(8):600–607, August 2010.

[36] Ron Dockhorn and Jens-Uwe Sommer. A model for segregation of chromatin after replication: segregation of identical flexible chains in solution. *Biophysical journal*, 100(11):2539–2547, June 2011.

[37] Manfred Bohn and Dieter W Heermann. Repulsive forces between looping chromosomes induce entropy-driven segregation. *PloS one*, 6(1):e14428, 2011.

[38] Barbara Di Ventura, BenoÃőt Knecht, Helena Andreas, William J Godinez, Miriam Fritsche, Karl Rohr, Walter Nickel, Dieter W Heermann, and Victor Sourjik. Chromosome segregation by the escherichia coli min system. *Molecular systems biology*, 9:686, 2013.

[39] Pui-Man Lam and Richard M. Neumann. Role of chain entropy in an analytic model of protein binding in single-DNA stretching experiments. *Physical Review E*, 84(3):032901, September 2011.

[40] Christoph J Feinauer, Andreas Hofmann, Sebastian Goldt, Lei Liu, Gabriell MÃąté, and Dieter W Heermann. Zinc finger proteins and the 3D organization of chromosomes. *Advances in protein chemistry and structural biology*, 90:67–117, 2013.

[41] Mariliis Tark-Dame, Roel van Driel, and Dieter W Heermann. Chromatin folding–from biology to polymer models and back. *Journal of cell science*, 124(Pt 6):839–845, March 2011.

[42] Geoffrey Fudenberg and Leonid A. Mirny. Higher order chromatin structure: bridging physics and biology. *Current opinion in genetics & development*, 22(2):115–124, April 2012.

[43] J.F. Marko. Micromechanical studies of mitotic chromosomes. *Chromosome Research*, 16(3):469–497, 2008.

[44] Michael R Lieber and Zarir E Karanjawala. Ageing, repetitive genomes and DNA damage. *Nature reviews. Molecular cell biology*, 5(1):69–75, January 2004.

[45] T.K. Pandita and C. Richardson. Chromatin remodeling finds its place in the dna double-strand break response. *Nucleic Acids Res*, 37(5):1363–1377, Apr 2009.

[46] Stephen P. Jackson and Jiri Bartek. The DNA-damage response in human biology and disease. *Nature*, 461(7267):1071–1078, October 2009.

[47] Lumir Krejci, Veronika Altmannova, Mario Spirek, and Xiaolan Zhao. Homologous recombination and its regulation. *Nucleic Acids Research*, 40(13):5795–5818, 2012.

[48] T. Helleday, J. Lo, D.C. van Gent, and B.P. Engelward. Dna double-strand break repair: From mechanistic understanding to cancer treatment. *DNA Repair*, 6(7):923–935, 2007. Replication Fork Repair Processes.

[49] Andrea Kinner, Wenqi Wu, Christian Staudt, and George Iliakis. Gamma-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin. *Nucleic Acids Res*, 36(17):5678–5694, October 2008.

[50] Sally A Amundson, Khanh T Do, Lisa C Vinikoor, R Anthony Lee, Christine A Koch-Paiz, Jaeyong Ahn, Mark Reimers, Yidong Chen, Dominic A Scudiero, John N Weinstein, Jeffrey M Trent, Michael L Bittner, Paul S Meltzer, and Jr Fornace, Albert J. Integrating global gene expression and radiation survival parameters across the 60 cell lines of the national cancer institute anticancer drug screen. *Cancer research*, 68(2):415–424, January 2008.

[51] Denis A Smirnov, Michael Morley, Eunice Shin, Richard S Spielman, and Vivian G Cheung. Genetic analysis of radiation-induced changes in human gene expression. *Nature*, 459(7246):587–591, May 2009.

[52] Jr Koshland, Daniel E. Special essay. the seven pillars of life. *Science (New York, N.Y.)*, 295(5563):2215–2216, March 2002. PMID: 11910092.

[53] Chris P McKay. What is life–and how do we search for it in other worlds? *PLoS biology*, 2(9):E302, September 2004. PMID: 15367939.

[54] Bruce Alberts. *Molecular biology of the cell*. Garland, New York, 4. ed. edition, 2002.

[55] Harvey F. Lodish. *Molecular cell biology*. Freeman, New York, NY, 6. ed. edition, 2008. Glossary S. G-1 - G-24; Index I-1 - I-52.

[56] Gerald Karp. *Molekulare Zellbiologie*. Springer-Lehrbuch. Springer, Berlin ; Heidelberg [u.a.], 1. dt. aufl. edition, 2005.

[57] G. Poeggel. *Kurzlehrbuch Biologie*. Thieme, Stuttgart, 2. überarb. aufl. edition, 2009.

[58] D.P. Clark. *Molecular Biology*. Easy Reading. Elsevier, Spektrum Akademischer Verlag, München; Heidelberg, 1 edition, 2006.

[59] G. Mendel. Versuche über pflanzenhybriden. http://www.mendelweb.org/MWGerText.html, February 1866.

[60] Wikipedia, the free encyclopedia. http://www.wikipedia.org.

[61] Benjamin Lewin. *Genes*. Pearson Prentice Hall, Upper Saddle River, NJ, [8. ed.], internat. ed. edition, 2004.

[62] Sandra C Moser and Jason R Swedlow. How to be a mitotic chromosome. *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology*, 19(3):307–319, April 2011.

[63] R. Dahm. Discovering DNA: friedrich miescher and the early years of nucleic acid research. *Human Genetics*, 122(6):565–581, 2008. 10.1007/s00439-007-0433-0.

[64] James D. Watson and Francis H.C. Crick. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171:737–738, 1953.

[65] F.H.C. Crick and J.D. Watson. The complementary structure of deoxyribonucleic acid. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 223(1152):80–96, 1954.

[66] J B Lawrence, R H Singer, and J A McNeil. Interphase and metaphase resolution of different distances within the human dystrophin gene. *Science (New York, N.Y.)*, 249(4971):928–932, August 1990.

[67] G Li, G Sudlow, and A S Belmont. Interphase cell cycle dynamics of a late-replicating, heterochromatic homogeneously staining region: precise choreography of condensation/decondensation and nuclear positioning. *The Journal of cell biology*, 140(5):975–989, March 1998.

[68] Andrew S Belmont. Mitotic chromosome scaffold structure: New approaches to an old controversy. *Proceedings of the National Academy of Sciences of the United States of America*, 99(25):15855–15857, 2002.

[69] Dieter W Heermann. Mitotic chromosome structure. *Experimental cell research*, 318(12):1381–1385, July 2012.

[70] J.T. Finch and A. Klug. Solenoidal model for superstructure in chromatin. *Proceedings of the National Academy of Sciences of the United States of America*, 73(6):1897–1901, 1976.

[71] P Suau, E M Bradbury, and J P Baldwin. Higher-order structures of chromatin in solution. *European journal of biochemistry / FEBS*, 97(2):593–602, July 1979.

[72] K. van Holde and J. Zlatanova. What determines the folding of the chromatin fiber? *Proceedings of the National Academy of Sciences of the United States of America*, 93(20):10548–10555, 1996.

[73] T. Cremer and C. Cremer. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature Review Genetics*, 2(4):292–301, 2001.

[74] Thomas Cremer and Marion Cremer. Chromosome territories. *Cold Spring Harbor Perspectives in Biology*, 2(3):a003889, March 2010.

[75] Ernest S Kawasaki. The end of the microarray tower of babel: will universal standards lead the way? *Journal of biomolecular techniques: JBT*, 17(3):200–206, July 2006.

[76] G. S. Stein, M. Montecino, A. J. van Wijnen, J. L. Stein, and J. B. Lian. Nuclear structure-gene expression interrelationships: implications for aberrant gene expression in cancer. *Cancer Res*, 60(8):2067–2076, April 2000.

[77] Uma T Shankavaram, William C Reinhold, Satoshi Nishizuka, Sylvia Major, Daisaku Morita, Krishna K Chary, Mark A Reimers, Uwe Scherf, Ari Kahn, Douglas Dolginow, Jeffrey Cossman, Eric P Kaldjian, Dominic A Scudiero, Emanuel Petricoin, Lance Liotta, Jae K Lee, and John N Weinstein. Transcript and protein expression profiles of the NCI-60 cancer cell panel: an integromic microarray study. *Molecular cancer therapeutics*, 6(3):820–832, March 2007.

[78] Michael J Overman, Jiexin Zhang, Scott Kopetz, Michael Davies, Jiang Zhi-Qin, Katherine Stemke-Hale, Petra Rümmele, Christian Pilarsky, Robert Grützmann, Stanley Hamilton, Rosa Hwang, James L Abbruzzese, Gauri Varadhachary, Bradley Broom, and Huamin Wang. Gene expression profiling of ampullary carcinomas classifies ampullary carcinomas into biliary-like and intestinal-like subtypes that are prognostic of outcome. *PloS one*, 8(6):e65144, 2013.

[79] Hongfang Liu, Petula D'Andrade, Stephanie Fulmer-Smentek, Philip Lorenzi, Kurt W Kohn, John N Weinstein, Yves Pommier, and William C Reinhold. mRNA and microRNA expression profiles of the NCI-60 integrated with drug activities. *Molecular cancer therapeutics*, 9(5):1080–1091, May 2010.

[80] U Maskos and E M Southern. Oligonucleotide hybridizations on glass supports: a novel linker for oligonucleotide synthesis and hybridization properties of oligonucleotides synthesised in situ. *Nucleic acids research*, 20(7):1679–1684, April 1992.

[81] Yuk Fai Leung and Duccio Cavalieri. Fundamentals of cDNA microarray data analysis. *Trends in genetics: TIG*, 19(11):649–659, November 2003.

[82] Adam G. West and Peter Fraser. Remote control of gene transcription. *Human Molecular Genetics*, 14(suppl 1):R101–R111, April 2005.

[83] Stavros Lomvardas, Gilad Barnea, David J Pisapia, Monica Mendelsohn, Jennifer Kirkland, and Richard Axel. Interchromosomal interactions and olfactory receptor choice. *Cell*, 126(2):403–413, July 2006.

[84] David Carter, Lyubomira Chakalova, Cameron S Osborne, Yan-feng Dai, and Peter Fraser. Long-range chromatin regulatory interactions in vivo. *Nature genetics*, 32(4):623–626, December 2002.

[85] Dirk A. Kleinjan and Veronica van Heyningen. Long-range control of gene expression: Emerging mechanisms and disruption in disease. *The American Journal of Human Genetics*, 76(1):8–32, January 2005.

[86] Stefan Schoenfelder, Ieuan Clay, and Peter Fraser. The transcriptional interactome: gene expression in 3D. *Current opinion in genetics & development*, 20(2):127–133, April 2010.

[87] Stefan Schoenfelder, Tom Sexton, Lyubomira Chakalova, Nathan F Cope, Alice Horton, Simon Andrews, Sreenivasulu Kurukuti, Jennifer A Mitchell, David Umlauf, Daniela S Dimitrova, Christopher H Eskiw, Yanquan Luo, Chia-Lin Wei, Yijun Ruan, James J Bieker, and Peter Fraser. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nature genetics*, 42(1):53–61, January 2010.

[88] Cameron S Osborne, Lyubomira Chakalova, Karen E Brown, David Carter, Alice Horton, Emmanuel Debrand, Beatriz Goyenechea, Jennifer A Mitchell, Susana Lopes, Wolf Reik, and Peter Fraser. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nature genetics*, 36(10):1065–1071, October 2004.

[89] Jennifer A Mitchell and Peter Fraser. Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes & development*, 22(1):20–25, January 2008.

[90] Peter R. Cook. A model for all genomes: The role of transcription factories. *Journal of Molecular Biology*, 395(1):1–10, 2010.

[91] Sarah C R Elgin and Shiv I S Grewal. Heterochromatin: silence is golden. *Current biology: CB*, 13(23):R895–898, December 2003.

[92] Danielle Vermaak, Kami Ahmad, and Steven Henikoff. Maintenance of chromatin states: an open-and-shut case. *Current Opinion in Cell Biology*, 15(3):266–274, June 2003.

[93] C Münkel, R Eils, S Dietzel, D Zink, C Mehring, G Wedemann, T Cremer, and J Langowski. Compartmentalization of interphase chromosomes observed in simulation and experiment. *Journal of molecular biology*, 285(3):1053–1065, January 1999.

[94] Guillaume J Filion, Joke G van Bemmel, Ulrich Braunschweig, Wendy Talhout, Jop Kind, Lucas D Ward, Wim Brugman, InÃłs J de Castro, Ron M Kerkhoven, Harmen J Bussemaker, and Bas van Steensel. Systematic protein location mapping reveals five principal chromatin types in drosophila cells. *Cell*, 143(2):212–224, October 2010.

[95] C. L. Woodcock and S. Dimitrov. Higher-order structure of chromatin and chromosomes. *Current Opinion in Genetics & Development*, 11(2):130 – 135, 2001.

[96] S. Almagro, D. Riveline, T. Hirano, B. Houchmandzadeh, and S. Dimitrov. The mitotic chromosome is an assembly of rigid elastic axes organized by structural maintenance of chromosomes (SMC) proteins and surrounded by a soft chromatin envelope. *Journal of Biological Chemistry*, 279(7):5118–5126, 2004.

[97] J.R. Paulson and U.K. Laemmli. The structure of histone-depleted metaphase chromosomes. *Cell*, 12(3):817–828, 1977.

[98] M. P. F. Marsden and U. K. Laemmli. Metaphase chromosome structure: Evidence for a radial loop model. *Cell*, 17(4):849 – 858, 1979.

[99] W.C. Earnshaw and U.K. Laemmli. Architecture of metaphase chromosomes and chromosome scaffolds. *The Journal of Cell Biology*, 96(1):84–93, 1983.

[100] W.C. Earnshaw, B. Halligan, C.A. Cooke, M.M. Heck, and L.F. Liu. Topoisomerase II is a structural component of mitotic chromosome scaffolds. *The Journal of Cell Biology*, 100(5):1706–1715, 1985.

[101] Y. Hirano and T.J. Mitchison. A heterodimeric coiled-coil protein required for mitotic chromosome condensation in vitro. *Cell*, 79(3):449 – 458, 1994.

[102] T. Hirano and T.J. Mitchison. Topoisomerase II does not play a scaffolding role in the organization of mitotic chromosomes assembled in xenopus egg extracts. *The Journal of Cell Biology*, 120(3):601–612, 1993.

[103] T. Hirano, R. Kobayashi, and M. Hirano. Condensins, chromosome condensation protein complexes containing XCAP-C, XCAP-E and a xenopus homolog of the drosophila barren protein. *Cell*, 89(4):511 – 521, 1997.

[104] T. Hirano. Condensins: Organizing and segregating the genome. *Current Biology*, 15(7):R265 – R275, 2005.

[105] Christine Michaelis, Rafal Ciosk, and Kim Nasmyth. Cohesins: Chromosomal proteins that prevent premature separation of sister chromatids. *Cell*, 91(1):35–45, 1997.

[106] Christian H Haering, Ana-Maria Farcas, Prakash Arumugam, Jean Metson, and Kim Nasmyth. The cohesin ring concatenates sister DNA molecules. *Nature*, 454(7202):297–301, July 2008.

[107] Nenggang Zhang, Sergey G Kuznetsov, Shyam K Sharan, Kaiyi Li, Pulivarthi H Rao, and Debananda Pati. A handcuff model for the cohesin complex. *The Journal of Cell Biology*, 183(6):1019–1031, December 2008.

[108] Matthew J Renshaw, Jonathan J Ward, Masato Kanemaki, Kayo Natsume, Fran\textbackslashtextbackslashtextbackslashccois J Nédélec, and Tomoyuki U Tanaka. Condensins promote chromosome recoiling during early anaphase to complete sister chromatid separation. *Developmental Cell*, 19(2):232–244, August 2010.

[109] Sara Cuylen and Christian H Haering. Deciphering condensin action during chromosome segregation. *Trends in cell biology*, 21(9):552–559, September 2011.

[110] U. Claussen, A. Mazur, and N. Rubtsov. Chromosomes are highly elastic and can be stretched. *Cytogenetics and Cell Genetics*, 66(2):120–125, 1994.

[111] B. Houchmandzadeh, J.F. Marko, D. Chatenay, and A. Libchaber. Elasticity and structure of eukaryote chromosomes studied by micromanipulation and micropipette aspiration. *The Journal of Cell Biology*, 139(1):1–12, 1997.

[112] B. Houchmandzadeh and S. Dimitrov. Elasticity measurements show the existence of thin rigid cores inside mitotic chromosomes. *The Journal of Cell Biology*, 145(2):215–223, 1999.

[113] M. Poirier, S. Eroglu, D. Chatenay, and J.F. Marko. Reversible and irreversible unfolding of mitotic newt chromosomes by applied force. *Mol. Biol. Cell*, 11(1):269–276, 2000.

[114] W. F. Marshall, J. F. Marko, D. A. Agard, and J. W. Sedat. Chromosome elasticity and mitotic polar ejection force measured in living drosophila embryos by four-dimensional microscopy-based motion analysis. *Current Biology*, 11(8):569 – 578, 2001.

[115] M.G. Poirier, S. Eroglu, and J.F. Marko. The bending rigidity of mitotic chromosomes. *Mol. Biol. Cell*, 13(6):2170–2179, 2002.

[116] M.G. Poirier, A. Nemani, P. Gupta, S. Eroglu, and J.F. Marko. Probing chromosome structure with dynamic force relaxation. *Phys. Rev. Lett.*, 86(2):360–363, January 2001.

[117] M.G. Poirier and J.F. Marko. Mitotic chromosomes are chromatin networks without a mechanically contiguous protein scaffold. *Proceedings of the National Academy of Sciences of the United States of America*, 99(24):15393–15397, 2002.

[118] Mingxuan Sun, Ryo Kawamura, and John F Marko. Micromechanics of human mitotic chromosomes. *Physical Biology*, 8(1):015003, February 2011.

[119] Michael R. Lieber. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual Review of Biochemistry*, 79(1):181–211, 2010.

[120] Noritaka Adachi, Hiromi Suzuki, Susumu Iiizumi, and Hideki Koyama. Hypersensitivity of nonhomologous DNA end-joining mutants to VP-16 and ICRF-193: implications for the repair of topoisomerase II-mediated DNA damage. *The Journal of biological chemistry*, 278(38):35897–35902, September 2003.

[121] Brandon J Lamarche, Nicole I Orazio, and Matthew D Weitzman. The MRN complex in double-strand break repair and telomere maintenance. *FEBS Lett*, 584(17):3682–3695, September 2010.

[122] Agnieszka Rupnik, Noel F Lowndes, and Muriel Grenon. MRN and the race to the break. *Chromosoma*, 119(2):115–135, April 2010.

[123] Benjamin E. Nelms, Richard S. Maser, James F. MacKay, Max G. Lagally, and John H. J. Petrini. In situ visualization of DNA double-strand break repair in human fibroblast. *Science*, 280(5363):590–592, 1998.

[124] Olga K. Mirzoeva and John H. Petrini. DNA damage-dependent nuclear dynamics of the mre11 complex. *Molecular and Cellular Biology*, 21(1):281–288, January 2001.

[125] C. Lukas, J. Falck, J. Bartkova, J. Bartek, and Jiri Lukas. Distinct spatiotemporal dynamics of mammalian checkpoint regulators induced by DNA damage. *Nat Cell Biol*, 5(3):255–260, March 2003.

[126] S. V. Costes, I. Chiolo, J. M. Pluth, M. H. Barcellos-Hoff, and B. Jakob. Spatiotemporal characterization of ionizing radiation induced DNA damage foci and their relation to chromatin organization. *Mutat Res*, 704(1-3):78–87, 2010.

[127] E. P. Rogakou, D. R. Pilch, A. H. Orr, V. S. Ivanova, and W. M. Bonner. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem*, 273(10):5858–5868, March 1998.

[128] Emmy P. Rogakou, Chye Boon, Christophe Redon, and William M. Bonner. Megabase chromatin domains involved in DNA double-strand breaks in vivo. *The Journal of Cell Biology*, 146(5):905–916, September 1999.

[129] J. Bewersdorf, B.T. Bennett, and K.L. Knight. H2AX chromatin structures and their response to DNA damage revealed by 4Pi microscopy. *Proceedings of the National Academy of Sciences*, 103(48):18137–18142, 2006.

[130] Edyta Marcon and Peter B Moens. The evolution of meiosis: recruitment and modification of somatic DNA-repair proteins. *BioEssays: news and reviews in molecular, cellular and developmental biology*, 27(8):795–808, August 2005.

[131] John R. Walker, Richard A. Corpina, and Jonathan Goldberg. Structure of the ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature*, 412(6847):607–614, August 2001.

[132] Jacob A Aten, Jan Stap, Przemek M Krawczyk, Carel H van Oven, Ron A Hoebe, Jeroen Essers, and Roland Kanaar. Dynamics of DNA double-strand breaks revealed by clustering of damaged chromosome domains. *Science (New York, N.Y.)*, 303(5654):92–95, January 2004.

[133] M.J. Kruhlak, A. Celeste, G. Dellaire, O. Fernandez-Capetillo, W.G. Müller, J.G. McNally, D.P. Bazett-Jones, and A. Nussenzweig. Changes in chromatin structure and mobility in living cells at sites of DNA double-strand breaks. *The Journal of Cell Biology*, 172(6):823–834, 2006.

[134] Evi Soutoglou, Jonas F Dorn, Kundan Sengupta, Maria Jasin, Andre Nussenzweig, Thomas Ried, Gaudenz Danuser, and Tom Misteli. Positional stability of single double-strand breaks in mammalian cells. *Nature Cell Biology*, 9(6):675–682, June 2007.

[135] B. Jakob, J. Splinter, M. Durante, and G. Taucher-Scholz. Live cell microscopy analysis of radiation-induced DNA double-strand break motion. *PNAS*, January 2009.

[136] Przemek M. Krawczyk, T. Borovski, J. Stap, T. Cijsouw, R. ten Cate, J. P. Medema, R. Kanaar, N. A. P. Franken, and J. A. Aten. Chromatin mobility is increased at sites of DNA double-strand breaks. *Journal of Cell Science*, 125(9):2127–2133, 2012.

[137] Nabieh Ayoub, Anand D Jeyasekharan, Juan A Bernal, and Ashok R Venkitaraman. HP1-beta mobilization promotes chromatin changes that initiate the DNA damage response. *Nature*, 453(7195), May 2008.

[138] Céline Baldeyron, Gaston Soria, DaniÃÍle Roche, Adam J L Cook, and GeneviÃÍve Almouzni. HP1alpha recruitment to DNA damage by p150CAF-1 promotes homologous recombination repair. *The Journal of cell biology*, 193(1):81–95, April 2011.

[139] Jung-Ae Kim, Michael Kruhlak, Farokh Dotiwala, André Nussenzweig, and James E. Haber. Heterochromatin is refractory to γ-H2AX modification in yeast and mammals. *The Journal of Cell Biology*, 178(2):209–218, 2007.

[140] Ian G Cowell, Nicola J Sunter, Prim B Singh, Caroline A Austin, Barbara W Durkacz, and Michael J Tilby. gammaH2AX foci form preferentially in euchromatin after ionising-radiation. *PLoS One*, 2(10):e1057, 2007.

[141] R. S. Vasireddy, T. C. Karagiannis, and A. El-Osta. gamma-radiation-induced gammaH2AX formation occurs preferentially in actively transcribing euchromatic loci. *Cell Mol Life Sci*, 67(2):291–294, January 2010.

[142] Burkhard Jakob, Jörn Splinter, Sandro Conrad, Kay-Obbe Voss, Daniele Zink, Marco Durante, Markus Löbrich, and Gisela Taucher-Scholz. DNA double-strand breaks in heterochromatin elicit fast repair protein recruitment, histone H2AX phosphorylation and relocation to euchromatin. *Nucleic Acids Research*, 39(15):6489–6499, August 2011.

[143] Irene Chiolo, Aki Minoda, Serafin U Colmenares, Aris Polyzos, Sylvain V Costes, and Gary H Karpen. Double-strand breaks in heterochromatin move outside of a dynamic HP1a domain to complete recombinational repair. *Cell*, 144(5):732–744, March 2011.

[144] Aaron A. Goodarzi and Penny A. Jeggo. The heterochromatic barrier to DNA double strand break repair: How to get the entry visa. *International Journal of Molecular Sciences*, 13(9):11844–11860, September 2012.

[145] R. K. Pathria. *Statistical Mechanics*. Butterworth Heineman, Oxford, UK, second edition, 1996.

[146] Evi Soutoglou and Tom Misteli. Mobility and immobility of chromatin in transcription and genome stability. *Current opinion in genetics & development*, 17(5):435–442, October 2007.

[147] Angelo Rosa and Ralf Everaers. Structure and dynamics of interphase chromosomes. *PLoS Computational Biology*, 4(8):e1000153, August 2008.

[148] Gert Strobl. *The physics of polymers*. Springer, Berlin ; Heidelberg [u.a.], 2. corr. ed. edition, 1997.

[149] Alexander Grosberg and Alexei R. Chochlov. *Statistical physics of macromolecules*. AIP series in polymers and complex materials. American Institute of Physics, New York, 1994.

[150] L.D. Landau and E.M. Lifshitz. *Elastizitätstheorie*. Number 7 in Lehrbuch der theoretischen Physik ; 7. Akad.-Verl., Berlin, 3. aufl. edition, 1970.

[151] Hsiao-Ping Hsu, Wolfgang Paul, and Kurt Binder. Standard definitions of persistence length do not describe the local "intrinsic"stiffness of real polymer chains. *Macromolecules*, 43(6):3094–3102, March 2010.

[152] William D.C. Moebs. A monte carlo simulation of chemical reactions. *Mathematical Biosciences*, 22:113–120, 1974.

[153] Pierre Bremaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer, May 1999.

[154] Chaitanya Athale. Monte carlo cell simulations. *Genome Biology*, 3(1):reports2001, December 2001.

[155] Peter Jäckel. *Monte Carlo Methods in Finance*. Wiley, April 2002.

[156] David Vose. *Quantitative risk analysis: a guide to Monte Carlo simulation modelling*. Wiley, Chichester, 1996.

[157] A. Berdondini, M. Bettuzzi, D. Bianconi, R. Brancaccio, F. Casali, S. Cornacchia, A. Flisch, J. Hofmann, N. Lanconelli, M.P. Morigi, A. Pasini, A. Rossi, C. Sauerwein, and M. Simon. Monte carlo optimization of an industrial tomography system. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 580(1):771–773, September 2007.

[158] Mark Chang. *Monte Carlo Simulation for the Pharmaceutical Industry: Concepts, Algorithms, and Case Studies*. CRC Press, October 2010.

[159] J. G. Park, C. H. Kim, M. C. Han, S. H. Jung, J. B. Kim, and J. Moon. Optimization of detection geometry for industrial SPECT by monte carlo simulations. *Journal of Instrumentation*, 8(04):C04006, April 2013.

[160] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, December 1953.

[161] Neal Madras and AlanD. Sokal. The pivot algorithm: A highly efficient monte carlo method for the self-avoiding walk. *Journal of Statistical Physics*, 50(1-2):109–186, 1988.

[162] A.D. Sokal. Monte carlo methods in statistical mechanics: Foundations and new algorithms. In *Functional Integration: Basics and Applications*. Springer, 1997.

[163] I. Carmesin and K. Kremer. The bond fluctuation method: a new effective algorithm for the dynamics of polymers in all spatial dimensions. *Macromolecules*, 21(9):2819–2823, 1988.

[164] H.P. Deutsch and K. Binder. Interdiffusion and self-diffusion in polymer mixtures: A monte carlo study. *The Journal of Chemical Physics*, 94(3):2294–2304, 1991.

[165] Dennis C. Rapaport. *The art of molecular dynamics simulation*. Cambridge Univ. Press, Cambridge [u.a.], 2. ed., 4. print. edition, 2009.

[166] H.J. Limbach, A. Arnold, B.A. Mann, and C. Holm. Espresso – an extensible simulation package for research on soft matter systems. *Computer Physics Communications*, 174(9):704–727, May 2006.

[167] S. P. Williams, B. D. Athey, L. J. Muglia, R. S. Schappe, A. H. Gough, and J. P. Langmore. Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length. *Biophysical Journal*, 49(1):233 – 248, 1986.

[168] J. R. Swedlow and T. Hirano. The making of the mitotic chromosome: Modern insights into classical questions. *Molecular Cell*, 11(3):557 – 569, 2003.

[169] A.L. Bak, J. Zeuthen, and F.H. Crick. Higher-order structure of human mitotic chromosomes. *Proceedings of the National Academy of Sciences of the United States of America*, 74(4):1595–1599, 1977.

[170] J. Sedat and L. Manuelidis. A direct approach to the structure of eukaryotic chromosomes. *Cold Spring Harbor Symposia on Quantitative Biology*, 42:331–350, 1978.

[171] M.O. Christensen, M.K. Larsen, H.U. Barthelmes, R. Hock, C.L. Andersen, E. Kjeldsen, B.R. Knudsen, O. Westergaard, F. Boege, and C. Mielke. Dynamics of human DNA topoisomerases IIÎś and IIÎš in living cells. *The Journal of Cell Biology*, 157(1):31–44, 2002.

[172] R. A. Oliveira, S. Heidmann, and C. E. Sunkel. Condensin i binds chromatin early in prophase and displays a highly dynamic association with drosophila mitotic chromosomes. *Chromosoma*, 116:259–274, 2007.

[173] N. Kireeva, M. Lakonishok, I. Kireev, T. Hirano, and A.S. Belmont. Visualization of early chromosome condensation. *The Journal of Cell Biology*, 166(6):775–785, 2004.

[174] A.S. Belmont and K. Bruce. Visualization of g1 chromosomes: a folded, twisted, supercoiled chromonema model of interphase chromatid structure. *The Journal of Cell Biology*, 127(2):287–302, 1994.

[175] Y.G. Strukov, Y. Wang, and A.S. Belmont. Engineered chromosome regions with altered sequence composition demonstrate hierarchical large-scale folding within metaphase chromosomes. *The Journal of Cell Biology*, 162(1):23–35, 2003.

[176] J.F. Marko and E.D. Siggia. Polymer models of meiotic and mitotic chromosomes. *Mol. Biol. Cell*, 8(11):2217–2231, 1997.

[177] J.F. Marko and M.G. Poirier. Micromechanics of chromatin and chromosomes. *Biochemistry and Cell Biology*, 81(3):209–220, 2003.

[178] J. Dekker, K. Rippe, M. Dekker, and N. Kleckner. Capturing chromosome conformation. *Science*, 295(5558):1306–1311, 2002.

[179] M. Simonis, P. Klous, E. Splinter, Y. Moshkin, R. Willemsen, E. de Wit, B. van Steensel, and W. de Laat. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nature Genetics*, 38(11):1348–1354, 2006.

[180] E. Lieberman-Aiden, N.L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B.R. Lajoie, P.J. Sabo, M.O. Dorschner, R. Sandstrom, B. Bernstein, M.A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L.A. Mirny, E.S. Lander, and J. Dekker. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 326(5950):289–293, 2009.

[181] P. Fraser. Transcriptional control thrown for a loop. *Current Opinion in Genetics and Development*, 16(5):490 – 495, 2006.

[182] Manfred Bohn, Dieter W. Heermann, and Roel van Driel. Random loop model for long polymers. *Physical Review E*, 76(5):051805, November 2007.

[183] Julio Mateos-Langerak, Manfred Bohn, W. de Leeuw, O. Giromus, E.M.M. Manders, P.J. Verschure, M.H.G. Indemans, H.J. Gierman, D.W. Heermann, R. van Driel, and S. Goetze. Spatially confined folding of chromatin in the interphase nucleus. *Proceedings of the National Academy of Sciences*, 106(10):3812–3817, 2009.

[184] Manfred Bohn and Dieter W. Heermann. Diffusion-driven looping provides a consistent framework for chromatin organization. *PLoS ONE*, 5(8):e12218, 2010.

[185] J.F. Marko. Linking topology of tethered polymer rings with applications to chromosome segratation and estimation of the knotting length. *Phys. Rev. E*, 79(5):051905, 2009.

[186] Manfred Bohn and Dieter W. Heermann. Topological interactions between ring polymers: Implications for chromatin loops. *The Journal of Chemical Physics*, 132(4):044904, 2010.

[187] Lisa H. Pope, Chee Xiong, and John F. Marko. Proteolysis of mitotic chromosomes induces gradual and anisotropic decondensation correlated with a reduction of elastic modulus and structural sensitivity to rarely cutting restriction enzymes. *Mol. Biol. Cell.*, 17(1):104–113, January 2006.

[188] R. Kawamura, L.H. Pope, M.O. Christensen, M. Sun, K. Terekhova, F. Boege, C. Mielke, A.H. Andersen, and J.F. Marko. Mitotic chromosomes are constrained by topoisomerase II–sensitive DNA entanglements. *The Journal of Cell Biology*, 188(5):653–663, 2010.

[189] G. Micheli, A. R. C. Luzzatto, M. T. CarrÃň, A. de Capoa, and F. Pelliccia. Chromosome length and DNA loop size during early embryonic development of xenopus laevis. *Chromosoma*, 102(7):478 – 483, 1993.

[190] A. Chiu, E. Revenkova, and R. Jessberger. DNA interaction and dimerization of eukaryotic SMC hinge domains. *Journal of Biological Chemistry*, 279(25):26233–26242, 2004.

[191] A. Sakai, K. Hizume, T. Sutani, K. Takeyasu, and M. Yanagida. Condensin but not cohesin SMC heterodimer induces DNA reannealing through protein-protein assembly. *EMBO Journal*, (11):2764–2775, 2003.

[192] S. Cocco, J. F. Marko, R. Monasson, A. Sarkar, and J. Yang. Force-extension behavior of folding polymers. *Eur. Phys. J. E*, 10(3):249–263, 2003.

[193] Yoshinori Watanabe. Sister chromatid cohesion along arms and at centromeres. *Trends in genetics: TIG*, 21(7):405–412, July 2005.

[194] Stephan Gruber, Christian H Haering, and Kim Nasmyth. Chromosomal cohesin forms a ring. *Cell*, 112(6):765–777, 2003.

[195] Ana-Maria Farcas, Pelin Uluocak, Wolfgang Helmhart, and Kim Nasmyth. Cohesin's concatenation of sister DNAs maintains their intertwining. *Molecular cell*, 44(1):97–107, October 2011. PMID: 21981921.

[196] T Tanaka, M P Cosma, K Wirth, and K Nasmyth. Identification of cohesin association sites at centromeres and along chromosome arms. *Cell*, 98(6):847–858, September 1999.

[197] Christian H Haering and Kim Nasmyth. Building and breaking bridges between sister chromatids. *BioEssays: news and reviews in molecular, cellular and developmental biology*, 25(12):1178–1191, December 2003.

[198] Armelle Lengronne, Yuki Katou, Saori Mori, Shihori Yokobayashi, Gavin P Kelly, Takehiko Itoh, Yoshinori Watanabe, Katsuhiko Shirahige, and Frank Uhlmann. Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature*, 430(6999):573–578, July 2004.

[199] Earl F Glynn, Paul C Megee, Hong-Guo Yu, Cathy Mistrot, Elcin Unal, Douglas E Koshland, Joseph L DeRisi, and Jennifer L Gerton. Genome-wide mapping of the cohesin complex in the yeast saccharomyces cerevisiae. *PLoS biology*, 2(9):E259, September 2004.

[200] El\textbackslashtextbackslashtextbackslashccin Unal, Jill M Heidinger-Pauli, and Douglas Koshland. DNA double-strand breaks trigger genome-wide sister-chromatid cohesion through eco1 (ctf7). *Science (New York, N.Y.)*, 317(5835):245–248, July 2007.

[201] Beom-Jun Kim, Yehua Li, Jinglan Zhang, Yuanxin Xi, Yumei Li, Tao Yang, Sung Yun Jung, Xuewen Pan, Rui Chen, Wei Li, Yi Wang, and Jun Qin. Genome-wide reinforcement of cohesin binding at pre-existing cohesin sites in response to ionizing radiation in human cells. *The Journal of biological chemistry*, 285(30):22784–22792, July 2010.

[202] Lena Ström, Charlotte Karlsson, Hanna Betts Lindroos, Sara Wedahl, Yuki Katou, Katsuhiko Shirahige, and Camilla Sjögren. Postreplicative formation of cohesion is required for repair and induced by a single DNA break. *Science (New York, N.Y.)*, 317(5835):242–245, July 2007. PMID: 17626884.

[203] Yang Zhang and Dieter W Heermann. Loops determine the mechanical properties of mitotic chromosomes. *PLoS ONE*, 6(12):e29225, 2011.

[204] Christopher L Woodcock and Rajarshi P Ghosh. Chromatin higher-order structure and dynamics. *Cold Spring Harbor perspectives in biology*, 2(5):a000596, May 2010.

[205] Mariano Barbieri, Mita Chotalia, James Fraser, Liron-Mark Lavitas, Josée Dostie, Ana Pombo, and Mario Nicodemi. Complexity of chromatin folding is captured by the strings and binders switch model. *Proceedings of the National Academy of Sciences of the United States of America*, 109(40):16173–16178, October 2012.

[206] Candace R Correa and Vivian G Cheung. Genetic variation in radiation-induced expression phenotypes. *American journal of human genetics*, 75(5):885–890, November 2004.

[207] Sylwia Kabacik, Alan Mackay, Narinder Tamber, Grainne Manning, Paul Finnon, Francois Paillier, Alan Ashworth, Simon Bouffler, and Christophe Badie. Gene expression following ionising radiation: Identification of biomarkers for dose estimation and prediction of individual response. *Int J Radiat Biol*, November 2010.

[208] Alexandre Mezentsev and Sally A. Amundson. Global gene expression responses to low- or high-dose radiation in a human three-dimensional tissue model. *Radiation Research*, 175:677–688, June 2011.

[209] Rogier Versteeg, Barbera D C van Schaik, Marinus F van Batenburg, Marco Roos, Ramin Monajemi, Huib Caron, Harmen J Bussemaker, and Antoine H C van Kampen. The human transcriptome map reveals extremes in gene density, intron length, GC content, and repeat pattern for domains of highly and weakly expressed genes. *Genome research*, 13(9):1998–2004, September 2003.

[210] Aurélien Bancaud, Sébastien Huet, Nathalie Daigle, Julien Mozziconacci, JoÃńl Beaudouin, and Jan Ellenberg. Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. *EMBO J*, 28(24):3785–3798, December 2009.

[211] Sandra Goetze, Julio Mateos-Langerak, Hinco J Gierman, Wim de Leeuw, Osdilly Giromus, Mireille H G Indemans, Jan Koster, Vladan Ondrej, Rogier Versteeg, and Roel van Driel. The three-dimensional structure of human interphase chromosomes is related to the transcriptome map. *Mol Cell Biol*, 27(12):4475–4487, June 2007.

[212] Chunhui Hou and Victor G Corces. Throwing transcription for a loop: expression of the genome in the 3D nucleus. *Chromosoma*, 121(2):107–116, April 2012.

[213] R K Sachs, G van den Engh, B Trask, H Yokota, and J E Hearst. A random-walk/giant-loop model for interphase chromosomes. *Proceedings of the National Academy of Sciences of the United States of America*, 92(7):2710–2714, March 1995.

[214] S de Nooijer, J Wellink, B Mulder, and T Bisseling. Non-specific interactions are sufficient to explain the position of heterochromatic chromocenters and nucleoli in interphase nuclei. *Nucleic acids research*, 37(11):3558–3568, June 2009.

[215] Ivan Junier, Olivier Martin, and FranÃğois KépÃĺs. Spatial and topological organization of DNA chains induced by gene co-localization. *PLoS computational biology*, 6(2):e1000678, February 2010.

[216] Hansjoerg Jerabek and Dieter W Heermann. Expression-dependent folding of interphase chromatin. *PloS one*, 7(5):e37525, 2012.

[217] Daan Noordermeer and Wouter de Laat. Joining the loops: beta-globin gene regulation. *IUBMB life*, 60(12):824–833, December 2008.

[218] Ji-Hoon Lee and Tanya T Paull. ATM activation by DNA double-strand breaks through the mre11-rad50-nbs1 complex. *Science*, 308(5721):551–554, April 2005.

[219] J. A. Downs, M. C. Nussenzweig, and A. Nussenzweig. Chromatin dynamics and the preservation of genetic information. *Nature*, 447(7147):951–958, June 2007.

[220] Judith Miné-Hattab and Rodney Rothstein. Increased chromosome mobility facilitates homology search during recombination. *Nature cell biology*, 14(5):510–517, May 2012. PMID: 22484485.

[221] Vincent Dion, Véronique Kalck, Chihiro Horigome, Benjamin D Towbin, and Susan M Gasser. Increased mobility of double-strand breaks requires mec1, rad9 and the homologous recombination machinery. *Nature cell biology*, 14(5):502–509, May 2012.

[222] Michael M Cox and John R Battista. Deinococcus radiodurans - the consummate survivor. *Nature reviews. Microbiology*, 3(11):882–892, November 2005.

[223] Smadar Levin-Zaidman, Joseph Englander, Eyal Shimoni, Ajay K Sharma, Kenneth W Minton, and Abraham Minsky. Ringlike structure of the deinococcus radiodurans genome: a key to radioresistance? *Science (New York, N.Y.)*, 299(5604):254–256, January 2003.

[224] Dea Slade, Ariel B Lindner, Gregory Paul, and Miroslav Radman. Recombination and replication in DNA repair of heavily irradiated deinococcus radiodurans. *Cell*, 136(6):1044–1055, March 2009.

[225] Bong-Gun Ju, Victoria V Lunyak, Valentina Perissi, Ivan Garcia-Bassets, David W Rose, Christopher K Glass, and Michael G Rosenfeld. A topoisomerase IIbeta-mediated dsDNA break required for regulated transcription. *Science (New York, N.Y.)*, 312(5781):1798–1802, June 2006.

[226] Michael C Haffner, Angelo M De Marzo, Alan K Meeker, William G Nelson, and Srinivasan Yegnasubramanian. Transcription-induced DNA double strand breaks: both oncogenic force and potential therapeutic target? *Clinical cancer research: an official journal of the American Association for Cancer Research*, 17(12):3858–3864, June 2011.

[227] Michael C Haffner, Martin J Aryee, Antoun Toubaji, David M Esopi, Roula Albadine, Bora Gurel, William B Isaacs, G Steven Bova, Wennuan Liu, Jianfeng Xu, Alan K Meeker, George Netto, Angelo M De Marzo, William G Nelson, and Srinivasan Yegnasubramanian. Androgen-induced TOP2B-mediated double-strand breaks and prostate cancer gene rearrangements. *Nature genetics*, 42(8):668–675, August 2010.

[228] John D. Weeks, David Chandler, and Hans C. Andersen. Role of repulsive forces in determining the equilibrium structure of simple liquids. *The Journal of Chemical Physics*, 54(12):5237–5247, June 1971.

[229] Kurt Kremer and Gary S. Grest. Dynamics of entangled linear polymer melts: A molecularâĂŘdynamics simulation. *The Journal of Chemical Physics*, 92(8):5057–5086, April 1990.

[230] G K Dasika, S C Lin, S Zhao, P Sung, A Tomkinson, and E Y Lee. DNA damage-induced cell cycle checkpoints and DNA strand break repair in development and tumorigenesis. *Oncogene*, 18(55):7883–7899, December 1999.

[231] Muriel Grenon, Chris Gilbert, and Noel F. Lowndes. Checkpoint activation in response to double-strand breaks requires the Mre11/Rad50/Xrs2 complex. *Nature Cell Biology*, 3(9):844–847, September 2001.

[232] Lisa Woodbine, H Brunton, A A Goodarzi, A Shibata, and P A Jeggo. Endogenously induced DNA double strand breaks arise in heterochromatic DNA regions and require ataxia telangiectasia mutated and artemis for their repair. *Nucleic acids research*, 39(16):6986–6997, September 2011.

[233] Pranav Oza, Sue L Jaspersen, Adriana Miele, Job Dekker, and Craig L Peterson. Mechanisms that regulate localization of a DNA double-strand break to the nuclear periphery. *Genes Dev*, 23(8):912–927, April 2009.

[234] John F. Ward. DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability. *Progress in Nucleic Acid Research and Molecular Biology*, 35:95–125, 1988.

[235] Judith Campisi and Fabrizio di Fagagna. Cellular senescence: when bad things happen to good cells. *Nature Reviews Molecular Cell Biology*, 8(9):729–740, September 2007.

[236] Marianne B Sowa, Wilfried Goetz, Janet E Baulch, Dinah N Pyles, Jaroslaw Dziegielewski, Susannah Yovino, Andrew R Snyder, Sonia M de Toledo, Edouard I Azzam, and William F Morgan. Lack of evidence for low-LET radiation induced bystander response in normal human fibroblasts and colon carcinoma cells. *International journal of radiation biology*, 86(2):102–113, February 2010.

[237] William M. Bonner, Christophe E. Redon, Jennifer S. Dickey, Asako J. Nakamura, Olga A. Sedelnikova, Stéphanie Solier, and Yves Pommier. ÎşH2AX and cancer. *Nature Reviews Cancer*, 8(12):957–967, 2008.

[238] Yong-Chul Kim, Gabi Gerlitz, Takashi Furusawa, Frederic Catez, Andre Nussenzweig, Kyu-Seon Oh, Kenneth H. Kraemer, Yosef Shiloh, and Michael Bustin. Activation of ATM depends on chromatin interactions occurring before induction of DNA damage. *Nature Cell Biology*, 11(1):92–96, January 2009.

[239] Martijn S. Luijsterburg, Christoffel Dinant, Hannes Lans, Jan Stap, Elzbieta Wiernasz, Saskia Lagerwerf, DaniÃńl O Warmerdam, Michael Lindh, Maartje C Brink, Jurek W Dobrucki, Jacob A Aten, Maria I Fousteri, Gert Jansen, Nico P Dantuma, Wim Vermeulen, Leon H F Mullenders, Adriaan B Houtsmuller, Pernette J Verschure, and Roel van Driel. Heterochromatin protein 1 is recruited to various types of DNA damage. *The Journal of Cell Biology*, 185(4):577–586, May 2009.

[240] Robert B. Winter and Peter H. Von Hippel. Diffusion-driven mechanisms of protein translocation on nucleic acids. 2. the escherichia coli lac repressor-operator interaction: equilibrium measurements. *Biochemistry*, 20(24):6948–6960, 1981.

[241] Robert B. Winter, Otto G. Berg, and Peter H. Von Hippel. Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. the escherichia coli lac repressor-operator interaction: kinetic measurements and conclusions. *Biochemistry*, 20(24):6961–6977, 1981.

[242] Michael Slutsky and Leonid A Mirny. Kinetics of protein-DNA interaction: facilitated target location in sequence-dependent potential. *Biophysical journal*, 87(6):4021–4035, December 2004.

[243] Chris A Brackley, Mike E Cates, and Davide Marenduzzo. Effect of DNA conformation on facilitated diffusion. *Biochemical Society transactions*, 41(2):582–588, April 2013.

[244] Maximilian Bauer and Ralf Metzler. In vivo facilitated diffusion model. *PloS one*, 8(1):e53956, 2013.

[245] C. A. Brackley, M. E. Cates, and D. Marenduzzo. Intracellular facilitated diffusion: Searchers, crowders, and blockers. *Physical Review Letters*, 111(10):108101, September 2013.

[246] Samuel A. Isaacson, David M. McQueen, and Charles S. Peskin. The influence of volume exclusion by chromatin on the time required to find specific DNA binding sites by diffusion. *Proceedings of the National Academy of Sciences*, February 2011.

[247] Yann von Hansen, Roland R. Netz, and Michael Hinczewski. DNA-protein binding rates: Bending fluctuation and hydrodynamic coupling effects. *The Journal of Chemical Physics*, 132:135103, 2010.

[248] Michael A. Lomholt, Tobias Ambjörnsson, and Ralf Metzler. Optimal target search on a fast-folding polymer chain with volume exchange. *Physical Review Letters*, 95(26):260603, December 2005.

[249] Dana Branzei and Marco Foiani. Regulation of DNA repair throughout the cell cycle. *Nature Reviews Molecular Cell Biology*, 9(4):297–308, April 2008.

[250] David W. Scott. *Multivariate density estimation*. Wiley series in probability and mathematical statistics ; A Wiley Interscience publication. Wiley, New York, 1992.

[251] Douglas B. West. *Introduction to Graph Theory (2nd Edition)*. Prentice Hall, August 2000.

[252] Alessandro Vespignani. Modelling dynamical processes in complex socio-technical systems. *Nat Phys*, 8(1):32–39, January 2012.

[253] Paolo Bajardi, Chiara Poletto, Jose J. Ramasco, Michele Tizzoni, Vittoria Colizza, and Alessandro Vespignani. Human mobility networks, travel restrictions, and the global spread of 2009 h1n1 pandemic. *PLoS ONE*, 6(1):e16591, 01 2011.

[254] Sergey V. Buldyrev, Roni Parshani, Gerald Paul, H. Eugene Stanley, and Shlomo Havlin. Catastrophic cascade of failures in interdependent networks. *Nature*, 464(7291):1025–1028, April 2010.

[255] Brian J. O'Roak, Laura Vives, Santhosh Girirajan, Emre Karakoc, Niklas Krumm, Bradley P. Coe, Roie Levy, Arthur Ko, Choli Lee, Joshua D. Smith, Emily H. Turner, Ian B. Stanaway, Benjamin Vernot, Maika Malig, Carl Baker, Beau Reilly, Joshua M. Akey, Elhanan Borenstein, Mark J. Rieder, Deborah A. Nickerson, Raphael Bernier, Jay Shendure, and Evan E. Eichler. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature*, 485(7397):246–250, May 2012.

[256] Arunachalam Vinayagam, Ulrich Stelzl, Raphaele Foulle, Stephanie Plassmann, Martina Zenkner, Jan Timm, Heike E. Assmus, Miguel A. Andrade-Navarro, and Erich E. Wanker. A directed protein interaction network for investigating intracellular signal transduction. *Science Signaling*, 4(189):rs8, September 2011.

[257] B. Delaunay. Sur la sphère vide. *Izvestia Akademii Nauk SSSR, Otdelenie Matematicheskikh i Estestvennykh Nauk*, 6:793–800, 1934.

[258] Franz Aurenhammer. Voronoi diagrams – a survey of a fundamental geometric data structure. *ACM Computing Surveys*, 23(3):345–405, 1991.

[259] Cedric Cagliero, Ralph S Grand, M Beatrix Jones, Ding J Jin, and Justin M O'Sullivan. Genome conformation capture reveals that the escherichia coli chromosome is organized by replication and transcription. *Nucleic acids research*, 41(12):6058–6071, July 2013.

[260] Stefan Grob, Marc W Schmid, Nathan W Luedtke, Thomas Wicker, and Ueli Grossniklaus. Characterization of chromosomal architecture in arabidopsis by chromosome conformation capture. *Genome biology*, 14(11):R129, November 2013.

[261] Songling Li and Dieter W. Heermann. Transcriptional regulatory network shapes the genome structure of saccharomyces cerevisiae. *Nucleus*, 4(3):216–228, May 2013.

[262] Tamara Salem, Tiphanie Gomard, Franck Court, Gabriel Moquet-Torcy, Frédérique Brockly, Thierry Forné, and Marc Piechaczyk. Chromatin loop organization of the junb locus in mouse dendritic cells. *Nucleic acids research*, 41(19):8908–8925, October 2013.

[263] Elnaz Alipour and John F Marko. Self-organization of domain structures by DNA-loop-extruding enzymes. *Nucleic acids research*, October 2012.

[264] Wulan Deng and Gerd A Blobel. Do chromatin loops provide epigenetic gene expression states? *Curr Opin Genet Dev*, 20(5):548–554, October 2010.

[265] Keishi Shintomi and Tatsuya Hirano. Sister chromatid resolution: a cohesin releasing network and beyond. *Chromosoma*, 119(5):459–467, October 2010.

[266] Toyoko Tsukuda, Kelly M. Trujillo, Emmanuelle Martini, and Mary Ann Osley. Analysis of chromatin remodeling during formation of a DNA double-strand break at the yeast mating type locus. *Methods*, 48(1):40–45, 2009.