Polymeric models for chromosome organization: Impact of cross-linkers, crowders and confinement

By

Amit Kumar
PHYS07201404012

Institute Of Physics, Bhubaneswar

A thesis submitted to the
Board of Studies in Physical Sciences
In partial fulfillment of requirements
For the Degree of

DOCTOR OF PHILOSOPHY
of
HOMI BHABHA NATIONAL INSTITUTE

August, 2020

Homi Bhabha National Institute

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Mr Amit Kumar entitled "**Polymeric models for chromosome organization : Impact of cross-linkers, crowders and confinement**", and recommend that it may be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Chairman -**Prof. Arun M. Jayannavar** Date:27/11/2020

Guide /Convener - Dr. **Debasish Chaudhuri** Date:27/11/2020

Co-guide (if any) NA Date:

Examiner - **Prof. Surajit Sengupta** Date:27/11/2020

Member 1 - **Dr. Goutam Tripathy** Date:27/11/2020

Member 2 - **Prof. Sudipta Mukherji** Date:27/11/2020

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the dissertation to HBNI.

I/We hereby certify that I/we have read this thesis prepared under my/our direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date: 27/11/2020
Place: Bhubaneswar

27/11/2020

Co-guide (if applicable) **Guide**

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfilment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgement the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Amit Kumar

24/11/2020

Amit Kumar

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

Amit kumar
24/11/2020
Amit Kumar

List of Publications arising from the thesis

Journal

[1] "Cross-linker mediated compaction and local morphologies in a model chromosome", **Amit Kumar** and Debasish Chaudhuri *Journal of Physics: Condensed Matter* **31**, 354001 (2019).

[2] "Impact of crowders on the morphology of bacterial chromosome", **Amit Kumar**, Pinaki Swain, Bela Mulder, and Debasish Chaudhuri, *Europhysics Letters* **128**, 68003 (2019).

Conferences

[1] International Conference on Statistical Physics, STATPHYS- KOLKATA X, Presidency University, Kolkata, India, November 26-29, 2019.

[2] Indian Statistical Physics Community Meeting (ISPCM-2019), International Center for Theoretical Sciences (ICTS), Bangalore, India, February 14-16, 2019.

[3] Puri Polymer Conference, Puri, India, December, 12-14, 2018.

[4] Soft Matter Young Investigator Meet (SYIM), Shimla, India, May 23-25, 2018.

Amit kumar
24/11/2020
Amit Kumar

DEDICATIONS

Dedicated to my guardians

Vivekanad Pandey, Chandeswar Mishra, and The Guru Maharaj

ACKNOWLEDGEMENTS

This thesis is an outcome of about five years of research work. The constant guidance, support and encouragement of many individuals have made this endeavor possible. It is my pleasure to have the opportunity now to express gratitude to all of them.

First and foremost, I would like to thank my thesis supervisor Prof. Debasish Chaudhuri for his teachings, ideas, and unconditional support. His constant guidance led me to develop molecular dynamics simulation skills from the scratch. He introduced me to subtle aspects of soft condensed matter physics, polymer physics, biological physics and other important academic knowledge from the beginning of my PhD setting a path for evolution of my research. His work ethics and ability to work on diverse problems persistently is something I want to inculcate in my life. His dedication to science remains an inspiration for me. I will always be grateful to him for all his contributions.

I thank Prof. Bela M. Mulder and Dr. Pinaki Swain for a fruitful collaboration. Pinaki helped me with setting up the earlier simulations which constitute major part of my work on bacterial chromosome organization.

My Doctoral Committee members, Prof. Goutam Tripathy, Prof. Arun M Jayannavar, and Prof. Sudipta Mukherji carefully examined my work during annual review seminar and provided valuable suggestions. I am grateful to them. It is my pleasure to thank all the faculty members of the Institute of Physics, Bhubaneswar for their timely advice.

The presented work in the thesis would not have been possible without the support of many friends and colleagues. I thank Biplab Bhattacherjee, Subhadip Ghosh, Amir Shee, Chaitra Hegde, Rajneesh Perhate for helping me with coding, conceptual aspects of scientific research and all the discussions. I would like to thank- Biswajit Das, Sujay Shil, Ganesh Paul, Partha Paul, Pratik, Vikas, Avnish, Chitrak, Arpan, Debjyoti, Sudharsan, Atanu, Shreyansh Da, Priyo da, Chandan Da, and Arpan Da for making the stay at IOP memorable. I am fortunate to have friends like Ashish Thakur, Avaneesh Dubey, Bhupendra Singh, Harshad Gupta, Shalik Ram Jhosi, Sudheer Pandey, Lalit, Ankit Mishra, Ashutosh Mishra, and Rishabh Shukla in my life. Thanks to them for all the support.

I would also like to acknowledge the Department of Atomic Energy, Government of India, and the Homi Bhabha National Institute for the research fellowship. I would like to thank Institute of Physics, Bhubaneswar for providing the research facilities and memorable stay. The simulations related to all my works were performed on SAMKHYA, the high-performance computing facility at IOP, Bhubaneswar. Special thanks to our system manager Makrand Siddhabhatti, who not only provided all the help with HPC facility (SHAMKYA) but treated me like a family member.

There are some people whom you can never repay for their contributions in your life. My uncles, Vivekanad Pandey and Chandeswar Mishra lie in that category. Finally, I would like to thank my parents Anil Kumar Mishra and Kavita Mishra, and my brothers Sumit Mishra and Mritunjay Mishra for their support and faith in me.

Amit Kumar, Bhubaneswar, August 24, 2020.

Contents

List of Figures

7 Conclusion of the thesis

In this thesis, we have investigated the chromosomal organization at various scales, particularly focusing on the impact of DNA- associated proteins, molecular crowders, and confinement. In the first half of the thesis, we have studied and analyzed the local morphological changes due to non-specific DNA- binding proteins. With increasing density, such cross-linkers fold the polymer leading to a coil- globule transition, along which the chromatin displays morphological changes quantified in terms of formation of contacts, loops and zippers. Having established the loop- formation on chromatin, in the second part of the thesis, we introduced the feather-boa model, a coarse-grained polymer model for the whole chromosome. Within this model, the chromosome is described as a backbone chain attached to equispaced side-loops. We have shown that such a feather-boa model with big enough side loops, spontaneously adopts a helical shape when confined within a cylindrical cell, recapturing observed chromosome morphologies in rod-shaped bacteria, e.g., E. coli and B. subtilis [12, 13]. We used this model to investigate the impact of cytosolic molecular crowders on the organization of the bacterial chromosome. We found that the crowders can segregate from the polymer longitudinally to stabilize the helicoid-shaped chromosome into a nucleoid-like sub-volume of the cell, in the absence of any bounding membrane. Further, crowders can assemble into a complementary helical organization with respect to the chromosomal helix in cylindrical cells. This purely entropic effect provides a possible physical explanation to similar organizations of ribosomes with respect to nucleoid, as observed in E. coli bacteria [11]. In the following, we present a brief summary of the individual chapters presented in this thesis.

In chapter 1, we introduced the system of prokaryotic and eukaryotic chromosomes and their cellular environment. We discussed chromosomal organization at various scales, and the roles played by various DNA- associated proteins and local environment. We presented simple polymer models utilized for the description of DNA or chromosome at various scales.

In chapter 2, we considered a self- avoiding polymer model of chromatin, segments of which can get cross-linked by non- specific DNA binding proteins that diffuse in the environment and undergo passive binding- unbinding kinematics. Increasing the concentration of such cross-linkers shows chromatin compaction. Using numerical simulations and a mean field model, we demonstrated that such compaction is a continuous coil- globule transition, characterized by a unimodal probability distribution of end-to-end separation, diverging fluctuations of chain extension at the transition, and a critical slowing down of the relaxation dynamics of chromatin extension. Along the transition, the model chromatin shows morphological transformations characterized by the formation of simple and higher-order loops, inter-segment contacts, zippers, and changing scaling of sub- chain extensions. With the folding of polymer, complex higher order loops form, however, simple loops always dominate the local morphology with the maximum number appearing at the critical point.

In chapter 3, we replaced the effect of cross-linkers by an additional effective attraction between the chromatin segments. Increasing the relative strength of attraction with respect to temperature leads to a continuous coil- globule transition. Across this transition, local morphologies are again quantified in terms of contacts, loops, zipper, and sub-chain extension. They show qualitative agreement with the observations in chapter 2.

In chapter 4, we introduced the feather- boa polymer model of bacterial chromosome. Within a cylindrical confinement, like in E.coli bacteria, such a model chromosome spontaneously adopt a large scale helical organization beyond a threshold side-loop size. Along the long axis of the cylinder, the local density of monomers showed a periodic oscillation, onset of which coappeared with the helical organization.

In chapter 5, to mimic the impact of cytosolic crowding, we introduced non- additive

crowders to the system of feather- boa polymer confined inside a cylindrical volume. We found that the crowders and chromosome adopt complementary helical organization and show out-of-phase density oscillations along the long axis of the cylinder. We further coarse-grained the feather- boa polymer to an effective chain by replacing the side loops with an additional Gaussian- core repulsion between backbone monomers. We probed the impact of change in crowder size on the morphology of chromosome. For small crowders, the model chromosome opens up and crowders get homogeneously distributed within the cell. Increasing the crowder size first causes transverse segregation, leading to the complementary helical organization of the chromosome and crowders. Further increase in crowder size leads to longitudinal segregation of crowders and the chromosome. In this regime, crowders accumulate near the top and bottom of the cylinder compressing the chromosome in the central region in a membrane-less sub-volume like in bacterial nucleoid. With further increase in crowder size, the number of helical turns increases and the chromosome extension decreases reducing the helical pitch.

In chapter 6, we investigated the impact of crowder density on the chromosomal morphology. The smallest crowders that can penetrate any intra-chromosomal region does not impact the chromosome shape significantly. On the other hand, increasing the density of larger crowders can influence the relative organization. Increasing the crowder density causes the segregation of monomers and crowders influencing the chromosomal morphology. Again, we used the effective model replacing side-loops by the additional Gaussian- core repulsion between backbone monomers. For intermediate crowder size, changing the crowder density leads to transverse followed by longitudinal segregations and associated shape changes. The change in helicoid shape is analyzed carefully using winding number and turning number.

Summary

The highly folded structure of chromosome involves DNA and associated proteins [1]. The DNA chains having lengths ranging from millimeters (bacteria) to meters (mammals) are organized inside micron sized cells in bacteria and nuclear volume in eukaryotic cells, respectively. This requires a tremendous amount of compaction, concomitant with functional organization of the chromosome to allow information processing in terms of gene expression, regulation and DNA replication [1–3]. The chromosomal organization displays local loop formation that involves passive binders or active loop extruders [4–8]. Apart from that, strong confinement, and molecular crowding in the environment play important role in the chromosomal organization [9–15].

In the first part of this thesis, we consider the impact of binder proteins on chromatin morphology. We describe the chromatin as a self- avoiding chain, and assume an attractive potential between the binders and monomers. Using molecular dynamics simulations in the presence of a Langevin heat bath, we investigate the impact of increasing binder density to find chromatin- folding. Using simulation results and a mean field theory we characterize this as a continuous transition. Mediated by the attraction to the polymer, the binders show local clustering and negative cross- correlation with monomer density. The local density of binders is restricted above due to the inter- binder repulsion. At the transition point, a finite size analysis suggests divergence of the polymer size fluctuations, relaxation time, and susceptibility. Thus, near the transition point, a little change in the binder density can generate significant change in the chromatin compaction, allowing a potentially useful control mechanism for the cell.

Along the transition, the chromatin undergoes local morphological changes which we characterize in terms of contact probabilities between chromatin segment, topological loops of various orders, and zippering. The contact probability and subchain extension show behaviors similar to that in interphase human chromosomes [16, 17]. With increasing binder density, the number of simple loops first increases to attain a maximum at the critical point, and then decreases as higher- order loops proliferate. Concomitantly, the mean separation between the first order loops first decreases to reach a minimum at the critical point, and then increases as the polymer folds further. With increasing binder density, the probability of higher order loops increases, but that of the simple loops dominate the local morphology. The zipper fraction of the chain increases monotonically as the polymer folds. The cross-linkers provide effective attraction between different segments of the model chromatin. Replacing the cross-linkers with increasing inter-segment attraction shows similar coil- globule transition, and change in local structures.

Having established the local loop morphology, in the second part of the thesis, we consider the feather-boa model, a coarse-grained model of the whole chromosome consisting of polymeric loops attached to a backbone chain [18, 19]. Within a cylindrical confinement as in rod-shaped bacteria like E. coli or B. subtilis, the model chromosome spontaneously adopt a helical morphology beyond a threshold size of the side-loops. This reproduces the

observed helicoid shapes in bacterial chromosome, providing a physical mechanism behind their emergence in different species [12, 13]. We use this model to investigate the joint effect of molecular crowders and confinement on chromosomal organization. The cytosolic crowders can be of various sizes. While the smallest ones can penetrate everywhere, relatively larger crowders undergo spatial segregation with respect to the chromosome. We investigate the impact of changing crowder size and density on the shape and size of the model chromosome. In order to do this systematically, we consider a further coarse- grained model of the chromosome, replacing the side- loops of the feather- boa model by an additional Gaussian-core repulsion between the backbone monomers [9, 10]. With increasing crowder-size, the chromosome and crowders first segregate in the radial direction, followed by a longitudinal segregation. The crowders localized near the top and bottom edges of the cylindrical confinement, can compress the chromosome further within a sub-volume of the cell localizing it near the cell center, as in bacterial nucleoid [12, 13]. The amount of such compression depends on the crowder density, with higher density supporting a higher degree of compression. We find that the radial segregation of crowders generates a complimentary helical localization of them around the helical nucleoid. This is associated with an out-of-phase density modulation of the chromosome and crowders along the cell length. Similar behavior has recently been observed for ribosomes in live E. coli bacteria [11].

Thus, we present a study of the chromosomal organization at different length scales. First, we explore the impact of cross-linker proteins on the chromatin folding and the related changes in local morphology. In the second part, using a polymer- based model of the whole bacterial chromosome, we present a systematic investigation of the impact of cytosolic crowders within the cylindrical confinement of rod-shaped bacteria.

1 Introduction

This thesis presents a study of chromosomal organization at various scales, particularly focusing on the impact of DNA- associated proteins, molecular crowders, and confinement. In the first half of the thesis, we present a study of how non- specific DNA binding proteins lead to local morphological changes, e.g., forming loops and thereby folding the chromosome. In the second half of the thesis, we assume a looped structure of the chromosome and demonstrate the impact of molecular crowding in the emergent shape, size and relative organization of the chromosome in confinement. The complex of DNA and associated proteins constitute chromosomes, which are strongly compacted inside the cell nucleus, a membrane-bound compartment, in eukaryotes, and form a membrane-less compact nucleoid in prokaryotes like bacteria. In eukaryotes, chromosomes live in an environment of molecular crowding of nucleoplasm inside the nuclear- membrane bound compartment. In bacteria the nucleoid is stabilized in an environment of diverse molecular crowding of cytosol, confined by the bacterial cell wall. In this chapter, we provide a brief description of chromosomes and their environment, physical approaches towards describing them using polymer models, and an outline of the thesis.

1.1 Eukaryotic and prokaryotic cells

The structural, functional and biological building block of all living organisms is the cell. One of the simplest living organisms that can perform all basic biological processes independently, e.g., metabolism to keep the cell alive, and cell divisions to hand down the genetic material, are bacteria, prokaryotic cells that are devoid of membrane- bound organelles [1, 27]. On the

other hand, more complex multicellular organisms, e.g., animals and plants consist of trillions of eukaryotic cells, which are distinguished from prokaryotes by the presence of a proper cell nucleus, and membrane-bound organelles. All the cells have a protective outer layer that consists of a cell membrane and, in some cases, cell wall. The cell membrane is mostly made up of lipids, a phospholipid bilayer. The cell walls are present in most prokaryotes and plants. Both eukaryotic and prokaryotic cells contain cytoplasm, a fluid dispersion made up of cytosolic proteins and lipids, which in eukaryotes include long filaments like microtubules, actin filaments and intermediate filaments [1].

In a eukaryotic cell, the nucleus is covered by a plasma membrane. The long deoxyribonucleic acid (DNA) chains that contain the genetic information, are folded, compacted and organized within the nucleus with the help of various DNA- associated proteins [28]. The inside environment of nucleus is made of a highly viscous nucleoplasm, which is a fluid containing various proteins and macromolecules. The nuclear membrane separates the DNA from rest of the cell. The typical size of eukaryotic cells may vary from $10 - 100\,\mu$m. While typical size of nucleus lie within the range of $5 - 8\,\mu$m. In human cells, 23 pairs of chromosomes, each of which are made with a DNA of length $\sim$ 2m, are packed inside the nucleus of diameter $\sim 5\,\mu$m, which requires a minimum 10^4- fold folding of the DNA. Apart from DNA, the other genetic component of the cell is the ribonucleic acid (RNA). They participate in various functions, including gene regulation and protein production. The messenger RNA convey genetic information controlling protein synthesis in ribosomes, which are 25-30 nm sized macromolecular machines made of RNA and protein complexes [29]. The protein synthesis at ribosome proceeds by ribosomal RNA linking amino acids delivered by transfer RNA. Ribosomal structure in eukaryotes and bacteria are quite similar with a slightly smaller size of 20 nm.

The cytoplasm is composed of almost 80% water in which large numbers of proteins, lipid molecules and ions float with typical concentration of various macromolecules being approximately 300 g/L [30, 31]. Several membrane bound organelles like mitochondria, endoplasmic

reticulum, Golgi apparatus etc. float in the cytoplasm and serve specialized functions critical for the survival of the cell. The major part of eukaryotic cytoplasm is its cytoskeleton, a dynamic network of filaments, motor proteins and cross-linkers. In eukaryotes, cytoskeleton consists of three kind of fibers - filamentous actin (F-actin), microtubules and intermediate filaments. These semiflexible filaments are assembled and disassembled using monomeric units at different parts of the cell, as required. F-actins are polymers made from globular actin (G-actin) subunits. The diameter of F-actin is $\sim$ 6 nm and is the thinnest of all cytoskeletal filaments [32]. Microtubules have a hollow cylindrical structure made from 13 protofilaments, each of which is made from α-β tubulin dimers. The cylindrical structure of microtubules has an inner radius of 18 nm and outer radius of 25 nm. The intermediate filaments have a diameter $\sim$ 10 nm.

Different components of the cell preform different functions. For example, ribosomes synthesize protein molecules which are the building blocks of cells. Various proteins work as structural support units, provide chemical catalysis, and act as molecular motors which provide transportation of essential units inside the cell. For example, mitochondria performs oxidation of sugar molecules to create ATP which serve as the energy currency of the cell. Hence, it is called the power house of the cell. The endoplasmic reticulum packages the proteins, and Golgi apparatus acts as a coordinator in these processes. The cytoskeleton provides structural support to cell, facilitate motion and provide tracks for movement of molecular motors which carry vesicles filled with cargo and other organelles [32]. The smaller F-actins along with myosin motors form the actomyosin network that predominantly localize near the cell cortex and around the nuclear membrane. They mediate cell motility. On the other hand, the longer microtubule network maintains the cell size providing it rigidity, and plays important role in chromosomal segregation during cell division. The cytoskeleton governs the mechanical properties of the cell [33, 34], and displays active visco-elastic response [35].

The prokaryotic cells, on the other hand, lack membraneous organelles. They do not have a separate nucleus. The prokaryotic chromosome floats in the cytosol of the cell [27, 36].

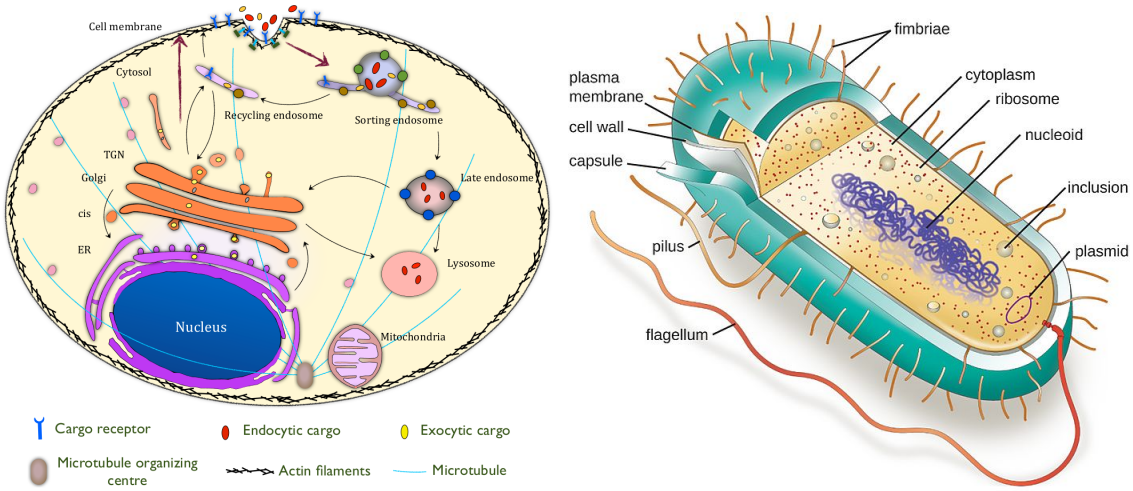


FIGURE 1.1: Schematics of eukaryotic cell with major components on left and E. coli cell on right. Figures are adopted from Ref. [20, 21].

However, the chromosome typically form a compact shape known as nucleoid which floats in the viscous cytosol at the middle of the cell [3, 25]. An example of prokaryotes which are studied extensively is the bacteria E. coli. It has a sphero-cylindrical cell with length 3-4 μm and diameter 0.8 μm. The chromosomal DNA in E. coli has a length of 1.6 mm and is compacted a 10^3-fold into a nucleoid that occupies only about 1/4-th of the cell volume. In the cytosol, various proteins molecules and other molecular machineries float creating an enormously crowded environment [37–39]. The glass-like environment due to large crowding density, strong polydispersity, and confinement gets fluidized by metabolic activity [40]. Abundant in them are, ribosomes which have typical size $\sim$ 20 nm. Apart from the nucleoid, bacteria contains a large number of small pieces of circular DNA called plamids, some of which carry significant genetic information requiring precise replication and segregation [41]. Bacterial cytosol contains homologues of all the cytoskeletal protein filaments, e.g., tubulin homologs like FtsZ, TubZ, actin- homologues like ParM and MreB, and intermediate filament homologues crescentin, FilP etc. [42, 43]. Bacterial cytosol do not contain large scale complex networks unlike eukaryotes, and they lack cytoskeletal motor proteins. However, the cytoskeletal filament homologues often perform similar functions

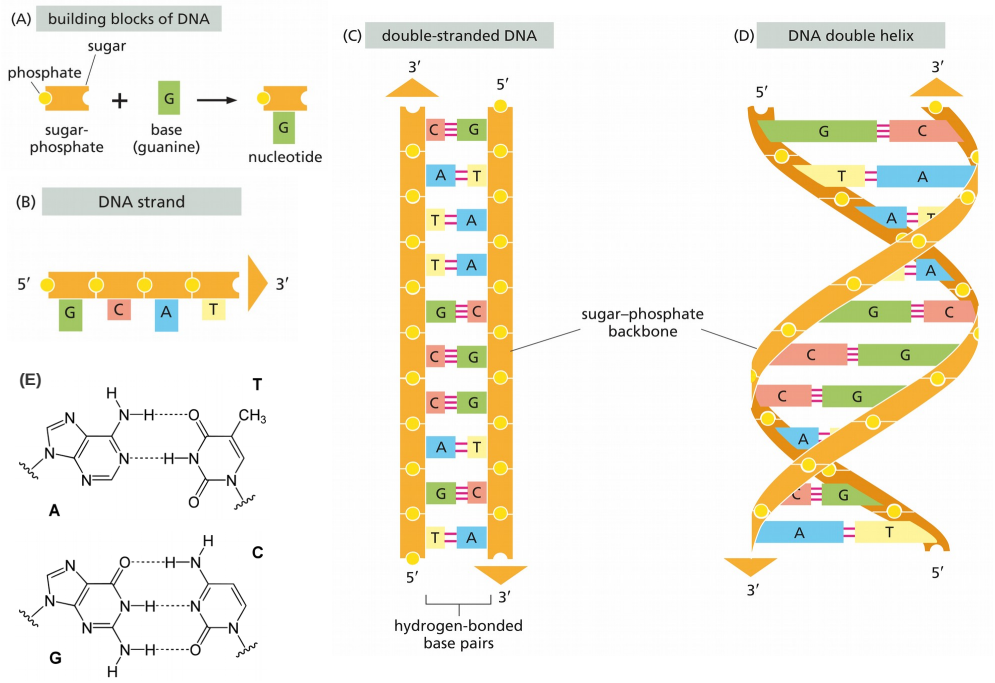


FIGURE 1.2: DNA structure: (A) Attachment of sugar phosphate and base to form a nucleotide. (B) Nucleotides string together to form a polynucleotide chain. (C) Two polynucleotides are held together by hydrogen bonds between two-bases of individual nucleotides. A pair of bound bases is called a base-pair (bp). (D) The two chains wrap around each other in a double-helix configuration reducing the Van der Waals base-stacking energy. (E) Hydrogen bonds between an A-T and G-C base pairs. Figures adopted from Ref. [1].

as in eukaryotes, for example, FtsZ forms Z-ring around which new cell wall forms during cell- division. The bi-directional elongation of ParM proteins mediate segregation of large plasmids [44].

1.2 DNA and its function

The double helix structure of DNA was determined from X-ray diffraction measurements by Watson and Crick in 1953. The DNA molecule consists of two polynucleotide chains running anti-parallel and wound around each other in a helical fashion [1, 28, 45]. Each

chain is made from sequence of four types of nucleotides which serve as monomeric units (see Fig.1.2). Each nucleotide is composed of three basic units- a five carbon sugar, a nitrogen base, and one or more phosphate groups. The phosphate groups and the nitrogen base are attached to a carbon-based sugar molecule. In DNA, deoxyribose is the sugar part and the nitrogen bases can be one of the four possibilities – two purines adenine (A), guanine (G), and two pyrimidines cytosine (C), and thymine (T) . The sugar and phosphate groups attach to each other with covalent bonds in alternating fashion making a backbone of sugar-phosphate repeats. To each sugar unit, one of the four bases are attached. The nucleotides are represented by their base units. Hence, symbols A, T, G or C are used to denote a nucleotide depending on its base unit. The two antiparallel chains of nucleotides are bound to each other by hydrogen bonds between base units. The bondings are such that pyrimidines bind to purines. The purines A and G are double-ring bases compared to the pyrimidines G and C which are single ring bases. A Purine always bind to a pyrimidine forming a complimentary base pair. The base A always bind to T with two bonds, and G bind to C with three bonds. Spatially, the base pair always lie inside the double helix, while the phosphate groups lie outside. The way in which sugar-phosphate units are linked together in individual polynucleoid chains gives a notion of direction in the chain. One end of a chain is called $3'$ hydroxyl and the other end is called $5'$ phosphate. In a DNA double- helix one chain runs from $3'$ to $5'$, while the other chain runs anti-parallel to it from $5'$ to $3'$. The pitch of the double-helix is ten base pair long. This is the length scale per turn. The separation between the two anti-parallel strands are almost constant and remains around 2 nm. This is the width of the DNA. Separation between two base pairs along the DNA is 0.34 nm. We often use this base pair (bp) size as a unit of the measure of DNA length in this thesis.

Primary function of the the DNA is to carry genetic information in form of genes. Different sequence of nucleotides A, T, G or C code different genetic information, like different sequence of alphabets constructing different words. A gene is a region of DNA (nucleotide sequence) which codes specific instruction for RNA molecules, which can be either messenger RNA

(mRNA) or functional RNA. DNA lengths depend on the species, e.g., the nucleus of a eukaryotic cell contains DNAs of length 3.2×10^9 bp ($\sim 2\,\mathrm{m}$). On the other hand, the E. coli cell has a single circular DNA of length 4.6×10^6 bp ($\sim 1.6\,\mathrm{mm}$). In human DNA, about 30,000 genes have been identified [1]. The 10^3 fold smaller DNA of E. coli bacteria carries around 4600 genes. The bacterial cells also contain small non-chromosomal DNA in the form of plasmids. Synthesis of RNA from the DNA occurs via a process called transcription. A chemical machinery called RNA polymerase, along with one or more transcription factors, binds to the promoter region, opens up the DNA double helix into a transcription bubble, reads out the nucleotide sequence of a gene and synthesizes a complimentary RNA strand by adding RNA nucleotides. The process of protein production is modulated by several mechanisms of gene regulation. This ranges from regulating transcriptional initiation, RNA processing and even post-transcriptional modification of proteins. One physical mechanism involves structural changes in the DNA organization. Strongly folded part of DNA has lesser frequency of transcription, e.g., heterochromatin, the transcriptionally less active high density regions of chromosome.

In the process of translation, messenger RNA (mRNA) is decoded by ribosome, building polypeptide chains by assembling different kinds of amino acids. Twenty different kind of amino acids strung in different sequence to form the protein molecules in live cells. The polypeptides fold into active proteins that perform their functions in the cell. In bacteria, translation occurs in cytosol as ribosomes bind to mRNAs. For functional efficiency, ribosomes gather near the nucleoid in bacteria. In eukaryotes, translation occurs in cytosol or on the membrane of endoplasmic reticulum to which ribosomes bind.

1.3 Chromosome

As we discussed earlier, the DNA is compacted by tremendous amount - at least 10^4 times for humans and 10^3 times for E. coli bacteria, to fit inside the cell. The nuclear or cellular

confinement brings a major degree of compaction of DNA, however, its local organization is largely achieved by various kind of packaging proteins which bind to DNA and organize it in either passive or active manner [1–3, 14, 25, 46]. This complex of DNA and proteins form the chromosome. The number and size of chromosomes greatly vary among different kind of cells. For example, human cells have 23- pairs of chromosomes, while E. coli bacteria have only one.

The amount of compaction of a chromosome in eukaryotic cells depends on the cell cycle. It mainly consists of two stages, the interphase and the mitotic phase. In the interphase the cell grows and gathers nutrients in Gap 1 (G1) phase, replicates its chromosomes in the Synthesis (S) phase, and a rapid cell growth in the Gap 2 (G2) phase, as the cell prepares for cell division. This is followed by a rapid cell division in mitosis. The sequence of events in mitosis is divided into prophase, pro-metaphase, metaphase, anaphase, and telophase. In the prophase, an early mitotic spindle is formed by the microtubules and centrosomes, as each replicated and bound sister chromatid pair starts to segregate from others, still within an intact nucleus. The nuclear membrane dissolves and sister chromatids get associated with the microtubule spindle in pro-metaphase. The chromosomes get into most compact configurations as they align around the mid-plane of the cell with the help of mitotic spindle in the metaphase. It is only in metaphase, that the more tightly packed chromosomes become normally visible under a light microscope. In anaphase, the daughter chromosomes segregate under the pulling force from the spindle. Finally the segregation completes in telophase and cytokinesis leads to formation of two daughter cells.

Visualization of chromosome in different stages of cell cycle has an interesting history. During the later half of nineteenth century, mitotic chromosomes were stained and observed under light microscopes [47, 48]. In the interphase, chromosomes spread out in the nucleus and made it difficult to visualize any structure. By 1956, work of Tjio and Levan clearly showed 46 human chromosomes during mitosis [49]. The current techniques of DNA

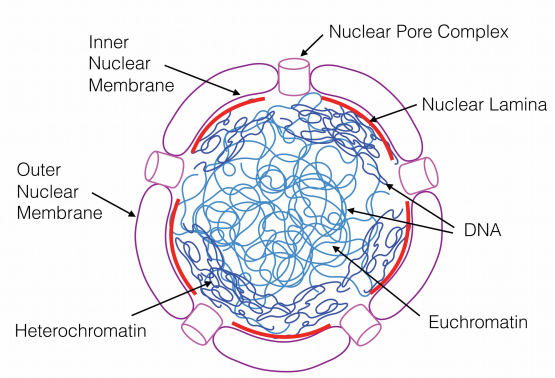


Figure 1.3: A schematic diagram of for interphase chromosomes inside nucleus. Figure adopted from Ref. [22].

visualization involves more sophisticated stainings, e.g., in using gene- specific (FISH) or aspecific (DAPI) fluorescence molecule tagging and using fluorescence microscopy or confocal microscopy.

In early 1980s, the fluorescent in situ hybridization (FISH) technique was developed in which a specific short sequence of DNA is used as a marker that identifies its complementary nucleotide sequence on the chromosome. First, these probe DNA pieces are made fluorescent by tagging them with specific fluorophores. The fluorophores are chemical compounds which can be excited by light, so that they re-emit light at a longer wavelength. The fluorophore tagging can be done with nick translation or polymerase chain reaction. In the second step, an in situ hybridization is performed. To do that, both the probe DNA segment and the chromosome are denatured, i.e, the double stranded helical segment is opened up into single strands using heat or chemical treatment. Then, the probe DNA fragments are mixed with chromosomal DNA. Consequently, the probe DNA fragments bind to their complimentary nucleotide sequence of the chromosomal DNA with hydrogen bonds at specific locations. This process is called molecular hybridization. The hybridized nucleotide sequence which were marked with probe DNA fragments, emit lights of different colors depending on the fluorophore used in tagging them. This is used in fluorescent microscopy or confocal

microscopy for better resolution.

Each chromosome can be painted with different colors using such a technique. A particular variant is known as multifluor FISH that makes different chromosome appear to be painted with different color. In this method, a collection of DNA sequence that acts as a probe is treated with fluorochromes that produce a single color. Afterwards hybridization with chromosome is performed which lead to appearance of different chromosomes with different colors. Using this technique it has been shown that the interphase chromosomes display a territorial organization in which different chromosomes occupy different spatial location inside the nucleus. Even pairs of homologous chromosomes are not generally located at the same position [1, 50].

The more densely packed heterochromatins are transcriptionally less active, compared to the normal euchromatins. Their predominantly peripheral localization [22] can be described in terms of activity dependent segregation of chromosomes as shown in recent numerical simulations [51]. A schematic diagram of nucleus encapsulating different chromosomes is illustrated in Fig. 1.3. In this diagram the euchromatin is represented by thin blue lines, and the heterochromatin by thick dark blue lines. The 46 chromosomes in human cells stay inside the tiny confinement of the nucleus with a typical DNA- density $\sim 10\,\mathrm{mg/mL}$ [52]. In such a high concentration of polymer solution, repulsion between individual sub-chains gets balanced by osmotic pressure by a sea of surrounding polymers [53]. As a result the individual chains are expected to behave as ideal chains. Such an ideal chain scaling behavior has been observed in FISH experiments for short DNA chains [54].

In contrast to eukaryotes, prokaryotic cells like bacteria have a single circular DNA which carries most of the genetic information. As has been pointed out earlier, the 4.6 Mbp (1.6 mm) long DNA in E. coli cell is compacted at least 1000 folds to form the nucleoid, that occupies 1/4-th of the cell volume having length 2-4 μm and diameter 0.8 μm. Initially, it was believed that the DNA is randomly organized inside bacterial cell. The compaction of bacterial chromosome with the help of various proteins can be described in terms of an

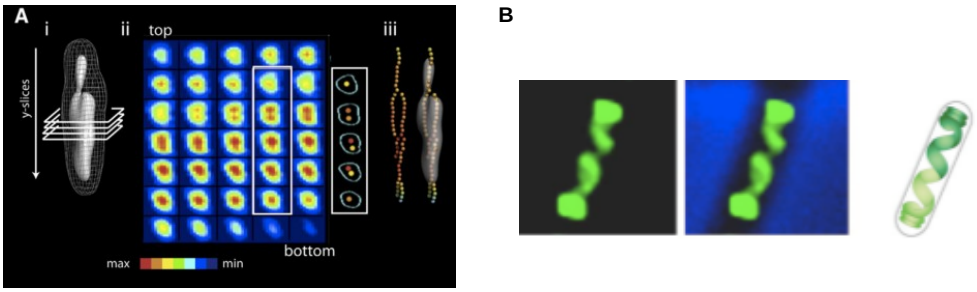


Figure 1.4: Experiments showing helical organization of chromosome in rod shaped bacteria. (A) E. coli nucleoid is scanned along long axis of the cell (*ii*). (*i*) and (*iii*) show the helicoid shape of the chromosome. Figure adopted from Ref. [12]. (B) B. subtilis chromosome shows helicoid shape. Figure adopted from Ref. [13].

effective poor solvent environment [55]. High resolution microscopy over the last one decade revealed the shape and size of the compacted nucleoid [12, 13, 56]. High resolution fluorescent microscopy of E. coli chromosomes reported a global helical organization of the nucleoid that lacks definite handedness [12, 57]. In the bacterium B. subtilis, newly replicated DNA was tagged with fluorescent DNA binding proteins (dNTP derivatives), which revealed helical organization of chromosome [13]. Further, whole cell cryo-electron tomography of the bacterium Bdellovibrio bacteriovorous suggests helical organization of the chromosome [58]. In E. coli, relaxing the cylindrical confinement of the cell leads to swelling of chromosome away from the original helical organization. This suggests that the shape and size of confinement plays a crucial role in dictating the emergent morphology of the chromosome [9]. Another recent study shows that the ribosomes crowd around the E. coli chromosome displaying modulating local densities along the long axis of the cell that are anti-correlated to the local density modulation of the chromosome [11].

Further small scale resolution is achieved using electron microscopy, that uses an electron beam instead of light, and magnetic coils instead of lens in performing the microscopy. To obtain contrast, electron- dense heavy metal staining is performed. The much smaller wavelengths allows a million fold magnification, and achieve resolution up to 1 nm in tunneling electron microscopy (TEM). The scanning electron microscopy (SEM) uses scattering from

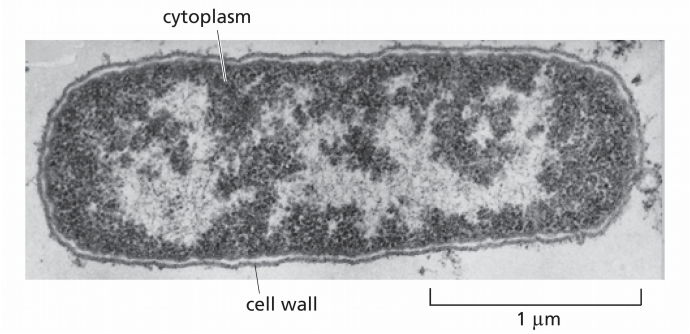


FIGURE 1.5: Electron micrograph of E. coli is shown with scale bar 1 μm. The light region shows the chromosome and the dark spots the crowders. Figure adopted from Ref. [1].

beam focussed on the sample. Using SEM, it is possible to create 3D images, however TEM works for 2D samples. Early electron microscopy of nucleoid revealed irregular organization of nucleoid. However, later electron microscopy and fluorescence based light microscopy revealed that the nucleoid occupies only a sub-volume of cell [11, 56, 59–61]. Fig. 1.5 shows an electron microscopy image of E. coli. The light portion shows the chromosome while dark dots show the crowders. It suggests that the crowders localize mostly near the two caps of the sphero-cylindrical volume, however small fraction of them also exist and penetrate the more central region primarily occupied by the nucleoid.

The chromosomal DNA in bacteria floats in the cytosol containing various kinds of crowder molecules including proteins, ions and other macromolecules. One of the bigger protein machineries are ribosome with typical size $\sim 20\,\mu$m which has about 55000 copies in E. coli cell [62–64]. In addition, almost 4600 copies of RNA polymerase (RNAPs) and about 221 copies of plasmids float in the cytosol [65]. The total number of proteins in E. coli cell is $\sim$ 3,000,000-4,000,000 [66], of which a fraction acts as DNA binding proteins and the rest as crowders. The overall steric repulsion from the crowders can potentially compress the chromosome. The effect of the crowders on relative organization of the chromosome is considered in the second half of this thesis.

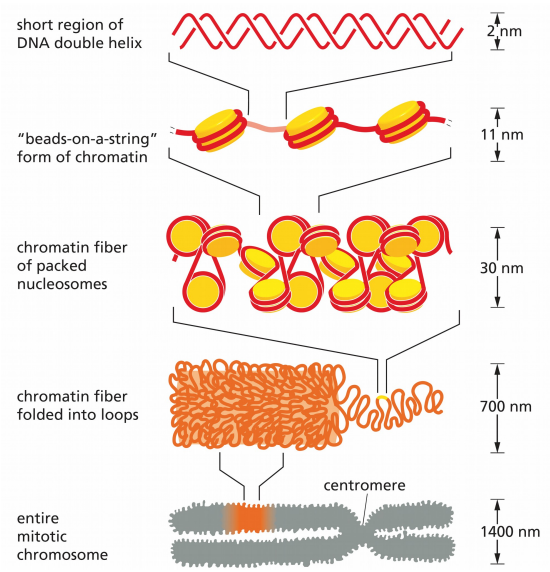


Figure 1.6: Hierarchical organization of eukaryotic chromosome from ds-DNA to mitotic chromosome via formation of nucleosome and 30nm chromatin fiber. Figure adopted from Ref. [1].

1.3.1 Protein- DNA association in eukaryotic chromosomes

DNA binds to various proteins towards condensing eukaryotic chromosomes. These proteins can have specific or non-specific affinity to DNA. The abundant proteins which generate packaging of DNA at the smallest length scales are histone proteins of which about 60 million copies are found in a typical eukaryotic cell [1]. Histone is a class of proteins with positive charge which allow them to bind with negatively charged DNA with electrostatic interaction in non-specific fashion. Histone proteins are found in five types and are conserved across all eukaryotic cells. There are also other non-histone proteins that can bind to DNA in gene- specific or non-specific fashion.

The double stranded DNA of a 147 bp stretch gets wrapped around the disk shaped octameric complex of histone proteins which contains two molecules of each histone proteins, H2A, H2B, H3 and H4. This disk shaped protein complex has a diameter of 11 nm. The 147 bp of DNA segment wraps around it about 1.7 times in a left-handed coil. Together

this unit is called a nucleosome (Fig. 1.6). Many nucleosomes form along the DNA, and consecutive nucleosomes remain connected to each other by linker DNA [67, 68], which can vary from few base-pairs up to 80 bp in length. This connected set of nucleosomes form a bead-on-string structure with diameter 11 nm and is known as chromatin thread. Such structures are observed in electron microscopy experiments as shown in Fig. 1.7(a_1). Formation of nucleosomes provide the first level of DNA compaction and reduces the chain length to one-third of its original value. In the next level, neighboring nucleosomes are packed together with the help of another histone protein H1 which is called the linker histone. It is believed that this class of histones attract several nucleosomes in close spatial proximity and arrange them in the form of a repeating array. Electron microscopy experiments reveal such structures as is shown in Fig. 1.7(a_2). This arrangement further compacts the DNA and form the 30 nm chromatin fiber. Next level of packing involves formation of looped structures by bringing together different segments of chromatin fiber. This looped structures are compacted to form interphase chromosome. The biggest loops are observed during the metaphase of cell division (Fig. 1.7(b)).

Major factors which facilitate such such loop formation are the DNA binding proteins or their complexes, which attach to multiple DNA segments to bring them into spatial proximity of each other, hence forming the loops [69]. Typical example involves transcription factors which are a class of proteins that bind to DNA at specific locations during the transcription. They are considered as major regulators of gene expression, impacting the 3D organization of chromosomes. They form loops on the chromatin while processing many genes simultaneously. Particular example of franscription factors (TF) is CCCTC-binding factor (CTCF), which binds to specific locations of DNA (three regular repeat of CCCTC sequence). A study showed, two CTCF proteins can bind together to form a dimer [70], which brings the DNA segments attached to them in spatial proximity forming loops [71]. However, study also showes the possibility of CTCF multimer formation. Subsequent experimental study reported evidence of genome-wide localization of CTCF with ring shaped cohesin

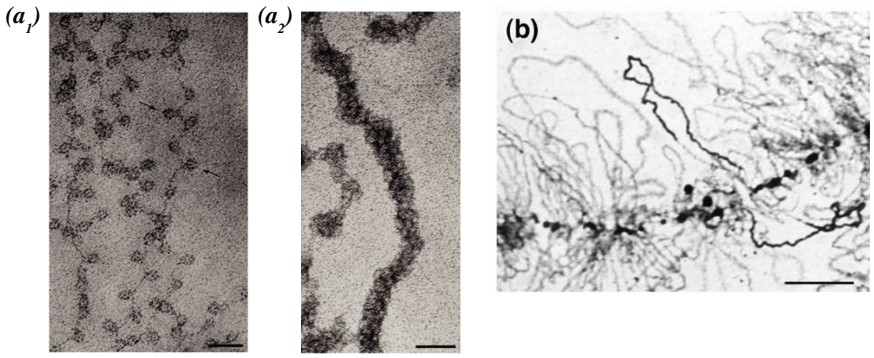


Figure 1.7: (a_1) Electron micrograph showing beads-on-string structure formed by nucleosomes [23]. Scale bar used is 30 nm. (a_2) 30 nm fibre of chromatin [23]. Scale bar used is 50 nm. (b) Part of Lamp brush chromosome X from Triturus carnifex after its attachment to a slide by centrifugation, followed by fixation and silver staining [24]. Scale bar used is 10 μm.

protein complexes. Based on this, loop a extrusion mechanism was proposed, which suggest that the cohesin proteins actively extrude the DNA loop like a motor, until it finds the properly oriented CTCF attached to DNA which works as a stopper [6, 7, 72]. However, the precise mechanism of this active loop extrusion process is yet to be fully established [22, 73].

Various kind of TFs, in addition, form complexes called Transcription factories, e.g., it may consist of RNA polymerase, RNA, transcription factors and a protein core complex as indicated in Fig. 1.8. The RNA polymerase reel the RNA molecule from one DNA strand. Other proteins and complexes like activators, co-activators, which form the protein core complex in transcription factory, provide assistance and regulation in the transcription process. Evidence suggests, a single transcription factory can regulate many genes simultaneously, which leads to co-localization of multiple genes forming chromatin loops as indicated in Fig. 1.8). At a time, the number of transcription factories can vary from few hundred to several thousands in a cell. For example, in human nucleus, cryo-section experiments suggest that the number of transcription factories are about 10000. The size of each transcription factory depends on the amount of RNA reeled, however, the typical size varies between 50-100 nm [74].

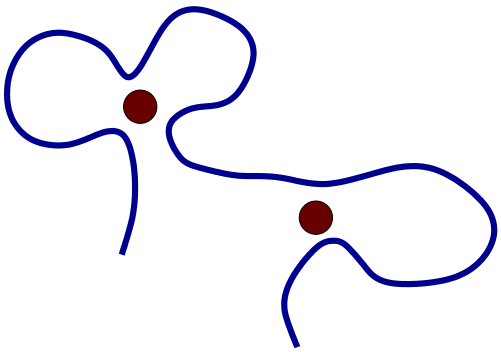


FIGURE 1.8: Schematics for DNA- binding protein or protein- complexes (red bead) cross-linking different chromatin segments (blue line) to form loops.

Experiments show plenty of evidence for chromatin loops. A spectacular example of loops are observed in the lampbrush chromosomes (lbc) of large oocytes in many vertebrates and invertebrates, as shown in Fig. 1.7(b). The open loops are of sister chromatids, they are transcriptionally active, and held together by the compacted and transcriptionally inactive chromomeres. Earlier electron microscopy of lysed human cells (HeLa) also showed evidence of loops [75]. Recently, chromosome conformation capture (CCC) experiments and its Hi-C variants show abundant number of chromatin loops in eukaryotic chromosomes [76–79]. In fact, a particular experimental Hi-C study reported evidence of 10000 chromatin loops in human cells [80].

Chromosome conformation capture (CCC) techniques are set of methods to quantify the contact probability between DNA segments which may be separated contour-wise, however, lie in spatial proximity of each other. There are three major steps: cross-linking the DNA segments, digesting cross-linked DNA and ligation. First, formaldehyde is used to cross-link the DNA segments which introduces permanent bond between those segments. Second, DNA is cut into pieces with the help of restriction enzymes. Size of these DNA fragments determine the resolution of method which are mainly determined by restriction enzymes used in the process. Smaller fragments leads to better resolution of interaction frequency. For example, restriction enzyme EcoR1 cut the DNA at the interval of 4 kbp yielding about

1 million DNA fragments in human nucleus. Third, ligation is performed in presence of T4 ligase, an enzyme that enables the binding of open ends of the cross-linked DNA strands. In this process, ligation between cross-linked DNA fragments are preferred over other DNA sequences. Finally, different CCC technique use this data differently. In 3C, ligated fragments are identified using polymearase chain reaction (PCR) with a known primer. This method provides information about interaction frequency of single DNA segments. Hi-C, on the other hand, uses a combination of high-throughput sequencing and paired end sequencing. In high-throughput sequencing, base-pair sequence of ligated fragments are identified while in paired end sequencing, the sequence belonging to small segments of each fragments are retrieved. The pair of nucleotide sequence is aligned along the entire DNA sequence, hence identifying the DNA sequence belonging to fragments involved in ligation. This allows one to determine all the pairwise interaction between different DNA fragments [16, 81–83]. With this information, a contact map of genomic interaction is created. This experimental technique allows one to find contact maps in all kinds of chromosomes, eukaryotic and prokaryotic.

1.3.2 Protein- DNA association in prokaryotic chromosomes

Unlike eukaryotic chromosomes, the prokaryotic chromosome lacks nucleosomes. Although, they do have several histone homologues, e.g., the histone like nucleoid structuring (HNS) proteins, H, H1, Hu, IHF etc. Bacterial cells have a single circular chromosomal DNA that folds at least 1000 times to fit inside the cell. There are at least five major physical and chemical forces which determine their organization. First, the confinement due to cell envelope bring about a major degree of compaction. However, this is not enough as the nucleoid occupies only 1/4-th of the cell volume, requiring further compaction. A second major compacting mechanism is due to the 5% negative supercoiling of the DNA. The supercoiling is maintained by activities of topoisomerases, transcription, and mucleoid associated proteins. It warps and folds various segments of the chain leading hair-braid like structures known as plectonemes [84]. Formation of plectonemes substantially decreases the effective size of the

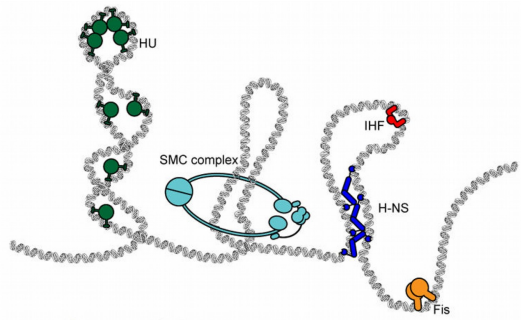


FIGURE 1.9: NAPs and SMC complexes can form loops via different mechanisms. Figure adopted from Ref. [25].

chromosome. Third, simultaneous binding of DNA segments to DNA binding proteins forms bridges and loops. A class of DNA-binding proteins called the nucleoid associated proteins (NAPs) have dominant impact on organization of prokaryotic chromosome. Like histones in eukaryotes, NAPs non-specifically attach to DNA, cross-linking, wrapping and bending DNA segments (Fig.1.9). Impact of local structures on DNA formed due to NAPs, determines the global organization of chromosome at large scales [85]. Examples of NAPs are-histone-homologues H-NS, HU, IHF (integration host factor), FIS (factor for inversion stimulation). Apart from this, the structural maintenance of chromosome (SMC) complexes that are an evolutionarily conserved family of protein complexes, including condensin, cohesin, Smc5/6, and others, are also found in bacteria. They do perform loop extrusion (Fig.1.9) as in eukaryotes [86, 87]. Fifth, the depletion effect due to cytosolic crowders can compress the chromosome to the nucleoid sub-volume. In the rest following, we describe the interaction of DNA with NAPs and SMCs in some further detail.

The H-NS are small proteins found typically in E. coli. A single H-NS can bind to multiple DNA segments, in turn bridging or cross-linking them, which brings distant DNA segments

close to each other spatially. As a result of cross-linking, loop structures emerge and volume of DNA-protein complex (chromosome) decrease. Experiments like single molecule and atomic force microscopy have revealed cross-linking of different DNA segments by H-NS proteins [46].

Another example of NAP is HU. Like H-NS, this is also small protein, of which approximately 30000 units exist in a single bacterial cell. HU can attract DNA segments around itself bending the DNA, showing some similarity to the action of histones proteins in eukaryotes [88, 89]. ChIP-Seq analysis in E. coli chromosome suggests, HU binds to DNA in non-specific manner and due to its large number, it can wrap almost 10% of the DNA [90]. Cells lacking HU produce daughter cells lacking chromosome, suggesting a major role of HU in chromosome segregation [91]. On the other hand, variants of HU with increased DNA affinity over-compact the chromosome [92].

IHF or integration host factor, binds to specific DNA sequences and bends it by 160° [93]. This bend brings DNA segments close to each other, which helps in formation of DNA loops. Similar to IHF, Fis also bends the DNA by $\sim 50 - 90°$ [94]. The non-specific attachment of Fis impacts transcription, replication and recombination.

Collectively, all the different kinds of NAPs and SMCs create DNA loops, which facilitate in chromosome compaction and organization. Evidence of such chromosome loops have been found in earlier electron microscopy (EM) experiments [26, 95–97] (see Fig. 1.10) and from complimentary experiments like 3C, Hi-C, ChIP-seq etc. showing local contact formation and interaction of DNA with proteins [98].

We have already discussed the EM, 3C and Hi-C methods. Before ending this section, here we briefly outline the ChIP-seq method. The ChIP-sequencing (ChIP-seq) method determines protein-DNA interaction. This technique uses combination of chromatin immunoprecipitation (ChIP) and massively parallel DNA sequencing to identify DNA sequences which bind to specific proteins. The technique involves mainly five steps. First, a cross-linking between DNA and associated protein is performed using formaldehyde. Second,

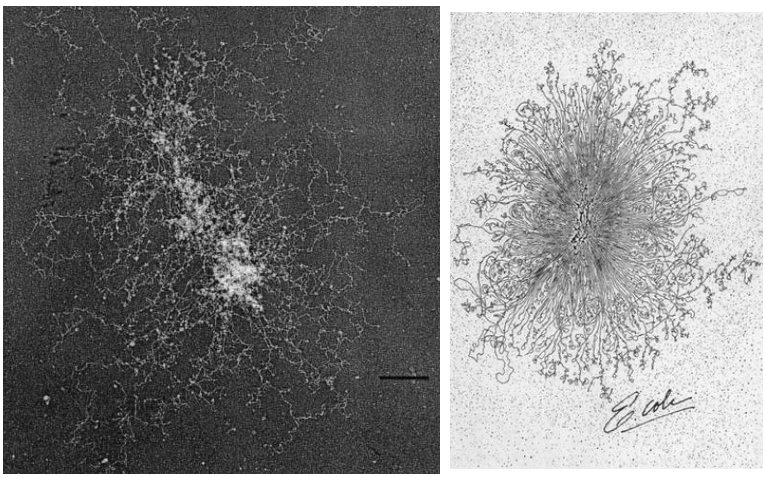


FIGURE 1.10: Electron micrograph shows presence of loops in E. coli chromosome. Scale bar is 500 nm. Figure adopted from Ref. [26]. The figure on the right hand side illustrates the looped structures.

the chromatin is fragmented into DNA pieces below 500 bp. Third, in the ChIP process, the cross-linked DNA-protein complex is augmented using antibody, followed by incubation and centrifugation process. During immunoprecipitation, a removal of non-specific binding sites are performed. In the forth step, cross-links between DNA and proteins are removed and the protein- attached DNA fragments are recovered. Fifth, is the analysis step. Using a massively parallel sequencing technique the gene positions of the protein-attached DNA fragments are identified [99–101].

1.4 Polymer models of DNA and chromosome

Given the polymeric nature of DNA and chromosome, it is only natural to model them using standard polymer physics [53, 102].

Freely jointed chain:

The simplest model of a polymer is an ideal chain with chain length $L = Nb$ of N-bonds each of length b connected as a freely jointed chain such that each bond can be at any

random orientation with respect to its adjacent bond. This ensures that the orientation of consecutive bonds are uncorrelated. The polymer configuration is defined by position vectors $\{\mathbf{r}_n\} = (\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_N)$ or alternatively with bond vectors $\{\mathbf{b}_n\} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N)$, where $\mathbf{b}_n = \mathbf{r}_n - \mathbf{r}_{n-1}$, for all $n = 1, 2, \ldots, N$. The distribution function of the polymer configuration is given by

$$\mathcal{P}(\{\mathbf{b}_n\}) = \prod_{n=1}^{N} \psi(\mathbf{b}_n) \tag{1.1}$$

where the distribution of the randomly oriented bond vector

$$\psi(\mathbf{b}) = \frac{1}{4\pi b^2}\delta(|\mathbf{b}| - b). \tag{1.2}$$

This is independent of other bonds, and normalizes to $\int d\mathbf{b}_n \psi(\mathbf{b}_n) = 1$. The end- to- end vector of the chain is

$$\mathbf{r} = \mathbf{r}_N - \mathbf{r}_0 = \sum_{n=1}^{N} \mathbf{b}_n. \tag{1.3}$$

The random orientation of bond vectors ensure $\langle \mathbf{b}_n \rangle = 0$, and therefore $\langle \mathbf{r} \rangle = 0$. However, $\langle \mathbf{r}^2 \rangle$ has a finite value:

$$\langle \mathbf{r}^2 \rangle = \sum_{n,m} \langle \mathbf{b}_n \cdot \mathbf{b}_m \rangle = Nb^2 \tag{1.4}$$

as the uncorrelated orientations ensure that $\langle \mathbf{b}_n \cdot \mathbf{b}_m \rangle = b^2 \delta_{n,m}$.

Given the probability distribution of the configurations, one can derive the probability

distribution of the end- to- end vector

$$\begin{aligned} P(\mathbf{r},N) &= \int \prod_{n=1}^{N} d\mathbf{b}_n \, \mathcal{P}(\{\mathbf{b}_n\}) \, \delta\left(\mathbf{r} - \sum_{n=1}^{N} \mathbf{b}_n\right) \\ &= \frac{1}{(2\pi)^d} \int d\mathbf{q} \prod_{n=1}^{N} d\mathbf{b}_n \; e^{i\mathbf{q}\cdot(\mathbf{r}-\sum_{n=1}^{N}\mathbf{b}_n)} \prod_{n=1}^{N} \psi(\mathbf{b}_n) \\ &= \frac{1}{(2\pi)^d} \int d\mathbf{q} e^{i\mathbf{q}\cdot\mathbf{r}} \left[\int d\mathbf{b} e^{-i\mathbf{q}\cdot\mathbf{b}} \psi(\mathbf{b})\right]^N \end{aligned} \tag{1.5}$$

In the second step, we used the Fourier representation of the Dirac-delta function and Eq.(1.1). Using Eq.(1.2) in the last step the integration over the bond vector yields the result $[\sin(qb)/qb]$. In the large N limit, the integrand $[\sin(qb)/qb]^N$ contributes if $qb \ll 1$, where it can be approximated to $[\sin(qb)/qb]^N \approx (1 - q^2b^2/6)^N \approx \exp(-Nq^2b^2/6)$. As a result one finds a Gaussian integral in the last step of Eq.(1.5). Performing this integration, the end- to- end distribution function in three dimensions

$$P(\mathbf{r},N) = \left(\frac{3}{2\pi N b^2}\right)^{3/2} \exp\left(-\frac{3\mathbf{r}^2}{2Nb^2}\right). \tag{1.6}$$

This derivation is a special case of the more general central limit theorem [102]. The distribution agrees with the exact results $\langle \mathbf{r} \rangle = 0$ and $\langle \mathbf{r}^2 \rangle = Nb^2$, but is strictly valid only for small extensions $|\mathbf{r}| \ll Nb$. It is interesting to note that the above distribution function suggests a free energy $\beta F(\mathbf{r},N) = -\ln P(\mathbf{r},N) \sim 3\mathbf{r}^2/2Nb^2$ where $\beta = 1/k_BT$.

Gaussian Polymer:

Relaxing the bond-length constraint, one can use a simpler Gaussian polymer model with bond vectors obeying the distribution

$$\psi(\mathbf{b}) = \left(\frac{3}{2\pi b^2}\right)^{3/2} \exp\left(-\frac{3\mathbf{b}^2}{b^2}\right) \tag{1.7}$$

such that $\langle \mathbf{b} \rangle = 0$ and $\langle \mathbf{b}^2 \rangle = b^2$. The probability distribution function for the chain conformations is given by,

$$\mathcal{P}(\{\mathbf{b}_n\}) = \prod_{n=1}^{N} \psi(\mathbf{b}_n) = \left(\frac{3}{2\pi b^2}\right)^{3N/2} \exp\left(-\sum_{n=1}^{N} \frac{3(\mathbf{r}_n - \mathbf{r}_{n-1})^2}{2b^2}\right) \tag{1.8}$$

This conformational distribution gives the same end- to- end distribution shown in Eq.(1.6). In the continuum limit, the polymer is described by a space curve $\mathbf{r}(s)$. Replacing the expression $(\mathbf{r}_n - \mathbf{r}_{n-1})/b$ by the partial derivative $\partial\mathbf{r}/\partial s$, the distribution function of polymer conformations can be expressed as

$$\mathcal{P}[\mathbf{r}(s)] = \text{const} \times \exp\left[-\frac{3}{2b}\int_0^L ds \left(\frac{\partial \mathbf{r}}{\partial s}\right)^2\right] \tag{1.9}$$

A mechanical model of beads connected by springs described by the Hamiltonian

$$H = \frac{3\, k_B T}{2b^2} \sum_{n=1}^{N} (\mathbf{r}_n - \mathbf{r}_{n-1})^2 \tag{1.10}$$

obeys the Boltzmann distribution given in Eq.(1.8). Similarly, the Hamiltonian in the continuum limit

$$\beta H = \frac{3}{2b}\int_0^L ds \left(\frac{\partial \mathbf{r}}{\partial s}\right)^2 \tag{1.11}$$

leads to the distribution in Eq.(1.9). In writing the above expression we used $\beta = 1/k_B T$.

Self- avoiding polymer:

Unlike the ideal chain, real chains are self avoiding polymers in which different segments do not cross each other. The first scaling theory describing its behavior is due to Flory. Within

this scaling theory the free energy of the self avoiding chain can be written down as

$$\beta F \sim \frac{r^2}{Nb^2} + b^d \frac{N^2}{r^d} \tag{1.12}$$

where r denotes the end- to- end separation. The first term accounts for entropic elasticity of the ideal chain, and the second term is due to the local repulsion between different segments. Minimizing the free energy with respect to r, one obtains the Flory scaling $r \sim bN^{\nu}$ where $\nu = 3/(d+2)$.

The Hamiltonian of an excluded volume chain was proposed by Edwards [103],

$$\beta H = \frac{3}{2b} \int_0^L ds \left(\frac{\partial \mathbf{r}}{\partial s} \right)^2 + \frac{v_0}{2} \int_0^L ds \int_0^L ds' \delta \left[\mathbf{r}(s) - \mathbf{r}(s') \right], \tag{1.13}$$

where the first term is the same as in ideal chain, and v_0 in the second term sets the strength of repulsion between different segments of the polymer. Edwards and Singh [103] proposed a self-consistent field theory approach to determine the size of the self-avoiding polymer. The result of this calculation agreed with Flory scaling.

In three dimensions, $\nu = 3/5$ leads to a substantial swelling of chain compared to the ideal polymer. Various polymeric models for chromosome has used a self-avoiding chain model for the long DNA or chromatin [104–106].

1.4.1 Semi-flexible polymer:

To incorporate the bending rigidity of DNA or chromatin, for chains with intermediate length scale, semiflexible polymers are considered [105]. The most commonly used model for semiflexible polymers is the worm- like- chain (WLC) model. Within this model the inextensible polymer is defined as a space curve $\mathbf{r}(s)$ such that the local tangent $\mathbf{u}(s) = \partial \mathbf{r}/\partial s$ is a unit vector $\mathbf{u}^2(s) = 1$. The change in orientation $\mathbf{u}(s)$ along the polymer contour costs

energy [107, 108],

$$\beta H = \frac{\kappa}{2} \int_0^L \left(\frac{\partial \mathbf{u}(s)}{\partial s} \right)^2 ds \tag{1.14}$$

where κ has the dimension of length and plays the role of bending rigidity. The tangent-tangent correlation function shows an exponential decay $\langle \mathbf{u}(s) \cdot \mathbf{u}(s') \rangle = e^{-|s-s'|/l_p}$, where $l_p = \frac{2\kappa}{d-1}$ represents the persistent length of the chain. In this expression d stands for the embedding dimension of the polymer.

DNA is a negatively charged polymer. In ambient condition it shows a persistence length $l_p \approx 50\,\text{nm}$ [109, 110]. The persistence length changes with changing concentration of counter ions in the medium [111, 112]. At the smallest length scales, DNA behaves like a stiff polymer and the local bending rigidity depends on the base pair sequence [113]. For chain lengths comparable to l_p, the bending rigidity dominates the statistics and mechanical properties. Force- extension measurements on DNA show remarkable agreement with the WLC model prediction [114–116]. The persistence length of the DNA used in the fit these results is around 53 nm [109]. At largest length scales, the DNA shows behavior similar to flexible chain [117–119].

Inside the cell, DNA interacts with proteins to form structures. For example, formation of nucleosome leads to beads-on-string structure which further packs to form 30 nm fiber via help of other protein. The complexities like constraints, crowding, aggregation of proteins along the DNA contour can potentially impact the persistence length of chain. High resolution in-situ fluorescence microscopy of budding yeast chromatin suggests, the 30 nm fibre has persistence length in the range of 170-220 nm with mass density $\sim$ 110-150 bp/nm. This leads to a packing of 7-10 nucleosomes per 11 nm turn in the fiber [120]. Simulation studies of chromatin in certain contexts used semi-flexible polymer models [106, 121]. However, longer chains $L \gg l_p$ are often considered as flexible chains of Kuhn segments $b = 2l_p$ [122].

Considering the DNA double helix as a fiber of diameter $\sigma = 2.1\,\text{nm}$, the separation

between the two strands of DNA, each bead of the polymer will contain only $\sigma/0.34 \approx 6.2$ bp of DNA. Within this picture, a persistence length of 50 nm translates to about $25\,\sigma$. Modeling an E. coli chromosome of length $L = 4.6$ Mbp would already require a large number of degrees of freedom containing $N = L/\sigma = 7.5 \times 10^5$ beads. The coarse-grained models avoid this, e.g., by modeling a chromatin fibre as a self-avoiding flexible chain consisting of beads that stand for 10-12 closely packed nucleosomes containing 2-2.5 kbp DNA segments having a size 20-40 nm [17]. In this level of coarse-graining a 500 bead polymer stands for ~ 1 Mbp chain. The interaction between DNA and proteins can be considered by explicit consideration of proteins, or using effective protein- mediated interaction between DNA segments [104–106]. We explore both these approaches in the first part of this thesis. In order to study the impact of confinement and molecular crowding on the overall organization of a chromosome, in the second part of the thesis, we consider a further coarse-grained feather-boa model of chromosome incorporating the chromosomal loop structures [26, 80], the details of which will be presented later.

1.5 Organization of the thesis

In chapter 2, we consider a self- avoiding polymer model for chromatin in the presence of DNA- binding proteins. The binding proteins repel each other, however has an attractive interaction with polymer beads. As a result, they (un-)bind (from) to the chain following Boltzmann rates that depend on the short ranges interaction. We perform molecular dynamics simulation to study the chromosome organization due to such binding proteins. Changing the density of binding proteins lead to continuous coil-globule phase transition of the model chromatin. The transition is described in terms a mean field model. The slow time scales for chromosome organization is captured in terms of a linear stability analysis. Along the transition, we analyze the morphological changes in terms of the formation of chromatin loops, changes in inter-segment contacts and sub- chain extensions.

The cross-linking of binder proteins with different segments of the chain can be viewed as an effective attraction between chain segments. In chapter 3, we use self-avoiding polymer model with a short ranged attractive tail in the interaction between chain segments. This reproduced most of the observed phenomenology of folded chromosome. Using a hetero-polymer variant of the model we demonstrate the formation of typical checker-board patterns in the chromosome contact map. These two chapters show formation of loops as a mode of chromosome compaction.

In what follows in this thesis, we focus on the organization of the whole chromosome using a feather-boa model, which incorporates chromosomal loops in terms of side-loops attached to a polymeric backbone, in a manner similar to the bottle-brush polymers. We use the model, particularly, to study chromosomal organization in rod-like bacteria, e.g., E. coli. In chapter 4, we introduce this model. Subjected to a cylindrical confinement the feather-boa polymer show spontaneous helicity. In this chapter we show how such organization depends on the relative size of the side loops with respect to the degree of confinement.

In the next two chapters we study the impact of cytosolic molecular crowders on the confined chromosome. These crowders come in various size and density. In Chapter 5, we vary the size of crowders. The smallest crowders can diffuse through the system almost freely, getting distributed uniformly. On the other hand, with increase in size crowders get segregated from the chromosome, first in the radial direction, and finally in the longitudinal direction. For intermediate crowder sizes, the system shows a complementary helical organization of chromosome and crowders. We study this problem in quantitative detail with the help of a further coarse-graining of the chromosome model.

In chapter 6, we present a study of the impact of changing crowder density. Changing the density of smallest crowders, that can penetrate the inter-filament gaps in the volume occupied by the chromosome, does not impact the organization significantly. To capture the polymeric nature of some of the larger crowders, we model them using effective Gaussian core repulsion. Changing their density generates spatial segregation of the feather- boa

chain from the crowders. The radial segregation is observed at a lower density, followed by longitudinal segregation at high densities. The size of the chromosome as a function of crowder density is obtained using a mean field argument. The molecular crowding impacts the detail chromosomal organization, by changing its helicity analyzed in terms of winding number, turn number and number of kinks.

Finally we conclude in Chapter 7.

2 Chromosomal compaction due to cross linkers: change in morphology

As it has been discussed in the previous chapter, two major components of chromosome are DNA and its associated proteins. The length of DNA varies between millimeters in prokaryotes to meters in the eukaryotic cells. A typical example of the former is E. coli bacteria, with a 1.6 mm (4.6 Mbp) long circular DNA packed inside a small cylindrical cell of diameter $0.8\,\mu$m and length $3 - 6\,\mu$m. An example of the latter is a human cell, with 46 DNAs of around 2 meters length packed inside nucleus of diameter $10 - 20\,\mu$m. The DNA has to fold at least 10^3 times to fit inside such strong confinements. This enormous compaction needs to allow efficient information processing in terms of gene expression and regulation. The DNA associated proteins plays crucial role in such organization [1, 2, 14].

At the small length scales, eukaryotic DNA wraps around histone octomers forming a "beads on a string" chromatin structure with connected set of nucleosomes [68]. The bacterial cells have histone like nucleoid structuring (H-NS) proteins. Their dimers bind to DNA nonspecifically to generate organizations at the smallest scale [46, 123]. The non-specific binding is often mediated by the negative charge on DNA and the positive charges on these proteins [1, 2]. The role of such passive nucleoid associated proteins (NAP) have been discussed in some detail in the previous chapter.

The next higher-order structures are formed when two or more distant segments of the chromatin fiber are brought in the spatial proximity of each other to form loops [80, 124–128]. These are observed in all domains of life, in bacteria [129, 130], archea [131] and eukaryotic cells [80, 132–134]. Such loops are often stabilized by protein cross-linkers [4,

5, 135, 136]. Some earlier numerical and theoretical studies used explicit binders [104–106, 121, 137, 138] or effective attraction between chromatin segments [139–141] to study the resultant compaction and its impact. Apart from this, active extrusion of loops, e.g., with the help of cohesin and CTCF have also been observed [6, 7, 80, 133, 134, 142, 143]. Chromosomal loops were directly observed in electron microscopy [26, 95–97], and through chromosome conformation capture experiments [16, 144, 145]. It is often observed that 3-D spatial structure of the chromosome determines the biological function of the chromosome. For example, contour- wise distant genes on the DNA contour are found to be in spatial proximity of each other, which are all regulated by the same transcription factory [85, 146–150].

In this chapter we use a homo-polymer model for chromatin fiber along with diffusible attractive binders to extract some of the generic features of such protein mediated compaction. The resultant effective attraction between segments is expected to lead to a coil-globule transition [53]. Earlier theories suggested first order or second order phase transition depending on parameter values [151–155]. A relatively recent numerical simulation and mean field theory of semi-flexible chain in the presence of binding proteins showed first order transition for stiff polymers and continuous transition for flexible polymers [105]. Another mean field study including fluctuations of co-solvent density, suggested that the nature of the coil-globule transition depends on polymer- co-solvent interaction. An attractive interaction between the polymer and co-solvent leads to a first order transition, while a repulsive interaction between the two leads to a continuous transition [156].

In this chapter, we consider a self-avoiding polymer model for the chromatin fiber. The binding proteins are modeled by explicit cross-linkers diffusing in the embedding 3-D space. They have local aspecific attractive interaction with all the segments of the chromatin fiber. Otherwise, the cross-linkers repel themselves with short range interaction. We observe, with the increase in density of cross-linkers, the polymer collapses. In this regard, the first question we ask is, what is the nature of the folding transition of polymer? Across the transition, we

investigate how the local morphology of the polymer changes. We have quantified this in terms of formation of loops, inter-segment contacts, and zippers. In addition, we also focus on the behavior of cross-linker clusters.

The details of the model and numerical simulations are presented in section 2.1. The simulation results regarding the coil-globule transition of the model chromatin along with the clustering of chromatin-bound cross-linkers are discussed in section 2.2. The transition and clustering are interpreted in terms of a mean field theory is presented in Sec. 2.2.6. In this same section, the emergent time-scales associated with filament relaxation is interpreted using a linear stability analysis. In Sec. 2.3, we characterize the local morphological changes across the transition in terms of contacts, loops, and zippers. Finally, we conclude in Sec. 2.5, presenting a discussion on experimentally verifiable predictions in terms of loop structures and relaxation time scales.

2.1 Model

We use a self-avoiding flexible chain model of chromatin. The bead size is assumed to be larger than the Kuhn length, twice the effective persistent length of a persistent chain [122]. The chain connectivity is maintained by finitely extensible nonlinear elastic (FENE) bonds between consecutive beads,,

$$U_{\mathrm{FENE}}(r_{i+1,i}) = -\frac{k}{2}R^2 \ln[1 - (r_{i+1,i}/R)^2]. \tag{2.1}$$

Here, k and R fix the bond, and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ denotes separation between i-th and j-th bead. The self avoidance between monomers is implimented by Weeks-Chandler-Anderson potential [157],

$$\begin{aligned} U_{\mathrm{WCA}}(r_{ij}) &= 4\epsilon[(\sigma/r_{ij})^{12} - (\sigma/r_{ij})^{6} + 0.25], \text{ for } r_{ij} < 2^{1/6}\sigma \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{2.2}$$

Thus the chain is defined by, [158],

$$U = [U_{WCA} + U_{FENE}]. \tag{2.3}$$

The potential U is shown in Fig. 2.1. The repulsion between cross-linkers are modeled through the same U_{WCA} interaction. The energy and length scales are set by ϵ and σ respectively. The FENE potential is set by $k = 30.0\,\epsilon/\sigma^2$, $R = 1.6\,\sigma$.

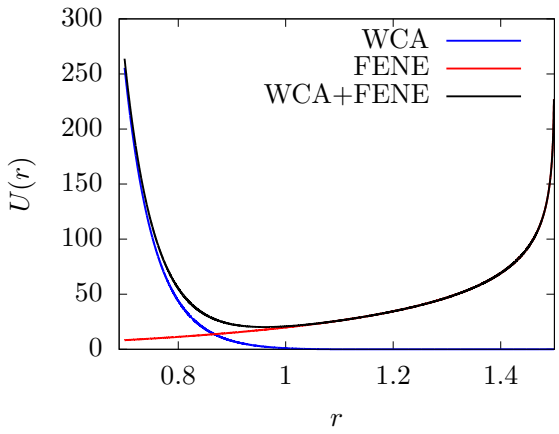


FIGURE 2.1: The WCA potential U_{WCA} (see Eq. 2.2), and the FENE potential U_{FENE} (see Eq. 2.1) are plotted as a function of separation r, and shown in blue and red solid line respectively. Together they define the self- avoiding chain (see Eq. 2.3) shown in black solid line. Parameter values used are $\epsilon = 1$, $\sigma = 1$ for WCA, and $k = 30.0\,\epsilon/\sigma^2$, $R = 1.6\,\sigma$ for FENE.

The interaction between cross-linkers and monomers is modeled through a truncated and shifted Lennard-Jones potential,

$$\begin{aligned} U_{\text{shift}}(r) &= U_{\text{LJ}}(r) - U_{\text{LJ}}(r_c) \qquad for \;\; r < r_c \\ &= 0 \qquad otherwise \end{aligned} \tag{2.4}$$

where $U_{LJ}(r) = 4\epsilon_m[(\sigma/r)^{12} - (\sigma/r)^6]$.

We set $\epsilon_m = 3.5\,\epsilon$, and $r_c = 1.5\,\sigma$. The potential is plotted in Fig. 2.2. The choice of ϵ_m is stronger than the typical hydrogen bonds ($1.2\,k_BT$) and provides better stability [159], e.g, as for transcription factors [104]. However, allows equilibration through attachment-detachment kinematics over the simulation time scales. The bond between a cross-linker and a monomer is formed if they come within the range of attraction $r_c = 1.5\sigma$. A single cross-linker may bind to multiple monomers, capturing the presence of multiple DNA binding domains in a number of regulatory proteins [104]..

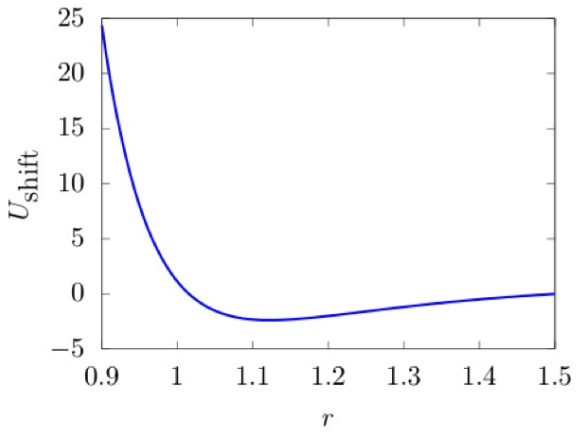


Figure 2.2: Shifted and truncated Lennard-Jones potential $U_{\text{shift}}(r)$ is plotted as a function of separation r. The cut-off range $r_c = 1.5\sigma$ and depth of potential $\epsilon_m = 3.5\ k_BT$ is chosen.

The molecular dynamics simulations are performed using the standard velocity-Verlet algorithm [160] using time step $\delta t = 0.01\tau$, where $\tau = \sigma\sqrt{m/\epsilon}$ is the characteristic time scale. The mass of the particles are chosen to be m = 1. The temperature of the system

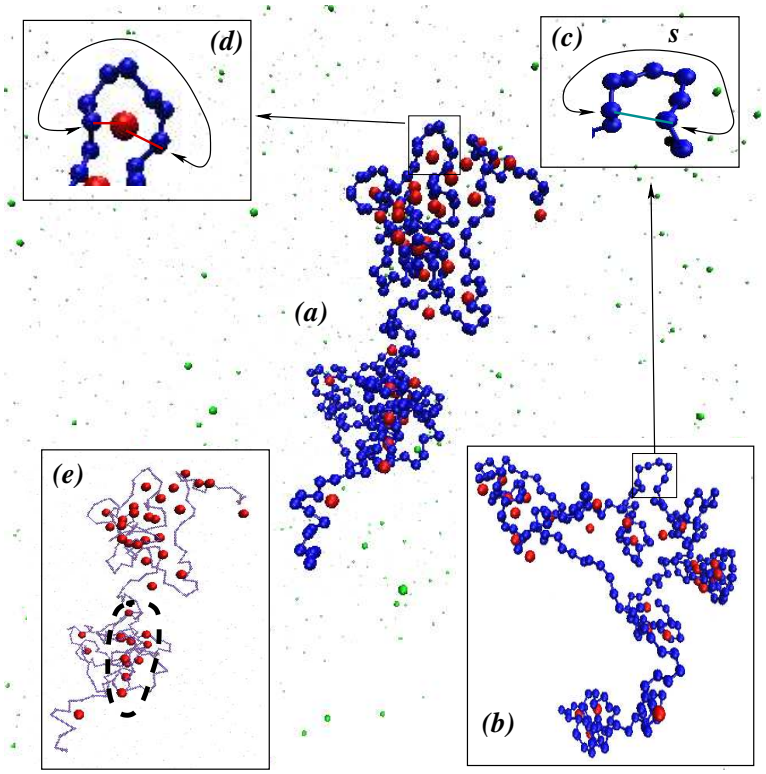


FIGURE 2.3: Configurations of the model chromosome and cross-linkers at the transition point $\phi_c = 1.57 \times 10^{-3}$, where large structural fluctuations are observed. (a) A relatively compact configuration. The chromatin filament is shown by blue monomers connected by bonds. The cross-linkers attached to the chromatin are shown as red beads, while the freely diffusing cross-linkers are shown as green beads. (b) A relatively open configuration of the model chromosome. (c) A magnified portion of (b) showing a contact formation denoted by the aqua-green bar. Here two monomers separated by a contour length s have come within the cutoff separation r_c forming contact. (d) A magnified portion of the chain in (a) shows loop formation by a polymer bound cross-linker (red bead). The red bars from the red bead indicate the bonds that it forms with the chromatin segments lying within the cutoff separation r_c. The line with arrows identifies a simply connected loop (for further details see Sec. 2.3.1). (e) Clusters of polymer bound cross-linkers as seen in the configuration (a). For better visibility of cross-linkers, the chromatin is shown here as a transparent chain. The thick dashed circle identifies one cluster of cross-linkers.

is kept constant at $T = 1.0\,\epsilon/k_B$ by using a Langevin thermostat [157] characterized by an isotropic friction constant $\gamma = 1/\tau$, as implemented by the ESPResSo molecular dynamics package [161]. Similar methods have been successfully used earlier in simulation of polymers in various contexts [162]. Note that the diffusion of a single bead over its size σ takes a time $\gamma\sigma^2/k_BT$, which is the same as the characteristic time τ .

Unless stated otherwise, in this chapter, we consider a $N = 256$ bead chain. Its typical size in absence of binders is given by the radius of gyration $R_g^0 = (13.02 \pm 2.65)\,\sigma$. The largest fluctuations in its end to end separation are restricted within $80\,\sigma$. To avoid any possible boundary effect, we perform simulations in a cubic volume of significantly bigger size with sides of $L = 200\,\sigma$, and implement periodic boundary condition. We vary the total number of cross-linkers from $N_c = 0$ to 6000 that changes the dimensionless cross-linker density from $\phi_c = \frac{4}{3}\pi\sigma^3 N_c/L^3 = 0$ to $\pi \times 10^{-3}$. The approach to equilibrium is followed over $10^6\,\tau$, longer than the longest time taken for equilibration near the transition point. The analyses are performed over further runs of $10^6 - 10^7\,\tau$. We also represent the cross- linker density by a short notation $\phi_r = \phi_c \times 10^3$. The system size dependence is studied using a restricted set of simulations, as simulating longer chains requires longer equilibration, larger simulation box and larger number of cross-linkers, increasing the simulation time significantly.

A couple of representative equilibrium configurations are shown using VMD [163] in Fig.2.3 illustrating polymer contacts, loop formation, and clustering of cross-linkers.

2.2 Results

2.2.1 The valency of cross-linkers

Fig. 2.3 (a) and (b) clearly shows the cross-linkers used in the simulations are potentially multivalent. Here the question we ask is how many monomers of the chain, a cross-linker attaches to simultaneously? From the simulations, we identify the polymer bound cross-linkers, count the number of monomers that lie within the range of attraction $r_c = 1.5\,\sigma$ identifying the instantaneous valency of a cross-linker, and compute the histogram over all the cross-linkers and time. This leads to the probability Π_v of valency v, normalized to $\sum_v \Pi_v = 1$. The maximum of the probability indicates the typical valency of cross-linkers at an ambient density ϕ_c (see Fig.2.4). At $\phi_c = 0.26 \times 10^{-3}(\triangle), 1.57 \times 10^{-3}(\square)$ and $3.14 \times 10^{-3}(\circ)$, the peak occurs at $4, 5$ and 6 respectively. Thus, as ϕ_c increases, the peak

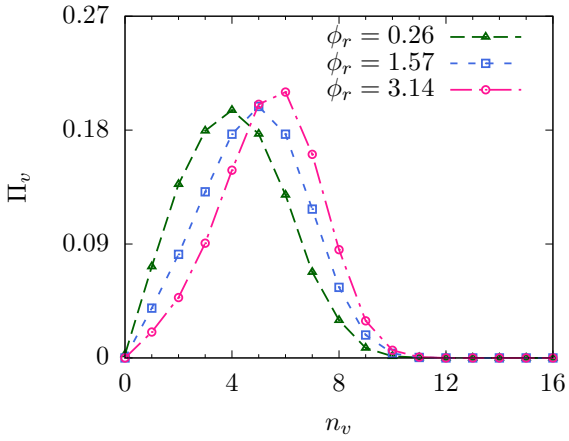


FIGURE 2.4: At the cross-linker density $\phi_r = \phi_c \times 10^3$, Π_v represents probability that a single cross-linker binds to n_v monomers of the polymer simultaneously.

shifts towards larger values. It means with increase of ϕ_c, a single cross-linker on an average binds to more number of monomers of the chain. In the fully compact state, at the largest ϕ_c, the typical valency we find is 6.

2.2.2 The coil-globule transition

The passive binders diffuse in three dimensions and attach to polymer segments following the Boltzmann weight. They are multi-valent, typically cross-linking multiple polymer segments. The probability of number of chromatin segments that a binder can cross-link simultaneously shows a maximum that increases from 4 to 6 as the average binder concentration is increased (see Fig.-2.4). This range overlaps with the typical multiplicity of binding factors like CTCF and transcription factories [104].

As different polymer segments start attaching to a cross-linker the local density of monomers increases, generating a positive feedback recruiting more cross-linkers and as a result localizing more monomers. Such a potentially runaway process gets stabilized, within our model, due to the inter-binder repulsion that ensures the binder-clusters are spatially extended. These clusters are identified using the clustering algorithm in Ref. [164], and the

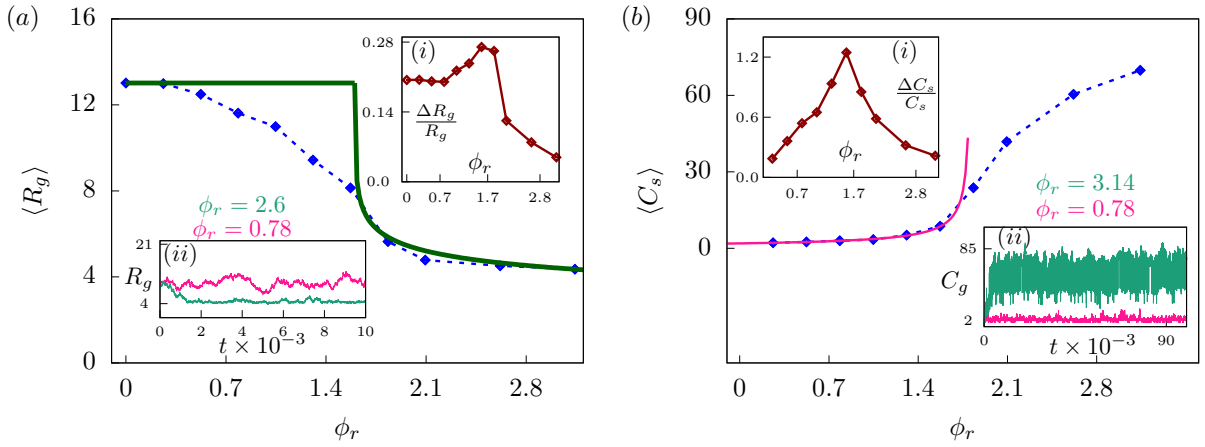


FIGURE 2.5: The coil-globule transition as a function of ambient cross-linker density ϕ_c, expressed in terms of $\phi_r = \phi_c \times 10^3$. (a) The decease in mean radius of gyration of polymer $\langle R_g \rangle$ with ϕ_c, the data are shown by $\diamond$ with standard errors, captures a coil-globule transition. The blue dashed line is a guide to eye. The mean field description predicts $\langle R_g \rangle$ = constant = $\langle R_g \rangle(\phi_c = 0)$ at $\phi < \phi_c^* = 1.57 \times 10^{-3}$. At higher densities, $\phi_c > \phi_c^*$, simulation data for $\langle R_g \rangle$ fits well with Eq.2.8 with fitting parameter $u/v = 0.1$. The two curves are shown by green solid lines. At the transition point ϕ_c^*, relative fluctuation of polymer size $\Delta R_g/R_g$ shows a maximum (inset (i)). The equilibration of R_g with time t at two densities $\phi_c = 0.78 \times 10^{-3}$, 2.6×10^{-3} are shown in inset (ii). (b) Mean size of the polymer bound cluster of cross-linkers $\langle C_s \rangle$ increases with ϕ_c, the data are shown by $\diamond$ with standard errors. The blue dashed line is a guide to eye. A fitting of the increase in $\langle C_s \rangle$ at $\phi_c < \phi_c^*$ with Eq.2.11 is shown by the pink solid line. The fitting parameter $\mathcal{A} = 1.9$. The relative cluster size fluctuation $\Delta C_s/C_s$ shows a sharp maximum at the transition point ϕ_c^* (inset (i)). Inset (ii) shows how the instantaneous mean cluster size C_g equilibrates with time t at two cross-linker densities $\phi_c = 0.78 \times 10^{-3}$, $\pi \times 10^{-3}$.

cluster-size is given by the total number of binders in a cluster. Concomitant with such clustering, the polymer gets folded undergoing a coil-globule transition.

In Fig. 2.5(a), we have presented the transition in terms of decrease in mean radius of gyration $\langle R_g \rangle$ of polymer with an increase in cross- linker density ϕ_c in the simulation box. The data points $(\diamond)$, plotted with their respective standard error, are averaged over 10^6 equilibrium configurations. Assuming x_1 represents the mean measured over x_2 number of data points, the corresponding standard error would be $x_1/\sqrt{x_2}$. In Fig. 2.5(a), the standard error of data points are smaller than symbol size due to large number of measurements. The solid (green) line shows the mean field prediction that we present in Sec. 2.2.6. The

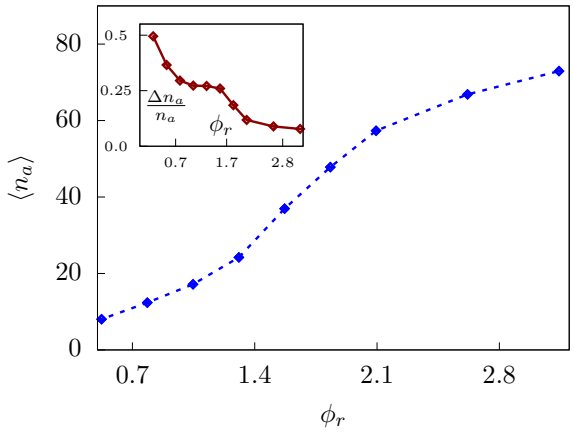


FIGURE 2.6: The increase in mean number of polymer-bound cross-linkers $\langle n_a \rangle$ with $\phi_r = \phi_c \times 10^3$. The data points shown in $\diamond$ are plotted with their respective standard errors and blue dashed line is guide to eye. Inset shows relative fluctuation of n_a as a function of ϕ_r.

transition point, $\phi_c^* = 1.57 \times 10^{-3}$, is characterized by a maximum in relative fluctuations of the polymer size $\Delta R_g/R_g = \sqrt{\langle R_g^2 \rangle - \langle R_g \rangle^2}/\langle R_g \rangle$, shown in the inset (i) of Fig. 2.5(a). The equilibrations of R_g at two representative binder concentrations ϕ_c are illustrated in the inset (ii). As we show in Fig. 2.12, the relative fluctuations $\Delta R_g/R_g$ near phase transition increases with polymer size N, suggesting divergence in the thermodynamic limit, a characteristic of continuous phase transitions. In Fig.2.3, the large fluctuations at the phase transition point are further illustrated with the help of two representative conformations: a relatively compact conformation in Fig.2.3(a), and a more open conformation in Fig.2.3(b).

The coil-globule transition occurs concomitantly with the formation of polymer-bound cross-linker clusters. At a given instant, several disjoined clusters may form along the model chromatin (see Fig.2.3(e)). The cluster size $\langle C_s \rangle$ is the average number of binders constituting the clusters. It grows significantly as ϕ_c approaches phase transition from below (Fig. 2.5(b)). The linear stability estimate of cluster size, as discussed in Sec. 2.2.6, is represented by the (pink) solid line in Fig. 2.5(b). The relative fluctuations in cluster size $\Delta C_s/C_s = \sqrt{\langle C_s^2 \rangle - \langle C_s \rangle^2}/\langle C_s \rangle$ show a sharp maximum at the phase transition point ϕ_c^* (inset (i)). Equilibration of the mean cluster size C_g, instantaneous average over all

clusters, at two ϕ_c values are shown in the inset(ii). Here we should reemphasize that the stable equilibrium cluster sizes that we find at all cross-linker densities, are maintained by the competition between monomer- cross-linker attraction, and the short ranged repulsion between cross-linkers limiting the upper bound for cluster density.

The average number of cross-linkers attached to chain $\langle n_a \rangle$ are plotted as a function of ϕ_c in Fig. 2.6. With increase in ϕ_c, $\langle n_a \rangle$ grows to finally saturate at largest cross- linker density. Around the transition point, ϕ_c^*, a sharp increase in $\langle n_a \rangle$ is observed. The inset shows relative fluctuation of number of polymer-bound cross-linkers, $\Delta n_a/n_a = \sqrt{\langle n_a^2 \rangle - \langle n_a \rangle^2}/\langle n_a \rangle$ as a function of ϕ_c. $\Delta n_a/n_a$ decreases with increase in ϕ_c, however shows a plateau around transition point $\phi_c^* = 1.57 \times 10^{-3}$. Note that $\langle n_a \rangle$ is always greater than $\langle C_s \rangle$, because all the polymer-bound cross-linkers may not result into a cluster.

To characterize the nature of the coil-globule transition, we study the probability distribution of the equilibrium radius of gyration $P(R_g)$ across the transition. As is shown in Fig. 2.7(a), the distribution remains unimodal at all values of cross-linker densities. This clearly displays absence of metastable phase on the other side of the transition, characteristic of the continuous transition. Moreover, the distribution has the largest variance at the transition point $\phi_c^* = 1.57 \times 10^{-3}$. Note that this observation is in contrast to mean field prediction of Ref. [156], while is in agreement with the numerical simulations in Ref. [105].

We also computed the probability distribution function of magnitude of end-to-end vector r. In Fig.2.7(b), the probability distribution is plotted by the quantity $P(r/R_F)$, where $R_F = \langle r^2 \rangle^{1/2} = \sqrt{\langle r \rangle^2 + \Sigma^2}$ denote the Flory radius of chain. Here, Σ is the standard deviation. Across the transition, $P(r/R_F)$ show unimodal distribution.

Further, the analytic form of the distribution function can be written as, $p_N(r) = \frac{1}{R_F^d} f_p(\frac{r}{R_F})$ for $\sigma << r << N\sigma$, with normalization $\int p_N(\mathbf{r}) d\mathbf{r} = 1$. Assuming, $x = \frac{r}{R_F}$, the function $f_p(x)$ in two opposite limits behave as, $\lim_{x\to 0} f_p(x) = \exp(-x^\delta) f_1(x)$ and $\lim_{x\to\infty} f_p(x) =$ constant x^g. One can capture the behavior of $f_p(x)$ in these two limits into a single function $\tilde{p}(x) = a x^g \exp(-x^\delta)$. For the self avoiding chain, the scaling analysis suggests, $g = \frac{\gamma-1}{\nu}$ and

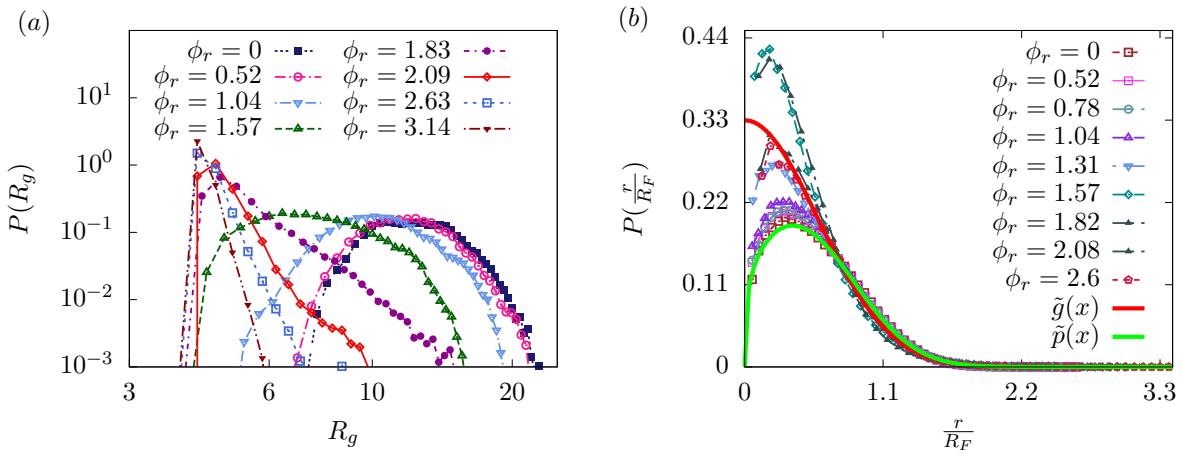


FIGURE 2.7: Probability distributions of the polymer radius of gyration, $P(R_g)$ in (a), and end-to-end separation $P(\frac{r}{R_F})$ in (b) are plotted at different $\phi_r = \phi_c \times 10^3$. Scale factor $R_F = \sqrt{\langle r^2 \rangle}$ is the Flory radius. Function $\tilde{p}(x) = 0.28x^{0.28} \exp(-1.2x^{2.43})$ and $\tilde{g}(x) = (3/2\pi)^{\frac{3}{2}} \exp(-\frac{3}{2}x^2)$, where $x = \frac{r}{R_F}$.

$\delta = (1-\nu)^{-1}$, that in 3-dimension leads to the value $\delta = \frac{5}{2}$ and $g = \frac{5}{18}$ [53]. The distribution function from our simulation $P(x)$ at $\phi_c = 0$ fits remarkably well with $\tilde{p}(x)$ that leads to $\delta = 2.43$ and $g = 0.28$, nearly capturing the mean field behavior of self avoiding polymer (green solid line in Fig. 2.7(b)). In addition, a function $\tilde{g}(x) = (3/2\pi)^{\frac{3}{2}} \exp(-\frac{3}{2}x^2)$ agrees with the simulation data in globule phase at $\phi_c = 2.6 \times 10^{-3}$ (red solid line in Fig. 2.7(b)).

The distribution of cross-linker cluster sizes $P(C_s)$, on the other hand, shows clear bimodality in much of the ϕ_c range scanned across the transition (Fig.2.8), capturing coexistence of clusters of small and large sizes. However, such clusters have similar densities and do not suggest phase coexistence. In fact, as we show in Sec.2.2.6, the assumption of constant particle density within clusters provides a good description of the growth of mean cluster size through Eq.(2.11) (see Fig.2.5(b)).

2.2.3 Dynamical analysis

The equilibrium dynamics in our system is characterized by following the evolution of the radius of gyration $R_g(t)$ and the total number of chromatin-bound cross-linkers $n_a(t)$. This

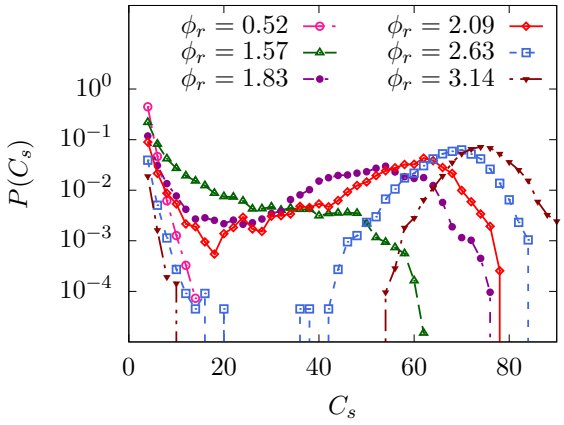


Figure 2.8: Probability distributions of the polymer bound cluster size of cross-linkers, $P(C_s)$, are plotted at different cross-linker densities $\phi_r = \phi_c \times 10^3$. The coexistence of small and large clusters are observed at densities $\phi_c \geq 1.83 \times 10^{-3}$.

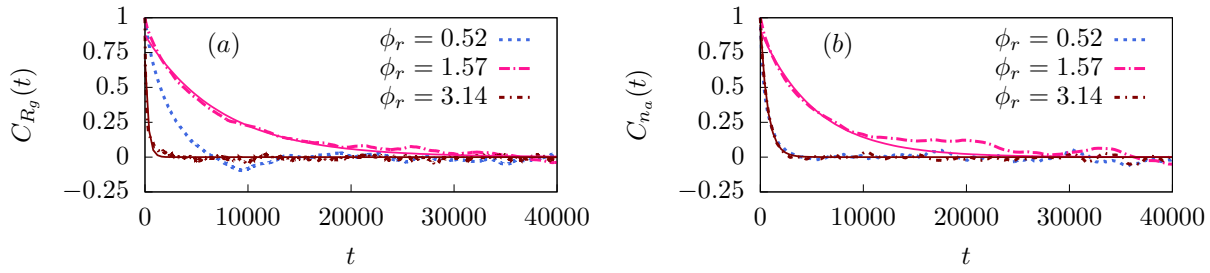


Figure 2.9: The auto-correlation functions $(a)\, C_{R_g}(t)$ of polymer radius of gyration R_g, and $(b)\, C_{n_a}(t)$ of the number of cross-linkers attached to the chain n_a, at three cross-linker densities $\phi_r = \phi_c \times 10^3$; t is expressed in unit of τ. Fitting them to exponential forms $\exp(-t/\tau_c)$ gives correlation time $\tau_c = \tau_{R_g}, \tau_{n_a}$ for R_g and n_a respectively. Two such fittings are shown in each plot by solid lines. The fitted correlation times are $\tau_{R_g} = 7371\,\tau$, $\tau_{n_a} = 5402\,\tau$ at $\phi_r = 1.57$, and $\tau_{R_g} = 307\,\tau$, $\tau_{n_a} = 720\,\tau$ at $\phi_r = \pi$.

is done using the two-time normalized auto-correlation functions of the polymer radius of gyration $C_{R_g}(t) = \langle \delta R_g(t) \delta R_g(0) \rangle / \langle \delta R_g^2 \rangle$, and the total number of bound cross-linkers $C_{n_a}(t) = \langle \delta n_a(t) \delta n_a(0) \rangle / \langle \delta n_a^2 \rangle$, where $\delta R_g(t) = R_g(t) - \langle R_g \rangle$ and $\delta n_a(t) = n_a(t) - \langle n_a \rangle$ denote instantaneous deviations of the two quantities around their respective mean values, and $\langle \delta R_g^2 \rangle$, $\langle \delta n_a^2 \rangle$ represent the corresponding standard deviation of the data. We compute respective $C_{R_g}(t)$ and $C_{R_g}(t)$ at various ϕ_c across the phase transition. In Fig. 2.9 (a) and

(b), we show the correlation functions $C_{R_g}(t)$ and $C_{n_a}(t)$ at $\phi_c = 0.52 \times 10^{-3}$ (blue dashed line), 1.57×10^{-3} (pink dashed line) and 3.14×10^{-3} (brown dashed line) respectively. The correlations show approximate exponential decay $\exp(-t/\tau_c)$ with correlation time τ_c denoted by τ_{R_g} for the polymer radius of gyration, and τ_{n_a} for the total number of polymer bound cross-linkers. The fitted correlation times τ_{R_g} and τ_{n_a} are plotted in Fig. 2.10 as a function of ϕ_c. For our finite sized chain, the corresponding correlation times $\tau_c = \tau_{R_g}$, τ_{n_a} show sharp increase at ϕ_c^* (Fig. 2.10), reminiscent of the critical slowing down [165]. As we show below in Sec. 2.2.6, scaling analysis is used to get the analytic expression of relaxation time given by Eq.(2.12). The corresponding scaling form $(1 - \phi_c/\phi_c^*)^{-3/2}$, shown in solid red line, agrees with data points for τ_{R_g} and τ_{n_a} below the transition point ($\phi_c < \phi_c^*$). Due to visibility reasons, we have shifted the analytic expression (red solid line) by adding a constant number.

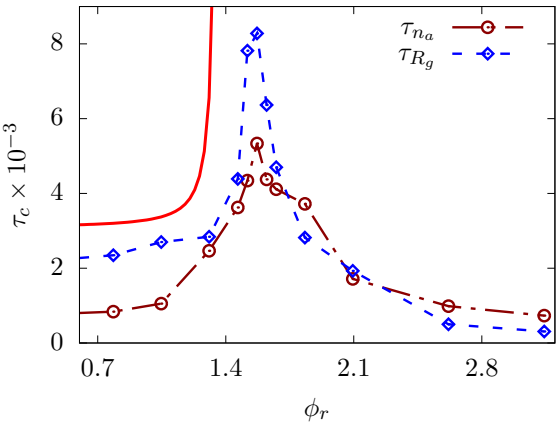


FIGURE 2.10: The correlation times $\tau_c = \tau_{R_g}$ ($\circ$), τ_{n_a} ($\diamond$) are obtained from two-time correlations of polymer radius of gyration R_g, and the total number of chromatin-bound cross-linkers n_a. They reach their maximum values at the transition point $\phi_c^* = 1.57 \times 10^{-3}$, with $\phi_r = \phi_c \times 10^3$. The solid red line is a shifted plot of the scaling form $(1 - \phi_c/\phi_c^*)^{-3/2}$ added to a constant background, with the shift aimed at better visibility. .

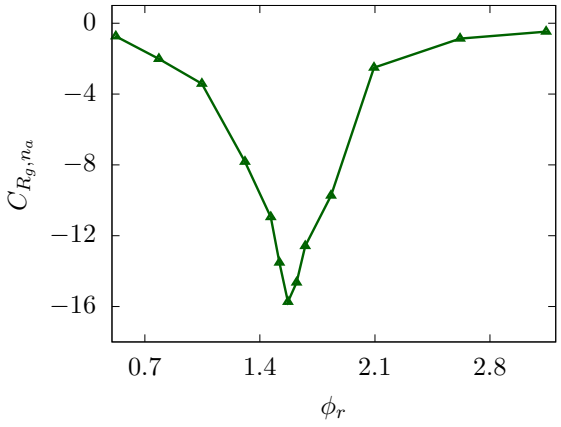


Figure 2.11: The negative values of the cross-correlation coefficient C_{R_g,n_a} between fluctuations in R_g and n_a show anti-correlation, with the amplitude maximizing at the transition point $\phi_c^* = 1.57 \times 10^{-3}$. Here $\phi_r = \phi_c \times 10^3$.

The fluctuations in n_a and R_g are anti-correlated in our system, because of the attractive interaction between the cross-linkers and monomers. The negative values of the cross-correlation coefficient at equilibrium $C_{R_g,n_a} = (1/\tau_p) \int^{\tau_p} dt \langle \delta R_g(t) \delta n_a(t) \rangle$ quantifies this anti-correlation. Remarkably, the amount of anti-correlation maximizes at the critical point ϕ_c^* signifying a large reduction in polymer size associated with a small increase of attached cross-linkers, and vice versa (Fig. 2.11). At this point, the model chromatin morphology is most susceptible to small variations in the number of attached cross-linkers. A living cell may utilize this physical property for easy conformational reorganization, useful for providing access to DNA-tracking enzymes in an otherwise folded chrmosome.

2.2.4 System size dependence at coil-globule transition

We performed simulations with various chain lengths N to check the system size dependence on the coil-globule transition. We have considered chains of sizes $N = 128, 256, 512$ and 1024, with corresponding box size large enough to nullify possible boundary effects. For the largest system size, the number of cross-linkers considered are up to 80000. A continuous change of $\langle R_g \rangle$ with cross-linker density $\phi_c = \phi_r \times 10^{-3}$ (Fig. 2.12) is observed across the various

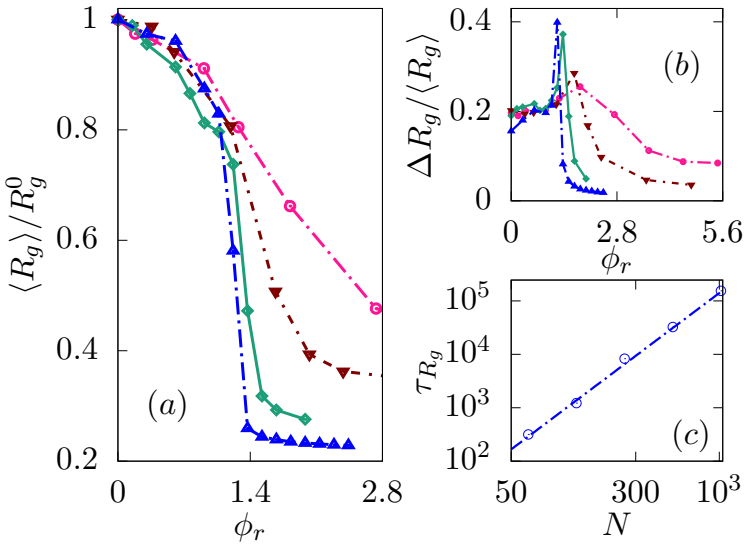


FIGURE 2.12: In (a) and (b) respectively, the scaled radius of gyration $\langle R_g\rangle/R_g^0$ and relative fluctuations of radius of gyration $\Delta R_g/\langle R_g\rangle$ is plotted with respect to $\phi_r = \phi_c \times 10^3$ for chain lengths $N = 128\,(\bigcirc),\ 256\,(\triangledown),\ 512\,(\diamond), 1024\,(\triangle)$. R_g^0 is radius of gyration for the particular chain in the absence of cross-linkers. In (c), the power law growth of correlation time at the transition point τ_{R_g} with increase of chain size N is represented. The dash-dotted line denotes $\tau_{R_g} \sim N^{9/4}$.

system sizes. The curves showing decrease of $\langle R_g\rangle$ with ϕ_c display sharper transitions with increasing system sizes N. The relative fluctuations $\Delta R_g/R_g$ at the transition point increases with system size N (Fig. 2.12(b)). The correlation time τ_{R_g} characterizing fluctuations in R_g at the transition point grow as the power law $\tau_{R_g} \sim N^{\zeta}$ with $\zeta \approx 9/4$ (Fig. 2.12(c)), thus suggesting divergence in the thermodynamic limit. In Fig. 2.13, we show probability distribution of radius of gyration $P(R_g)$ for the largest system size $N = 1024$. Across the transition, unimodal distribution is observed corroborating the continuous transition.

2.2.5 Clustering of cross-linkers

Till now, we have discussed the cluster size in terms of the number of cross-linkers in the cluster. Here we consider the cluster size in terms of their spatial extension. Due to repulsion between the cross-linkers, the clusters have spatial extension that we represent in terms of their radius of gyration R_g^c. In Fig. 2.14(a), we plot change of $\langle R_g^c\rangle$ as a function of

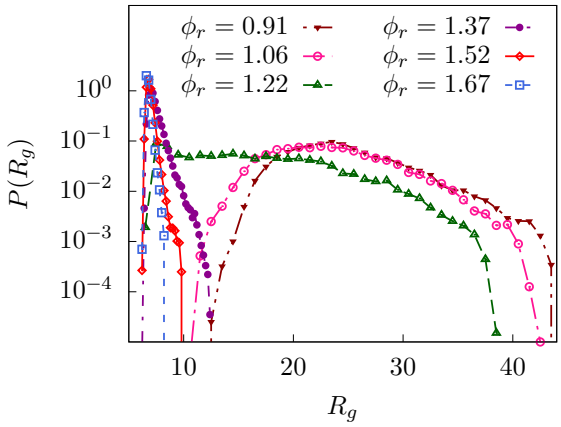


Figure 2.13: The probability distribution function for radius of gyration of the chain, $P(R_g)$ is plotted across the coil-globule transition for the chain size $N = 1024$.

$\phi_c = \phi_r \times 10^{-3}$ in the environment. $\langle R_g^c \rangle$ increases significantly as ϕ_c approaches phase transition from below, a behavior similar to $\langle C_s \rangle$ (see Fig. 2.5(b)). At higher ϕ_c, saturation in R_g^c is observed. The corresponding relative fluctuation $\Delta R_g^c / R_g^c$ is plotted as a function of ϕ_c (shown in the inset). At the transition point $\phi_c^* = 1.57 \times 10^{-3}$, the relative fluctuation has a maximum similar to previous measure of cluster size fluctuation $\Delta C_s / C_s$ ((see Fig. 2.5(b) inset (i)). The corresponding probability distribution $P(R_g^c)$ shows single peak at low ϕ_c, however bimodal distribution at densities $\phi_c \geq 1.83 \times 10^{-3}$ (see Fig. 2.14(b)), a behavior consistent with $P(C_s)$ (see Fig. 2.8(b)).

2.2.6 Mean field description

In the view of above phenomenology, we present a mean field model in terms of two coupled fields, the cross-linker density $\phi_c(\mathbf{r})$, and the deviation of monomer density due to cross-linkers $\rho(\mathbf{r}) = \rho_m(\mathbf{r}) - \rho_b$. Here, bare monomer density is given as $\rho_b = \sigma^3 N / (R_g^0)^3$, where R_g^0 denotes the radius of gyration of the open chain in absence of cross-linkers, and $\rho_m(\mathbf{r})$ represents monomer density in the presence of cross-linkers (defined later). A fraction of total cross-linkers are in polymer bound state $\phi(\mathbf{r})$, and the rest of them diffuse in space

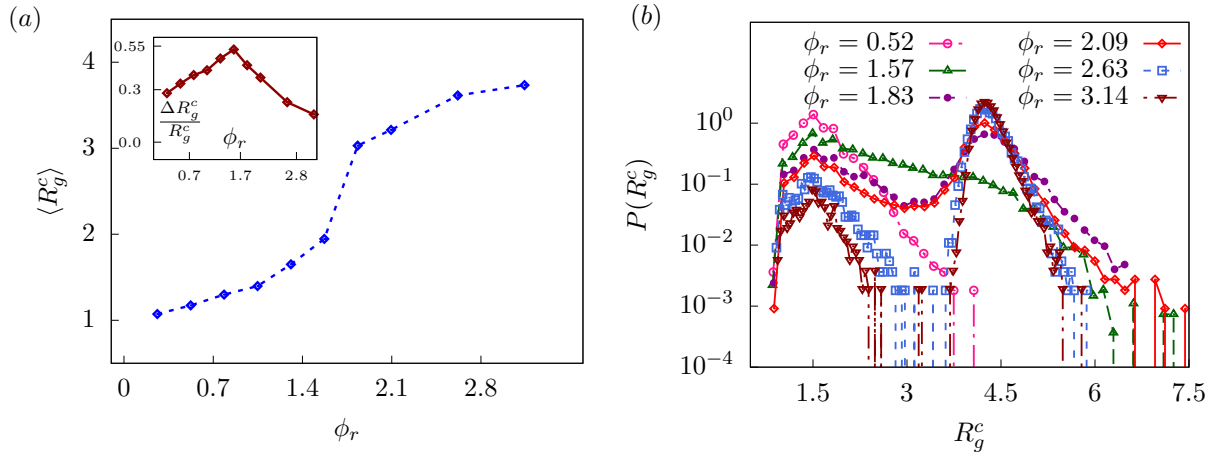


FIGURE 2.14: (a) The average radius of gyration of the polymer bound cross-linker clusters, $\langle R_g^c \rangle$ increases with $\phi_r = \phi_c \times 10^3$. The data are shown by $\diamond$ with the corresponding standard error. Blue dashed line is guide to eye. Inset (i) shows relative cluster size fluctuation $\Delta R_g^c / R_g^c$ plotted against ϕ_r. (b) Corresponding probability distributions for mean cluster radius of gyration, $P(R_g^c)$ is plotted at various ϕ_r across the transition.

constituting the detached fraction. We adopt the following free energy density [165],

$$\beta f = \frac{1}{2} u \left(1 - \frac{\phi}{\phi_*}\right) \rho^2 + \frac{v}{4} \rho^4 + \frac{\kappa}{2} (\nabla \rho)^2 + \frac{1}{2} w \phi^2. \tag{2.5}$$

The direct repulsion between polymer segments and between cross-linkers are captured by free energy costs $u\rho^2/2$ and $w\phi^2/2$ respectively. The bond formation between two polymeric segments via cross-linker proteins is captured by the three body term $\rho\,\phi\,\rho$ with strength $-u/2\phi_*$. The quartic energy cost $v\rho^4/4$ is introduced to provide thermodynamic stability. The coefficient κ in the gradient term adds free energy cost to the formation of sharp interfaces in local monomer-density. The evolution of coupled fields are represented by [165],

$$\begin{aligned}\frac{\partial\rho}{\partial t} &= M_\rho\nabla^2\left[u\left(1-\frac{\phi}{\phi_*}\right)\rho+v\rho^3-\kappa\nabla^2\rho\right]\\ \frac{\partial\phi}{\partial t} &= M_\phi\nabla^2\left[-\frac{u}{2\phi_*}\rho^2+w\phi\right]-r(\phi-\phi_0),\end{aligned} \tag{2.6}$$

where, the second term in the right hand side of second equation accounts for the turnover between attached and detached fraction of cross-linkers. Here, $r = (r_a + r_d)$, $\phi_0 = \Omega\phi_c$ with $\Omega = r_a/(r_a + r_d)$. The attachment and detachment rates, r_a and r_d respectively, are determined by interactions between the particles and the detailed balance condition. The coefficients M_ρ and M_ϕ denote mobalities of the two conserved fields ρ and ϕ respectively. A similar approach was used earlier in Ref. [106]. In the uniform equilibrium state $\phi = \phi_0$, and $\rho = \rho_0$. Using $\phi_0 = \Omega\phi_c$ and $\phi_* = \Omega\phi_c^*$, if $\phi_c < \phi_c^*$ the solution $\rho_0 = 0$, else

$$\rho_0^2 = \frac{u}{v}\frac{(\phi_c-\phi_c^*)}{\phi_c^*}. \tag{2.7}$$

Chromosome size

The mean monomer density $\rho_m = \sigma^3 N/\langle R_g\rangle^3 = \rho_0 + \rho_b$. As $\phi_c \geq \phi_c^*$, using Eq.(2.7) one obtains

$$\langle R_g\rangle = R_g^0\left[1+N^{4/5}\left(\frac{u}{v}\frac{\phi_c-\phi_c^*}{\phi_c^*}\right)^{1/2}\right]^{-1/3}. \tag{2.8}$$

This shows reasonable agreement with simulation results with fitting parameter $u/v = 0.1$ (Fig. 2.5(a)), as fluctuations are suppressed in the globule phase [154]. In the limit of $\phi_c \gg \phi_c^*$, $\langle R_g\rangle \approx N^{1/3}\sigma[(u/v)(\phi_c-\phi_c^*)/\phi_c^*]^{-1/6}$, i.e., an equilibrium globule with $\langle R_g\rangle \sim N^{1/3}\sigma$ gets further compacted with cross-linker density as $[(u/v)(\phi_c - \phi_c^*)/\phi_c^*]^{-1/6}$. The solution $\rho_0 = 0$ at $\phi_c < \phi_c^*$ corresponds to an open chain following Flory scaling $R_g^0 \approx \sigma N^{3/5}$. One may include the bilinear coupling between the monomer and the cross-linker density fields.

Considering such coupling and restricting the terms up to quadratic order gives

$$\beta f = \frac{1}{2}u\,\rho^2 + \frac{1}{2}w\phi^2 - z\rho\phi + \frac{\kappa}{2}(\nabla\rho)^2.$$

This does not describe the phase transition, however, suggests a uniform mean field solution $\rho_0 = z\phi_0/u = z\Omega\phi_c/u$. Thus, before the transition, mean radius of gyration is expected to decrease with ϕ_c as $\langle R_g\rangle = R_g^0[1 + N^{4/5}(z\Omega/u)\phi_c]^{-1/3}$.

Cluster size

As we have observed earlier, with an increase of cross-linkers density in the bulk, structures start to grow in the form of cross- linker clusters which folds the polymer. An estimate of the increase in the cluster size of the polymer-bound cross-linkers can be obtained by performing linear stability analysis of Eq.(2.6) around a homogeneous state of $\rho = \bar{\rho}$ and $\phi = \bar{\phi}$. To characterize the dynamics, we use small deviations about a homogeneous state as, $\rho = \bar{\rho} + \delta\rho(\mathbf{r})$, $\phi = \bar{\phi} + \delta\phi(\mathbf{r})$. The dynamics in Eq.(2.6) for these small deviations become

$$\begin{aligned}
\partial_t\delta\rho &= D_\rho\nabla^2\delta\rho - M_\rho\kappa\,\nabla^4\delta\rho - M_\rho\chi\,\nabla^2\delta\phi \\
\partial_t\delta\phi &= D_\phi\,\nabla^2\delta\phi - M_\phi\chi\,\nabla^2\delta\rho - r\,\delta\phi,
\end{aligned}$$

where

$$D_\rho = M_\rho u\left[\left(1 - \frac{\bar{\phi}}{\phi_*}\right) + 3\frac{v}{u}\bar{\rho}^2\right],$$

and $D_\phi = M_\phi w$ are the effective diffusion constants of the two components, and $\chi = u\bar{\rho}/\phi_*$ denote the strength of cross-coupling. In the above equations, the partial derivative with respect to time t is represented as ∂_t. Expressing time in units of inverse turnover rate,

$\tau_u = 1/r$ and lengths in units of $x_u = \sqrt{M_\phi w/r}$, one finds

$$\begin{aligned} \partial_\tau \delta\rho &= \mathcal{D}_0 \nabla_\xi^2 \delta\rho - \mathcal{K} \nabla_\xi^4 \delta\rho - \mathcal{C} \nabla_\xi^2 \delta\phi \\ \partial_\tau \delta\phi &= \nabla_\xi^2 \delta\phi - \mathcal{C}' \nabla_\xi^2 \delta\rho - \delta\phi, \end{aligned} \tag{2.9}$$

with control parameters of the dynamics $\mathcal{D}_0 = D_\rho/M_\phi w$, $\mathcal{K} = \frac{M_\rho}{M_\phi^2}\frac{\kappa r}{w^2}$, $\mathcal{C} = \frac{M_\rho}{M_\phi}\frac{\chi}{w}$, and $\mathcal{C}' = \frac{\chi}{w}$. The dimensionless time and length scales are denoted by $\tau = t/\tau_u$, and $\xi = x/x_u$, respectively.

Fourier transform of above equation leads to evolution of modes as matrix equations $\partial_\tau \left(\delta\rho_q, \delta\phi_q\right) = \mathcal{M}\left(\delta\rho_q, \delta\phi_q\right)$, where,

$$\mathcal{M} = \begin{pmatrix} -q^2(\mathcal{D}_0 + \mathcal{K}q^2) & \mathcal{C}q^2 \\ \mathcal{C}'q^2 & -(q^2+1) \end{pmatrix}.$$

The eigenvalues of matrix $\mathcal{M}$ are given by

$$\lambda(q^2) = \frac{1}{2}\left\{ \mathrm{Tr}\mathcal{M} \pm \sqrt{(\mathrm{Tr}\mathcal{M})^2 - 4\,\mathrm{det}\mathcal{M}} \right\}$$

As trace of this matrix

$$\mathrm{Tr}.\mathcal{M} = -q^2(\mathcal{D}_0 + \mathcal{K}q^2) - (q^2+1) \; < \; 0,$$

the only way of having instability (one of the eigenvalues becomes positive) is if the determinant

$$\mathrm{det}\mathcal{M} = q^2(q^2+1)(\mathcal{D}_0 + \mathcal{K}q^2) - \mathcal{C}\mathcal{C}'q^4 < 0.$$

This last criterion leads to, $\mathcal{C}\mathcal{C}' > F(q^2)$, where $F(q^2) = (1 + \frac{1}{q^2})(\mathcal{D}_0 + \mathcal{K}q^2)$. This will be satisfied irrespective for any q^2 if even the minimum of $F(q^2)$ obeys this inequality. One can

easily show that $F(q^2)$ is minimized at $q_0^2 = \sqrt{\mathcal{D}_0/\mathcal{K}}$ and $F(q_0^2) = (\sqrt{\mathcal{D}_0} + \sqrt{\mathcal{K}})^2$. Thus the instability criterion becomes,

$$\sqrt{\mathcal{C}\mathcal{C}'} > (\sqrt{\mathcal{D}_0} + \sqrt{\mathcal{K}}).$$

Here $\mathcal{C}$ and $\mathcal{C}'$ represent the coupling coefficients between evolution of the two fields ρ and ϕ as given in Eq.(2.9). Following the inequality, one can find a minimal coupling strength χ that is required to generate instability towards formation of cross-linker clusters,

$$\chi > \sqrt{\frac{\kappa r}{M_\phi}} + \sqrt{uw\left[\left(1 - \frac{\bar{\phi}}{\phi_*}\right) + 3\frac{v}{u}\bar{\rho}^2\right]}.$$

Once this condition is satisfied, instability in the form of clustering of cross-linkers, mediated by the attractive interaction with monomers, arise. The fastest growing mode $q_0 = (\mathcal{D}_0/\mathcal{K})^{1/4}$ predicts the most unstable length scale $\ell_0/x_u = 2\pi/q_0 = 2\pi(\mathcal{K}/\mathcal{D}_0)^{1/4}$, which gives the mean extension of the polymer-bound clusters

$$\ell_0 = 2\pi\left[M_\phi \frac{\kappa w}{ru}\frac{1}{\left(1 - \frac{\bar{\phi}}{\phi_*}\right) + 3\frac{v}{u}\bar{\rho}^2}\right]^{1/4}. \tag{2.10}$$

Assuming the constant density of the cross-linkers in the polymer-bound clusters, the average cluster size behave like $\langle C_s\rangle \sim \ell_0^3$, leading to

$$\langle C_s\rangle = \mathcal{A}\left[\left(1 - \frac{\phi_c}{\phi_c^*}\right) + 3\frac{v}{u}\bar{\rho}^2\right]^{-3/4}. \tag{2.11}$$

Replacing $\bar{\rho} = (z\Omega/u)\phi_c$, the dependence $\langle C_s\rangle = \mathcal{A}[(1-\phi_c/\phi_c^* + \mathcal{B}\phi_c^2)]^{-3/4}$ reasonably captures the growth of average cluster size in the numerical simulation with $\mathcal{A} = 1.9$ and small enough $\mathcal{B} = (3vz^2\Omega^2/u^3)$ such that $\mathcal{B}\phi_c^2 \ll 1$, as the coil- globule transition is approached from below (Fig. 2.5(b)).

Time scale

The diverging time-scales observed in simulations can be undersood using the following scaling argument based on Eq.(2.6). For this purpose, we use the length scale associated with the unstable mode l_0. Eq.(2.6) suggests a relaxation time $\tau_r \approx (\ell_0^2/M_\rho u)(1-\phi_c/\phi_c^*)^{-1}$. Using $\gamma\phi_c^2 << 1$ the realtion simplifies to

$$\tau_r \approx \frac{4\pi^2}{M_\rho u}\left(M_\phi \frac{\kappa w}{ru}\right)^{1/2}\left[1-\frac{\phi_c}{\phi_c^*}\right]^{-3/2}, \tag{2.12}$$

suggesting a divergence of correlation times as $(1-\phi_c/\phi_c^*)^{-3/2}$ near the critical point. For finite sized chains, while the time scales do not diverge, they show significant increase near criticality (Fig.2.10). Added with a constant background, Eq. 2.12 provides a reasonable description of the simulation results. As is shown in Fig. 2.12(c), the correlation time at criticality increases with chain length with an approximate power law $\sim N^{9/4}$ indicating divergence.

2.3 Local morphology

The binder mediated chromosomal compaction is associated with local morphological changes. The cross-linking due to binders may cause loop formation. In chromosomes, formation of such loops are expected to be highly complex, involving polydispersity of loop-sizes. The cross-linkers may also form zipper between contiguous polymeric segments. These, in turn, would enhance contact formation, and as a result modify subchain extensions. In this section, we discuss the change in all of these three aspects along the phase transition described above.

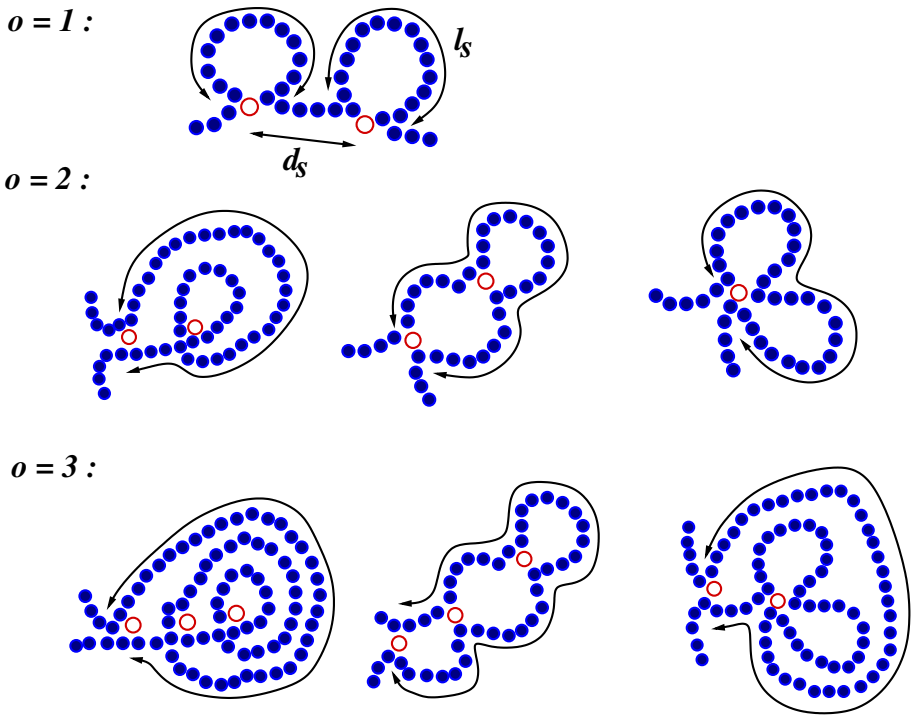


FIGURE 2.15: Schematics of loop topologies of order o: Polymer segments are indicated by blue beads and polymer-bound cross-linkers are shown by red open circles. (a) Simply connected loops of order $o = 1$. Two first order loops of size l_s are separated by a gap of size d_s. (b) Three examples of $o = 2$ loops. In the first two cases, the second order loop has one $o = 1$ loop embedded inside. The third case shows two embedded $o = 1$ loops. (c) Three examples of $o = 3$ loops. In the first two cases, the third order loop has a $o = 1$ and a $o = 2$ loop embedded. The third case shows two first order loops and a second order loop embedded inside the $o = 3$ loop.

2.3.1 Loops

We describe the possible loop-topologies with the help of Fig.2.15. The simply connected loop is called the loop of first order $o = 1$. A simply connected, or, first order loop is formed by a cross-linker binding two segments of the polymer in such a way that if one moves along the chain from one such segment to the other, no other contact or cross-link is encountered on the way. With removal of the cross-linker- bond stabilizing such a loop, the first order loop itself disappears (Fig.2.15: $o = 1$). In the figure, $\langle l_s \rangle$ and $\langle d_s \rangle$ denote loop-size and gap-size between such loops, respectively. In numerical evaluation of mean $\langle d_s \rangle$, all intermediate higher order loops are disregarded.

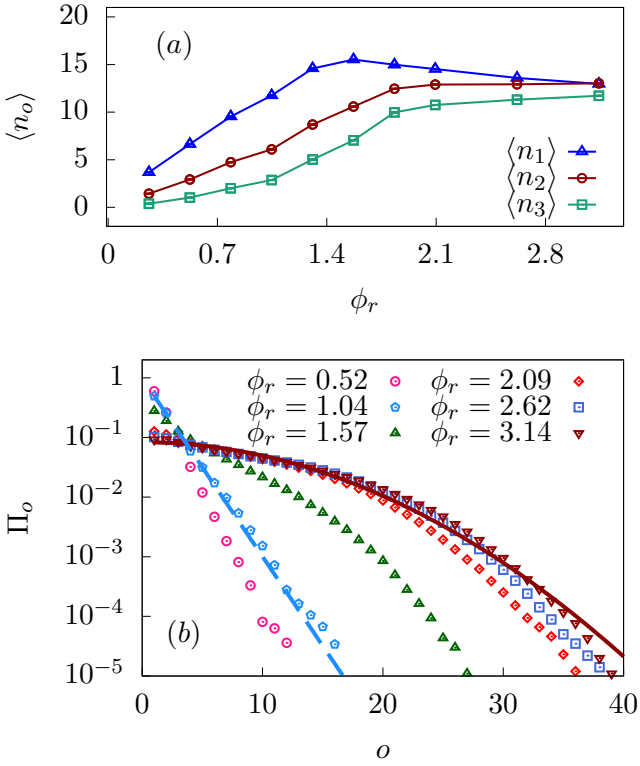


Figure 2.16: (*a*) Mean number of o-th order loops $\langle n_o \rangle$ as a function of density of cross-linkers $\phi_r = \phi_c \times 10^3$. Here $\langle n_{1,2,3} \rangle$ denote the mean number of first, second and third order loops. (*b*) Probability of o-th order loop Π_o is plotted on semi-log scale, for various cross-linker densities denoted in the labels. At $\phi_r = 3.14$, the probability of higher order loops decays with an approximate Gaussian form $\exp(-o^2/2g^2)$ where the standard deviation $g = 9.83$ (solid brown line). For $\phi_r = 1.04$, the probability of higher order loops decays exponentially as $\exp(-o/\bar{o})$ with $\bar{o} = 1.45$ (dashed blue line).

A higher order loop denoted by order $o = n$, embeds all possible lower order loops $o = 1, \ldots, (n-1)$ within it. In Fig.2.15: $o = 2$, three examples of second order loops are shown. In the first two examples removing one cross-linker reduces the second order loop to a first order loop. In the third example of $o = 2$ loop, three bonds of a single cross-linker maintains the loop, and with its removal the whole loop structure disappears. In Fig.2.15: $o = 3$ we show three examples of third order loops. Note that the first order and higher order loops identified here are related to the serial and parallel topologies described in Ref. [166]. As it has been shown before, consideration of chromosomal loops is crucial in understanding of

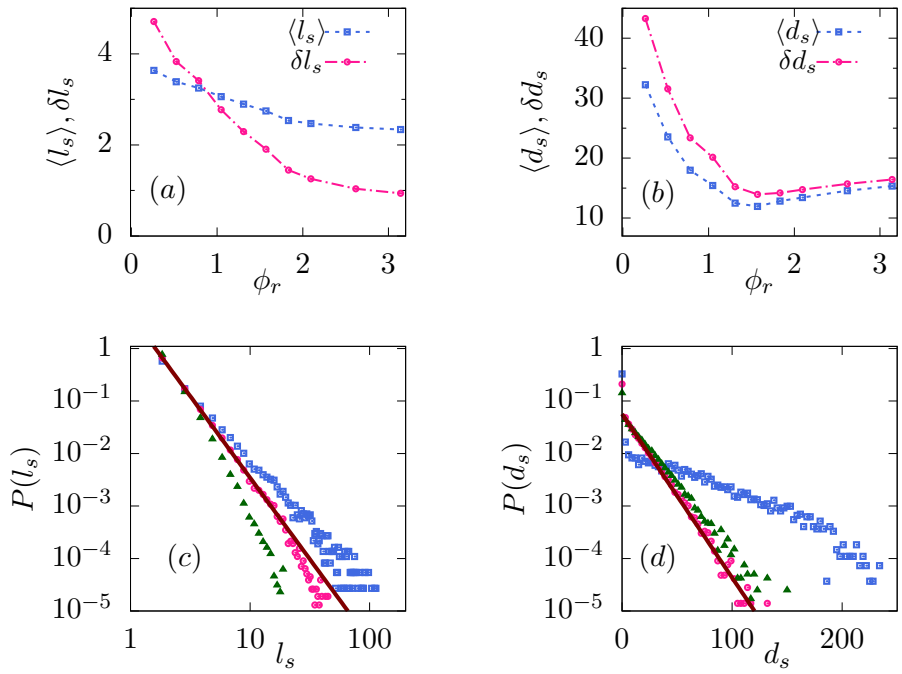


FIGURE 2.17: (a) Decrease of the mean size of first order loops $\langle l_s \rangle$ and its fluctuations δl_s with $\phi_r = \phi_c \times 10^3$. (b) Non-monotonic variation of mean separation between first order loops $\langle d_s \rangle$ and its fluctuations δd_s with ϕ_r. Probability distributions of the size of first order loops $P(l_s)$ and gaps between them $P(d_s)$ are plotted in (c) and (d) at $\phi_r = 0.26\,(\square), 1.57\,(\circ), \pi\,(\triangle)$. At the transition point $\phi_r = 1.57$, $P(l_s) \sim l_s^{-3.3}$, and $P(d_s) \sim \exp(-d_s/\lambda)$ with $\lambda = 13.8\,\sigma$, shown by the solid (brown) lines in (c) and (d) respectively.

its emergent behavior [9, 10, 18, 19]. In this chapter, we restrict ourselves to the relative importance of different orders of loops in local chromosomal morphology.

In Fig.2.16(a) the mean number of loops $\langle n_o \rangle$ of order $o = 1, 2, 3$ are shown against the cross-linker density ϕ_c. All through, $\langle n_1 \rangle$ remains larger than $\langle n_{o=2,3} \rangle$ corresponding to higher order loops that show a sigmoidal dependence on ϕ_c. Interestingly, $\langle n_1 \rangle$ maximizes at the phase transition point ϕ_c^*. Thus at the critical point the local morphology of the model chromosome is dominated by the first order loops.

Fig.2.16(b) shows the probability Π_o of a loop to be of o-th order. At small cross-linker densities $\phi_c < \phi_c^*$, the probability of higher order loops fall exponentially as $\Pi_o = \exp(-o/\bar{o})$. This behavior changes qualitatively after the coil-globule transition ($\phi_r = 1.57$) to a Gaussian

profile $\exp(-o^2/2g^2)$, as is shown in Fig.2.16(b) .

Given that loop sizes could be measured from electron microscopy [26], we further analyze the statistics of loop-sizes and inter-loop gaps corresponding to the first order loops in Fig. 2.17. With increasing cross-linker density ϕ_c, the mean size of first order loops $\langle l_s \rangle$ decreases (Fig. 2.17(a)), as their number increases (Fig.2.16(a)) reducing the mean gap size $\langle d_s \rangle$ (Fig.2.17(b)). However, increased ϕ_c stabilizes the loops, shown by decreased fluctuation of loop-sizes $\delta l_s = \sqrt{\langle l_s^2 \rangle - \langle l_s \rangle^2}$. The mean gap size $\langle d_s \rangle$ and its fluctuation $\delta d_s = \sqrt{\langle d_s^2 \rangle - \langle d_s \rangle^2}$ reach their minimum at the transition point $\phi_c^* = 1.57 \times 10^{-3}$. The increase in the inter-loop separation $\langle d_s \rangle$ beyond this point is due to the increase in probability of higher order loops in the local morphology of the model chromatin.

Fig.2.17(c) and (d) show the probability distributions of first order loop sizes $P(l_s)$, and separation between consecutive first order loops $P(d_s)$, respectively. For all ϕ_c values, $P(l_s) \sim l_s^{-\mu}$, with μ increasing with ϕ_c in a sigmoidal fashion, giving $\mu = 3.3$ at the critical point $\phi_c^* = 1.57 \times 10^{-3}$. The power law distribution of $P(l_s)$ shows that their is no characteristic loop size, and loops of all possible lengths are present. On the other hand, the gap size distributions follow an approximate exponential form $P(d_s) \approx (1/\langle d_s \rangle) \exp(-d_s/\langle d_s \rangle)$.

2.3.2 Zippering

The binders can also zipper different segments of the polymer. The inset of Fig.2.18 shows one such zipper maintained by cross-linkers. The zipper fraction of a conformation is given by $Z_p = (1/N) \sum_{\xi,i} N_i^\xi$, where N_i^ξ are the number of monomers involved in forming ξ-th zipper, and N is the total number of monomers in the chain. Fig.2.18 shows variation of ensemble averaged zippered fraction with the cross-linker density. The zipper fraction increases non-linearly to saturate in the equilibrium globule phase to a value that remains within 60% of the completely zippered filament $\langle Z_p \rangle = 1$. Near the critical point of the coil-globule transition $\langle Z_p \rangle \approx 0.3$, half the saturation value.

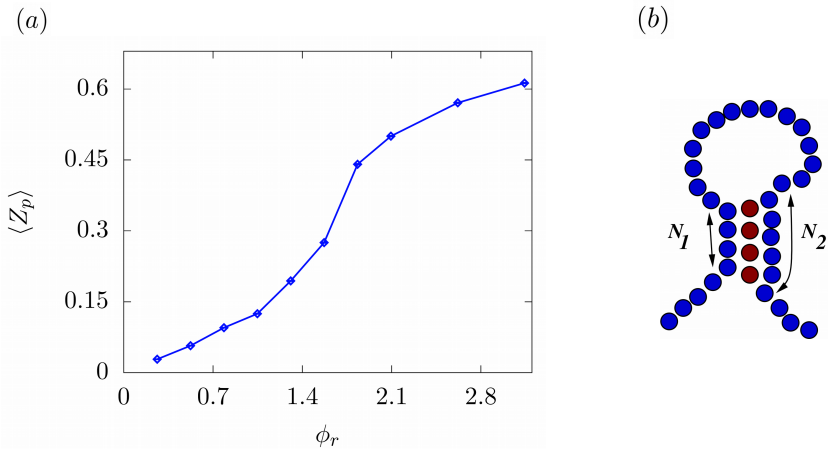


FIGURE 2.18: (a) The zippered fraction of chain $\langle Z_p \rangle$ as a function of cross- linker density $\phi_r = \phi_c \times 10^3$. (b) Schematics shows two contiguous segments containing N_1 and N_2 monomers (blue beads) forming a zipper via binders (red beads). The corresponding zipper fraction is $Z_p = (N_1 + N_2)/N$.

2.3.3 Contact probability and map

Conformational fluctuations of a chain bring its contour- wise distant segments in the spatial proximity of each other, even in the absence of cross-linkers. This leads to the formation of contact between spatially close monomers. An example of such contact is shown in Fig.2.3(c). Further, cross-linking of chain segments due to cross-linkers and the formation of zippers augment the number of contacts in the polymer conformation. To analyze contact formation from simulations one requires a finite cutoff length such that if two monomers fall within such a separation they are defined to be in contact. In our simulations, we chose a cut-off value $r_c = 1.5\sigma$ for this purpose. We have checked that our main results do not depend on the precise choice of this length scale.

The contour wise separation between two monomers, s, defines the genomic distance between chromatin segments. The contact probability $\Pi_c(s)$ is a mesure of two segmenst to be in contact. In absence of cross-linkers, we get $\Pi_c(s) \sim s^{-\alpha}$ with $\alpha \approx 2.1$, as expected for

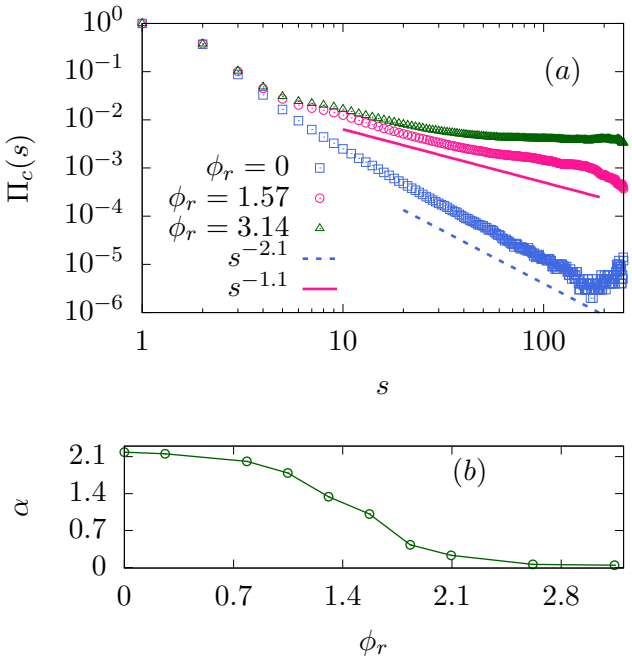


Figure 2.19: (a) Contact probability $\Pi_c(s)$ at different cross-linker densities $\phi_r = \phi_c \times 10^3$. They follow asymptotic power law profiles $\Pi_c(s) \sim s^{-\alpha}$ is shown at all ϕ_c, with α being a function of ϕ_c. (b) The decrease of asymptotic exponent α with increasing ϕ_r is related to the coil- globule transition.

self- avoiding chains [53]. Even in the presence of cross-linkers, the asymptotic power law persists with ϕ_c dependent α (see Fig. 2.19). At the critical point, $\phi_c^* = 1.57 \times 10^{-3}$, the simulation results are consistent with $\alpha \approx 1.1$, a number that agrees well with the prediction of the fractal globule model [17, 167]. It is interesting to note that $\alpha \approx 1.1$ is close to the average exponent found across all human cell chromosomes, in the genomic distances of 0.5-10 Mbp range [16, 17], and belongs to the range of exponents observed in individual mammalian chromosomes [16, 104, 168]. At large ϕ_c values, after the completion of the coil-globule transition, contact probabilities at large s plateaus to a constant, indicating $\alpha = 0$. As a function of ϕ_c, the asymptotic exponent α reveals a continuous decrease (see Fig. 2.19(b)), capturing the change in polymeric organization in the course of the coil-globule transition.

In Fig.2.20 we present ensemble averaged contact maps over equilibrium configurations

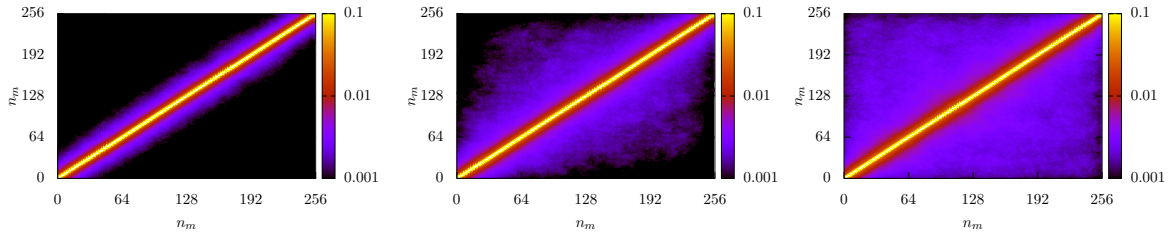


FIGURE 2.20: From left to right, the contact maps are shown at cross-linker densities $\phi_r = 1.31, 1.57$, and 1.83. n_m represents the monomer index along the chain contour. The color code captures the contact frequency and is shown in log scale.

at different cross-linker densities. Such maps represent probability measures of two chromatin segments to be in spatial proximity. At small cross-linker density ϕ_c, the contour-wise neighboring segments make frequent contacts with each other, as polymer adopt open conformations (Fig.2.20). This can be seen in the left map in the Fig.2.20, showing contacts only near the diagonal region. At the coil-globule transition $\phi_c = \phi_c^*$, the contact shows emergence of local pattern, indicating enhanced probability of contour-wise well separated segments to come into spatial proximity. In the compact phase at $\phi_c = 1.83 \times 10^{-3}$, the chromosomal contacts spread over the whole chromatin chain.

2.4 Extension of subchains

Here we consider the scaling behavior of subchain extensions, measured in terms of the mean squared end to end distance $\langle r^2(s)\rangle$ in subchains of contour length s. We observe three different scaling behaviors across the coil-globule transition (see Fig. 2.21(a)).

A sub-chain inside a compact equilibrium globule is expected to behave like a random walk due to strong screening of interaction by large monomeric density. Thus $\langle r^2(s)\rangle \sim s$, before the globule boundary is encountered. Multiple reflections from the globule boundary, as $s > \langle r^2(s)\rangle \sim N^{2/3}$, fills the space inside the globule uniformly, so that it becomes equally likely to find the other end of the subchain anywhere inside the globule, saturating $\langle r^2(s)\rangle$ to a

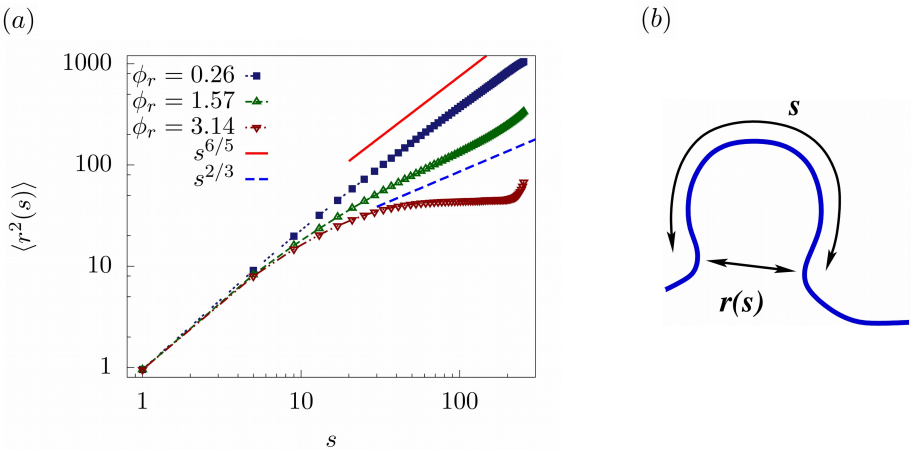


Figure 2.21: (*a*) The scaling behavior of end to end separation of subchain $\langle r^2(s)\rangle$. The dependence is shown at three cross-linker concentrations $\phi_r = \phi_c \times 10^3$, before, at and after the coil-globule transition. At low densities, $\langle r^2(s)\rangle \sim s^{6/5}$ follows Flory scaling (red solid line). At the transition point $\phi_c = 1.57 \times 10^{-3}$, the asymptotic behavior of $\langle r^2(s)\rangle \sim s^{2/3}$ agrees with the fractal globule estimate (blue dashed line). At the highest concentrations we find asymptotic plateauing, a characteristic of equilibrium globule. (*b*) Schematic showing a subchain of contour size s has spatial separation $r(s)$.

constant. Thus in equilibrium globules $\langle r^2(s)\rangle \sim s$ up to $s < N^{2/3}$, and saturates beyond that length scale [17, 53]. The random loop model, with fixed probability of attraction between monomers, shows all the features of equilibrium globule in final configurations [169, 170]. On the other hand, the fractal globule is space filling at all scales, such that $\langle r^2(s)\rangle \sim s^{2/3}$ [17, 167].

At small ϕ_c ($= 0.26 \times 10^{-3}$), we find a behavior typical of open chains, $\langle r^2(s)\rangle \sim s^{6/5}$, that follows Flory scaling (see Fig. 2.21(*a*)). In the fully folded compact phase at high ϕ_c ($= \pi \times 10^{-3}$), $\langle r^2(s)\rangle$ shows plateauing at large s as in compact equilibrium globules, and random loop models [17, 153, 169, 170]. Such plateauing was earlier related to folding of chromosome into territories [146]. In the compact phase, the molecular cross-linkers may not only pull different segments close to each other, by doing so, they may displace well separated parts further away from each other [135], reflected in the eventual increase of $\langle r^2(s)\rangle$ as s

approaches the full length N, e.g., at highest ϕ_c. At the critical point, ϕ_c^* ($= 1.57 \times 10^{-3}$), simulation results for subchain extensions is consistent with $\langle r^2(s)\rangle \sim s^{2/3}$ as in fractal globules [167]. This is close to the threshold-exponent predicted in [104] $\langle r^2(s)\rangle \sim s^{2\nu}$ with $\nu = 0.39$. Thus with increasing cross-linker density, the model chromatin morphology changes from an open chain to compact equilibrium globule, via an intermediate fractal globule behavior observed at the critical point. The sustenance of fractal globule like non-equilibrium behavior at the critical point can be understood in terms of the super-slow relaxation.

2.5 Conclusion and outlook

In summary, using an off-lattice model of self avoiding polymer and diffusing protein binders cross-linking different segments of the chromatin fibre, we have presented an extensive characterization of the continuous chromatin folding transition, and analyzed the associated changes in chromatin morphology in terms of formation of loops, zippering and contacts. The criticality is characterized by unimodal distributions, divergent fluctuations and critical slowing down. The negative maximum in the cross-correlation between the number of attached binders and chromosome size, at criticality, might be utilized by living cells for easy switching between folded and open conformations, providing easy access to DNA-tracking enzymes. This is suggestive of a possibility that chromosomes might be poised at criticality [171], vindicated further by the similarity of the calculated contact probability at the critical point with the average behavior of human chromosomes. Although the local chromatin morphology does show highly complex loop structures, at criticality, it is dominated by simply connected loops.

Each coarse-grained chromatin bead in our model can be considered as $10 - 12$ closely packed nucleosomes containing around $2 - 2.5\,\mathrm{kbp}$ DNA-segments having a diameter $\sigma \approx 20 - 40\,\mathrm{nm}$ [17, 172]. The dimensionless critical volume fraction ϕ_c is equivalent to a

concentration $[\phi_c^*/(4\pi\sigma^3/3)]$, which can be expressed in terms of molarity by dividing it by the Avogadro number. This leads to the estimate of critical concentration between $\sim 60\,\text{nmol/l} - 470\,\text{nmol/l}$. The mean size of the first order loops observed at criticality translates to $4-7\,\text{kbp}$. The estimated ratio of this loop size and inter-loop gaps is $\langle l_s \rangle : \langle d_s \rangle \approx 1:5$ at this concentration.

In the chromosomal environment having viscosity η, the dissipation constant $\gamma = 3\pi\eta\sigma$. As it has been observed, the nucleoplasm viscosity η felt by objects within the cell nucleus depends on their size [173, 174]. Using the measured viscosity $\sim 10\,\text{Pa-s}$ felt by solutes having $\sim 10\,\text{nm}$ size [174] for the $\sigma = 20\,\text{nm}$ beads, the characteristic time which is the same as the time required to diffuse over the length-scale σ can be determined by using the relation $\tau = \gamma\sigma^2/k_BT = 0.2\,\text{s}$. Thus, the simulated correlation time τ_{R_g} denoting chromosomal relaxation over $\sim 0.5 - 0.6\,\text{Mbp}$ translates to $\approx 22\,\text{minutes}$ at the critical point. While some of our predictions appear to compare well with experiments, others involving cross-linker clusters, relaxation time, and loop morphology are amenable to experimental verifications.

Here we should reemphasize that our study represents an average description of chromosomes using a coarse grained homopolymer model. This approach did not aim to distinguish interaction between specific protein types and gene sequences. However, the Hi-C map of chromosomes shows checkerboard pattern in the contact matrix at large genomic separation revealing the presence of alternate A(active) and B(inactive) type compartments [16, 22, 175]. Transcriptionally active, gene rich and relatively less compact euchromatin is represented by A type and less transcribed, highly compact heterochromatin is shown by B type compartments respectively. On the other hand, Hi-C map at high resolution (genomic separation $\lesssim 1$ Mbp) shows TADs of high contact frequency. Recent studies suggest, TADs are maintained by looping of chromatin where the base of majority of such loops are bound by CTCF proteins and cohesin subunits [84, 176, 177]. Such a compartmental organization occurs due to specific interaction between DNA and associated proteins leading to spatial separation of

heterochromatin and euchromatin, and formation of TADs. Note that Hi-C maps for different cell-types vary from each other. Apart from that, gene regulation depends on interaction between promoters and their regulatory elements like enhancers, repressors and insulators. Such specific interactions are mediated by transcription factors and other regulatory proteins forming large loops of size thousands of nucleotides, leading to cell type specific chromatin folding. Considering the sequence specific interaction between model chromatin and cross-linkers can potentialy lead closer to experimentally observed chromosome organization.

It is known that various DNA associated proteins have finite number of DNA binding domains. For example, binding multiplicity of CTCF proteins and HP1 with chromatin are six and two respectively [104, 178–180]. However, the simple model presented in the current chapter uses cross-linkers with a distribution of valency as a function of their density in the environment. Having a range of valency is similar to that in real chromatin. However, in the absence of specificity and heterogeneity of polymer, while our model provides a general physical basis for local chromatin folding, in detail it cannot capture any specific contact map, giving only an average picture. It is possible to incorporate such details within the model we presented, and compare with particular Hi-C maps. This remains an interesting future direction of study.

3 Chromosome compaction and morphologies due to effective inter-segment attraction

In the previous chapter, using a self avoiding polymer model of chromatin, we studied the coil- globe transition and associated morphological changes of chain due to DNA- binding proteins. Such proteins are modeled as cross-linkers which can simultaneously bind to multiple chromatin segments generating an effective attraction between them. In the current chapter, we replace the cross-linkers by introducing an additional local attractive tail in the interaction between chromatin segments. We present a study of chromosome compaction due to increasing attraction strength, and characterize the corresponding changes in local morphology in terms of inter-segment contacts, loops, and zippering, as in the previous chapter.

The effective attraction between polymer segments is known to lead to coil-globule transition [53, 122]. This transition has been studied in detail in polymer physics [181–184]. In typical situations studied in polymers, such effective attraction might stem from changing solvent quality, controlled by ionic concentration in the solvent, temperature, or pH [185, 186]. In a good solvent, polymer segments like to be in contact with solvent and experience an effective repulsion between themselves. Consequently, the chain adopts an open coil configuration following the Flory scaling of extension $R \sim N^{3/5}$ as a function of the number of segments N. This is significantly larger than the ideal chain. In contrast, in a poor solvent, the polymer segments dislike to be in contact with solvent. This leads to an effective attraction between the polymer segments finally leading to a compact globule phase following $R \sim N^{1/3}$, the size of which is significantly smaller than the ideal chain. With gradual change

in the solvent quality, the coil-globule transition occurs through a competition between entropy and energy at an intermediate point, called the θ point, at which the chain follows the ideal polymer behavior $R \sim N^{1/2}$. In contrast, in the context of chromatin, protein binders provide the effective attraction between different segments.

In this chapter, we consider a self- avoiding chain model for chromatin, considering an additional local attraction between non- bonded monomers. To mimic the effect of increasing cross-linker density, we increase the relative attraction between the polymer segments. In the competition between energy and entropy, at higher strengths of attraction, energy dominates to lead to a coil to globule transition. We compare this scenario with the other possibility of varying the ambient temperature, keeping energy scales unchanged. We perform molecular dynamics simulations in the presence of Langevin heat bath to keep the temperature constant. We find a continuous coil- globule transition, along which we analyze the change in polymer morphology.

To this end, we first analyze the contact probability between polymer segments. At small attraction strength between monomers, the contact probability show a power law decay with an exponent consistent with the self- avoiding chain estimate. As the polymer folds, this exponent monotonically reduces to vanish, capturing the polymer reorganization. At the critical point, the exponent shows a value very close to the fractal-globule estimate. The sub-chain extension follows a scaling behavior with an exponent capturing the Flory estimate in the coil phase and a fractal-globule behavior at the critical point. The zipper fraction of the chain monotonically increases as the polymer folds.

We further investigate the loop structures formed across the coil-globule transition. Topologically, such loops can be either simple loops or complex higher order loops. The mean number of simple or first order loops show a monotonic decrease as polymer folds. However, the mean number of second and third order loops show initial increase followed by saturation. In addition, the mean size of first order loops and its fluctuation decrease across the phase transition. Apart from that, the mean separation between the first order loops, and

its fluctuations increase monotonically. As the chromatin folds, the loop structures become more and more complex, and the relative probability for the formation of higher order loops increases.

At last, we focus on the contact map. Across the folding of polymer, contacts between various segments spread over the whole chain. In the globule phase, such contacts fill approximately all the map. However, using a homopolymer with the same local attraction between any two segments, we do not find any territorial organization, observed in chromosomes. The territorial organization appears due to inhomogeneous nature of the chromosome, with cross-linkers attaching to specific gene sequences. This we capture by considering a hetero-polymer model of chromatin, with differential interaction between different parts of the chain. As we show, such a model displays a checker- board like pattern in the contact map typical of the formation of topologically associating domains.

In Sec. 3.1, we describe the model and details of numerical simulation. In Sec. 3.2, we discuss the simulation results about the phase transition of model chromatin. In Sec. 3.3, we discuss local morphological changes along the coil- globule transition in terms of contacts probability, loop structures, subchain extension, and contact map. Specifically, in Sec. 3.3.5, we describe the heteropolymer model capturing formation of topologically associated domains. Finally, we conclude this chapter in Sec. 3.4.

3.1 Model

The chromatin is modeled by self interacting polymer which is realized with a bead spring model. Such a model assumes the consecutive beads of the polymer are connected by springs, that we model as harmonic springs, with potential

$$U_{Harmonic}(r_{i+1,i}) = -\frac{k}{2}(r_{i+1,i} - R_0)^2, \tag{3.1}$$

where, k and R_0 represent the bond parameters, and $r_{i+1,i} = |\mathbf{r}_{i+1} - \mathbf{r}_i|$ denote separation between consecutive monomers, where $\mathbf{r}_i$ denotes the position vector of i-th monomer. Here, $\beta = 1/k_BT$, where k_B is the Boltzmann constant and T denotes the temperature of the heath bath. The form of the potential is plotted in Fig. 3.1(b). The interaction between non-bonded beads of the polymer is given by a truncated and shifted Lennard-Jones potential with attractive tail followed by a repulsive core. The potential is represented as,

$$\begin{aligned} U_{ljs}(r) &= (U_{lj}(r) - U_{lj}(r_{thr})), \text{ for } r < r_{thr} \\ &= 0 \qquad \text{otherwise,} \end{aligned} \tag{3.2}$$

where, $U_{lj}(r) = 4\epsilon[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^{6}]$ and $r_{thr} = 2.5\sigma$. The unit of length is set by σ. We fix $R_0 = 1.0\sigma$. Let us first focus on the situation in which the temperature k_BT is held constant. This sets the unit of energy. The form of the shifted Lennard-Jones potential is shown in Fig. 3.1(a) at different values of interaction strength ϵ. Thus, polymer is defined by the Hamiltonian,

$$U(r) = U_{ljs}(r) + U_{Harmonic}(r). \tag{3.3}$$

To study the folding transition of this chain, we perform molecular dynamics simulations in the presence of Langevin heat bath. Accordingly, the stiffness for bond potential is given by, $k = 100.0\, k_BT/\sigma^2$. The integrations are performed using the velocity-Verlet algorithm with time step $\delta t = 0.01\tau$, where $\tau = \sigma\sqrt{m/k_BT}$ sets the unit of time. Mass of each bead is chosen to be $m = 1$. The simulations are performed using the implementation of the ESPResSo molecular dynamics package. To compare the results, we also perform an independent set of simulations keeping the interaction strength ϵ constant, and vary the temperature T. In this case ϵ sets the unit of energy, and as a result $\tau = \sigma\sqrt{m/\epsilon}$ sets the unit of time. We keep $k = 100.0\,\epsilon/\sigma^2$ constant. We use a smaller time-step $\delta t = 0.005\,\tau$ to

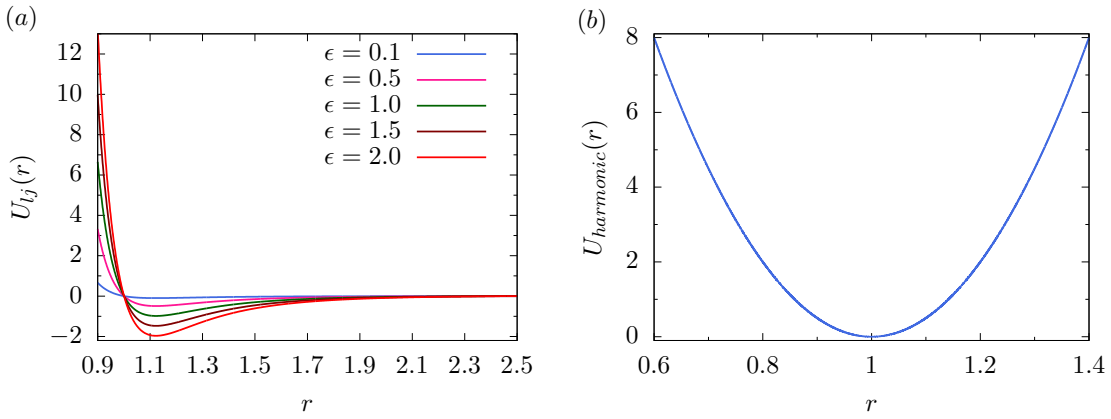


Figure 3.1: (a) The Lenard-Jones potential, $U_{ljs}(r)$ is shown at different strength ϵ, with $r_{thr} = 2.5\sigma$. The lengths are expressed in units of σ. (b) The harmonic potential, $U_{Harmonic}(r)$ is plotted with respect to r with $R_0 = 1.0\sigma$ and $k = 100\epsilon/\sigma^2$.

avoid bond breaking at higher temperatures in the latter case. The results are expressed in terms of the relative strength $\beta\epsilon = \epsilon/k_BT$.

We consider a polymer consists of $N = 256$ monomers for all the simulations. In the open extended state at $\beta\epsilon = 0.1$, the radius of gyration of the chain is given by $R_g^0 = (11.96 \pm 2.41)\,\sigma$. This limit of interaction strength $\beta\epsilon$ represents the good solvent condition allowing the coil state of the chain as a stable phase. The size of the cubic simulation box is chosen as $\mathcal{L} = 46\sigma$. This particular choice is larger than four times the radius of gyration and more than the end-to-end separation for the chain in the coil state. To study the folding behavior, first, we change our tuning parameter $\beta\epsilon$ from 0.1 to 0.8 by varying ϵ, keeping k_BT constant. Secondly, we vary $\beta\epsilon$ from 0.1 to 0.8 by changing the temperature T of heat bath keeping ϵ constant. The equilibration run is performed over $10^6\tau$, much longer than the longest relaxation time. The data is sampled over the next $10^6 - 10^7\tau$ for analysis.

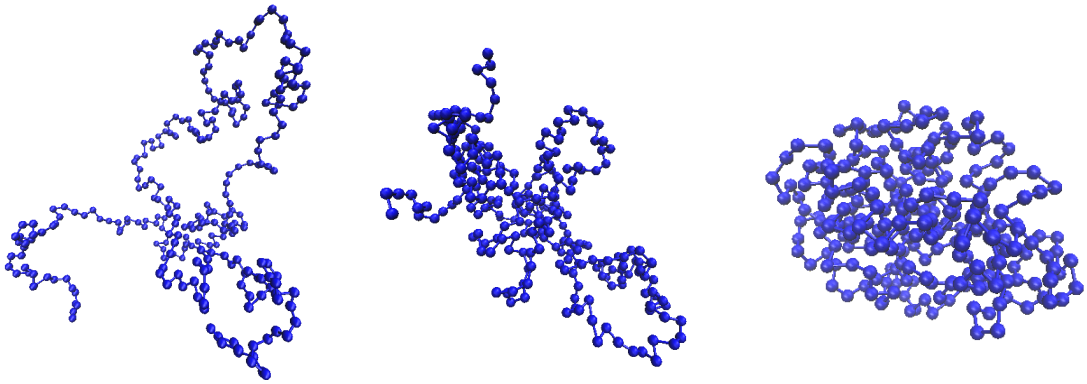


FIGURE 3.2: From left to right, snapshots of polymer configurations are shown at $\beta\epsilon = 0.1, 0.38$ and 0.5 respectively.

3.2 Results

3.2.1 The coil-globule transition

Coupled to the heat bath, the chain mimicking the chromatin undergoes structural fluctuations governed by Boltzmann rate. These thermal fluctuations can bring distant segments along the chain in spatial proximity of each other. The spatially close monomers attract each other depending on the depth of Lenard-Jones interaction as shown in 3.1(a). Probability of finding two monomers at separation r is proportional to Boltzmann factor $\sim \exp(-\beta U_{ljs}(r))$. At large $\beta\epsilon$, due to the formation of a deeper attractive well in the potential profile, this probability increases for separations within the range of attractive well of the potential $r < r_{thr}$, thus effectively forming a bound state by cross-linking the monomers. Changing the relative attraction $\beta\epsilon$ changes the effective excluded volume characterizing the interaction between non-bonded monomers [122].

This can be seen using the Mayer function

$$f(r) = \exp(-U_{ljs}(r)/k_B T) - 1. \tag{3.4}$$

This function quantifies the difference between probabilities of finding particle at separation

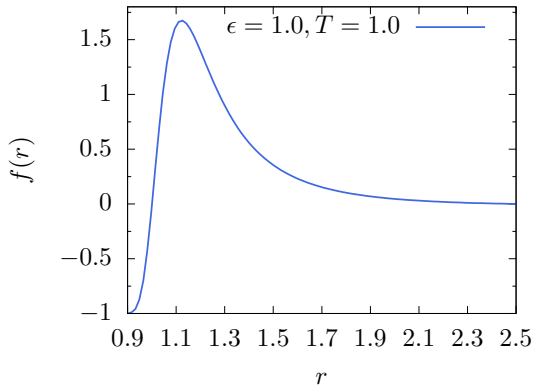


Figure 3.3: The Mayer function $f(r)$ corresponding to the Lennard-Jones potential plotted against separation the r, using $k_BT = 1$ and $\epsilon = 1$.

r and at very large separation (infinite) where interaction potential vanishes. For $\epsilon = k_B$ and $T = 1.0/k_B$, $f(r)$ is plotted in Fig. 3.3 as a function of r. It is interesting to note that for a hard-core interaction $U_{hc}(r) = \infty$ for $r \leq \sigma$ and $U_{hc}(r) = 0$ otherwise, the Mayer function $f(r) = -1$ for $r < \sigma$ and $f(r) = 0$ otherwise. This allows one to express the excluded volume around each hard-core particle as $v_{hc} := -4\pi \int_0^\infty f(r) r^2 dr = 4\pi \int_0^\sigma r^2 dr = (4/3)\pi\sigma^3$. One can extend this notion to other potentials, e.g., U_{ljs}. To that end, writing $f(r) \approx -1$ for $r \leq \sigma$, and $f(r) \approx -\beta U_{ljs}$ beyond that, one obtains a measure of an effective excluded volume,

$$\begin{aligned} v &:= -4\pi \int_0^\infty f(r) r^2 dr \\ &\approx 4\pi \int_0^\sigma r^2 dr + \frac{4\pi}{k_B T} \int_\sigma^\infty U_{ljs}(r) r^2 dr \\ &\approx \left[1 - \frac{\theta}{T}\right] \frac{4}{3}\pi\sigma^3 \end{aligned} \tag{3.5}$$

The first term is due to the effective hard core. The second term provides temperature

dependence of v, in which

$$\theta = -\frac{3}{\sigma^3 k_B}\int_{\sigma}^{\infty} U_{ljs}(r) r^2 dr \tag{3.6}$$

In the range $\sigma \leq r < \infty$, the potential $U_{ljs}(r)$ is negative leading to a $\theta > 0$. Thus,

$$v \approx \frac{T-\theta}{T} v_{hc} \tag{3.7}$$

For a positive v, different segments of the polymer effectively repel each other supporting a coil state. On the other hand, when $v < 0$, the segments attract each other preferring the globule state. The value of v can be changed from positive to negative by either changing the strength of the attractive potential ϵ changing θ, or changing the temperature T. For $T > \theta$, the repulsive core of the potential dominates leading to effective repulsion between monomers characterized by $v > 0$. This the regime of good solvent. The behavior of the self-avoiding polymer remains dominated by entropy. At $T = \theta$, the transition point $v = 0$ due to cancellation of the contribution from repulsive core and attractive well of potential $U_{ljs}(r)$. This characterizes the θ solvent in which the polymer behaves like an ideal chain. Finally, at $T < \theta$, dominant contribution comes from the attractive well yielding effective attraction between the monomers denoted by $v < 0$. This is the so called bad solvent regime, in which polymer gets into the globule phase. The tuning parameter for v is the dimensionless ratio θ/T.

In Fig. 3.2(a) we show the coil-globule transition in terms of the mean radius of gyration $\langle R_g \rangle$ of the chain with changing degree of relative inter-segment attraction $\beta\epsilon = \epsilon/k_B T$. The large $\langle R_g \rangle$ near $\beta\epsilon = 0.1$ correspond to the open and unfolded coil state of the polymer. This collapses to a globule state with small $\langle R_g \rangle$ at large $\beta\epsilon \gtrsim 0.5$. The figure shows two sets of data corresponding to change in ϵ and T separately.

As mentioned in the model section, in the first set of simulations, we varied ϵ fixing the T. Secondly, we have fixed ϵ and varied T. The coil-globule transition is presented as a

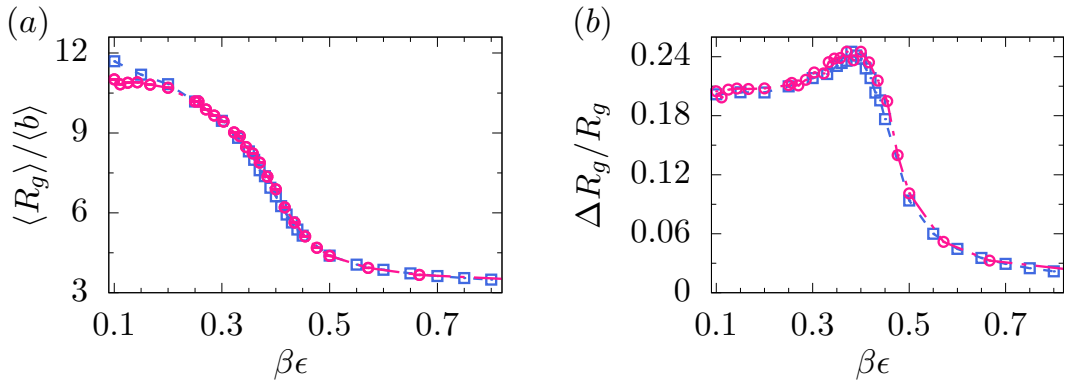


Figure 3.4: (a) The mean radius of gyration $\langle R_g \rangle$ scaled by mean bond length $\langle b \rangle$, plotted as a function of $\beta\epsilon$. The data points in blue $\square$ and pink $\circ$ represent the variation of ϵ and T respectively. (b) The relative fluctuation $\Delta R_g / R_g$ is plotted with $\beta\epsilon$. The data points follow the same notation.

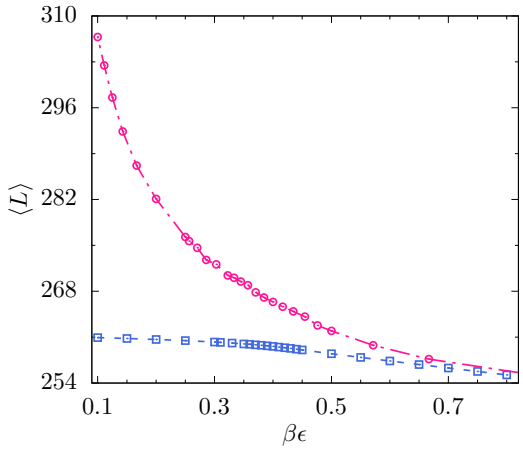


Figure 3.5: The average chain length $\langle L \rangle = (N-1)\langle b \rangle$ is plotted against $\beta\epsilon$. The data points in blue $\square$ and pink $\circ$ represent the variation of ϵ and T respectively.

variation of $\beta\epsilon$ in Fig. 3.4(a) for both the cases. The size of the chain is measured in terms of the average radius of gyration $\langle R_g \rangle$ scaled by mean bond length $\langle b \rangle$ of the chain. The data points from the two set of simulations show good agreement and display a monotonic decrease in the radius of gyration with increase in $\beta\epsilon$. Each data point in the plot is averaged over 10^6 equilibrium configurations which are separated from each other by $10^2\tau$ in time. The corresponding standard errors are smaller than the symbol size

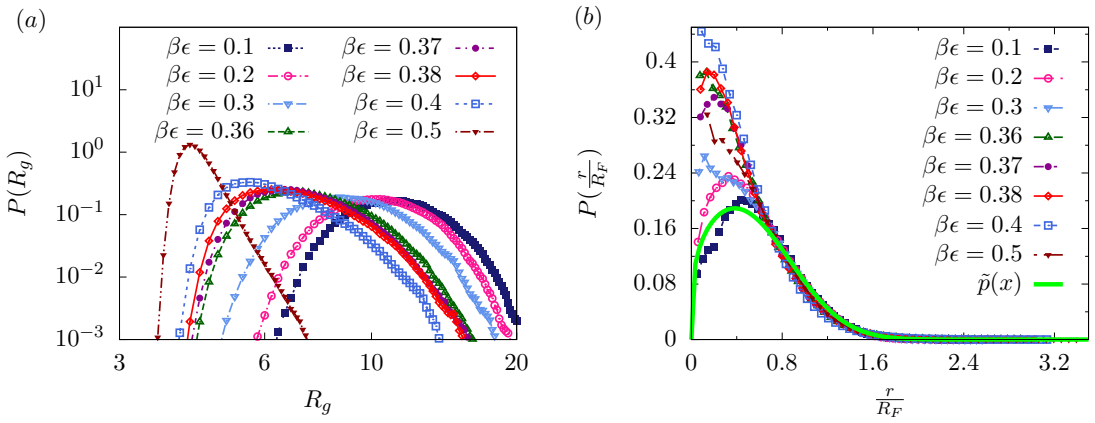


FIGURE 3.6: (*a*) Probability distribution function for R_g is plotted for various $\beta\epsilon$ from 0.1 to 0.5. (*b*) Probability distribution function for end to end separation scaled by its Flory radius, $\frac{r}{R_F}$, is plotted for the same span of $\beta\epsilon$. Function $\tilde{p}(x) = 0.28x^{0.28}\exp(-1.2x^{2.43})$, where $x = \frac{r}{R_F}$.

Fig. 3.4(*b*) shows the relative fluctuation of radius of gyration, $\Delta R_g/R_g = \sqrt{\langle R_g^2\rangle - \langle R_g\rangle^2}/\langle R_g\rangle$ as a function of $\beta\epsilon$. With increase in $\beta\epsilon$, it first increases, attains a maximum and finally decreases. The value $\beta\epsilon = \beta\epsilon^* = 0.38$ at which relative fluctuation attains its maximum is identified as the phase transition point for both set of simulations.

In Fig. 3.5, we show the mean chain length $L = (N-1)\langle b\rangle$ of the polymer as a function of $\beta\epsilon$, where $\langle b\rangle$ denotes the mean bond length. It is clear that increasing T impacts all the relative energy scales, including the bond energy. Thus we find a significant increase in bond length at higher temperatures. Although the change in ϵ impacts the non-bonded potential alone, thereby leaving the mean bond length $\langle b\rangle$ relatively unchanged. The coil-globule transition as shown in Fig. 3.4 is essentially the same for the two methods of changing $\beta\epsilon$. Thus, from now on, we analyze the results obtained from the first set of simulations performed at constant k_BT and changing ϵ.

To characterize the order of the coil-globule transition, we investigate the probability distribution function of polymer size, as in the previous chapter. The radius of gyration and end-to-end separation are two equivalent measures of the size of the polymer. In Fig. 3.6(*a*), the probability distribution of radius of gyration $P(R_g)$ is shown across the phase transition.

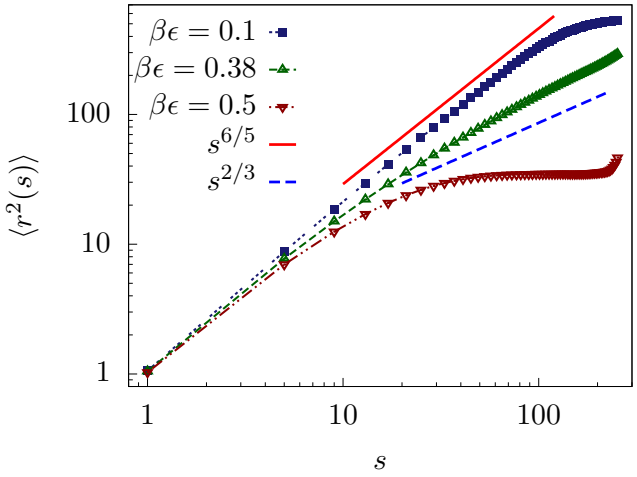


Figure 3.7: Subchain extension, $\langle r^2(s)\rangle$ is plotted at three $\beta\epsilon$, before, after and at the critical point. Flory scaling $\langle r^2(s)\rangle \sim s^{6/5}$ is shown for $\beta\epsilon = 0.1$ (solid red line). Around $\beta\epsilon_* = 0.38$, data follows scaling relation $\langle r^2(s)\rangle \sim s^{2/3}$ consistent with fractal globule estimate (shown in blue dashed line). At $\beta\epsilon = 0.5$, typical characteristic of an equilibrium globe is observed.

From small to large $\beta\epsilon$, the uni-modal distribution of $P(R_g)$ persists. At the phase transition point, $\beta\epsilon_* = 0.38$, the distribution turns out to be the widest.

The other measure of size used is the magnitude of end-to-end vector r. In Fig.3.6(b), the probability distribution is plotted by the quantity $P(r/R_F)$, where $R_F = \langle r^2\rangle^{1/2}$ denotes the Flory radius of the chain. The function $\tilde{p}(x) = 0.28x^{0.28}\exp(-1.2x^{2.43})$, with $x = \frac{r}{R_F}$ denotes the probability distribution of scaled end-to-end separation of self- avoiding chain [53]. We plot it here to compare with $P(r/R_F)$ at $\beta\epsilon = 0.1$ (coil phase). Across the coil globule transition, the distribution $P(r/R_F)$ remains unimodal. The absence of any metastable maximum across the phase transition point $\beta\epsilon = 0.38$ is a signature of the absence of phase coexistence, a characteristic of the continuous transition.

3.3 Local morphologies

3.3.1 subchain extension

First, we analyze the sub-chain extension of the polymer. This was used in the previous chapter. For the sake of completeness, we discuss it once again. The question we ask is, if one considers a chain segment of contour length s, how the average extension $\langle r^2(s)\rangle$ scales with s? Here, we show this in Fig. 3.7 in three regimes of $\beta\epsilon$. At $\beta\epsilon = 0.1$, the sub-chain extension shows $\langle r^2(s)\rangle \sim s^{2\nu}$ with $\nu = 3/5$, in agreement with Flory scaling for self-avoiding chains embedded in 3-D. The power law growth shown by red solid line is shifted for better visibility. With increase of $\beta\epsilon$, the polymer folds and it is reflected by a decrease in ν. At the critical point, $\beta\epsilon = \beta\epsilon_* = 0.38$, the scaling of sub-chain extension $\langle r^2(s)\rangle \sim s^{2\nu}$ with $\nu = 1/3$, consistent with the fractal-globule picture [17, 187]. Further, at the largest $\beta\epsilon$, in the globule phase, $\langle r^2(s)\rangle$ shows an initial growth followed by saturation. Inside the core of the globule, a large density of monomers screens the interaction causing the polymer to behave like an ideal chain [187] so that $\langle r^2(s)\rangle \sim s$ is expected until the boundary of globule is encountered. Multiple reflections from the globule boundary fill up the space uniformly over a contour length $s \sim N^{2/3}$. Thus segments beyond this contour separation are equally probable leading to the saturation in $\langle r^2(s)\rangle$. The small increase at the end of plateauing of $\langle r^2(s)\rangle$ at largest s is due to the fact that interactions does not only bring the contour-wise distant segments close to each other, as a result it may effectively push far apart segments along the contour further away from each other spatially. Experimental data from the FISH technique for yeast chromosome marked at telomere and centromere display emergence of a plateauing behavior similar to that of equilibrium globule [188, 189]. On the other hand, in the span of large genomic separations $s > 10$ Mbp, human chromosome 4 shows resemblance with $\nu = 0.32$ [190, 191], consistent with our estimate at critical point. However, in many cases, noise in the data makes it difficult to clearly differentiate the exponent values between 1/2 and 1/3 [189].

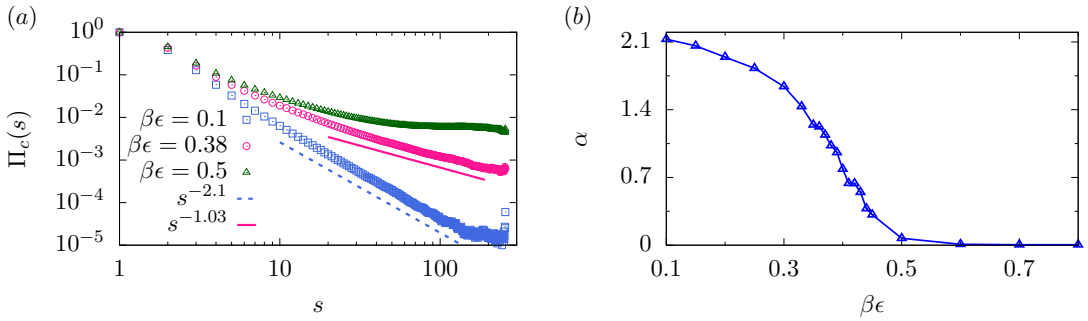


Figure 3.8: (a) The contact probability $\Pi_c(s)$ is plotted for different values of $\beta\epsilon$, where s represent separation along the polymer contour. The power law decay in the asymptotic limit given by $\Pi_c(s) \sim s^{-\alpha}$, is shown at $\beta\epsilon = 0.1$ and 0.38. (b) The asymptotic decay exponent α is plotted with respect to ϵ.

3.3.2 Contact analysis

Once the spatial separation between the contour-wise distant monomers of the chain is less than a cutoff value $r_c = 1.5\,\sigma$, we count a contact between chain segments. The results remain qualitatively same with small variations of r_c. Note that, the chosen value of r_c is smaller than the cutoff separation of Lenard-Jones potential ($U_{ljs}(r)$), which is $r_{thr} = 2.5\sigma$. Thus, monomers participating in contacts always lie within the range of attraction of others. Similar to the results in the previous chapter, the probability of contact formation decays with power law $\Pi_c(s) \sim s^{-\alpha}$ in the asymptotic limit (large s) (Fig.3.8). For the self avoiding chain, earlier results predicted $\alpha = 2.1$ [53], which compares well with our simulation results at small $\beta\epsilon = 0.1$. As $\beta\epsilon$ increases, the monomers of the chain start making frequent contacts with each other, augmenting the value of exponnet α. At the critical point, $\beta\epsilon_* = 0.38$, the exponent $\alpha \sim 1.1$, consistent with the fractal-globule estimate [17, 187]. Interestingly, this exponent is very close to what is measured for human chromosomes in the Hi-C experiments over the genomic separation $0.5 - 10$ Mbp [17]. Further, in the large $\beta\epsilon$ limit, in equilibrium globule phase, contact probability develops a plateau in the asymptotic limit yielding $\alpha = 0$. In Fig. 3.8(b), the change of asymptotic decay exponent α is plotted with $\beta\epsilon$. The non-linear decay in the exponent recaptures the coil-globule transition.

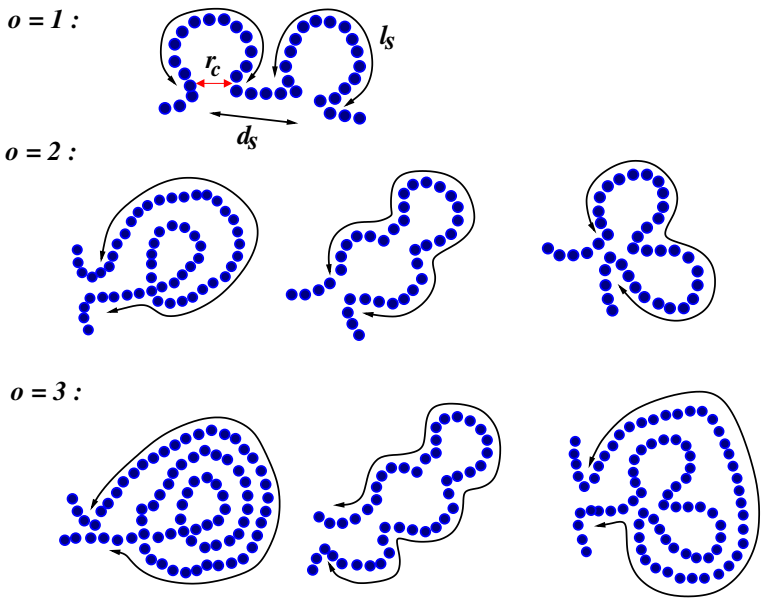


Figure 3.9: Schematics of loop topologies of order o: The monomers of the polymer are shown by blue beads. (a) The red arrow shows cut-off value of separation r_c below which loop forms between monomers. First order loops $o = 1$ are shown. The two first order loops with size l_s are separated from each other by distance d_s along the contour. (b) Three schematics for second order loops $o = 2$ are shown. (c) Three examples of third order loops $o = 3$ are shown.

3.3.3 Loop analysis

We further quantify the morphology of chain in terms of loop structures which are formed due to contact between different monomers. In above, it is already discussed that thermal motion and interaction facilitates contact formation between various different segments of the chain. In the entire space of contacts between monomers, we apply specific criteria to identify them as loops of various order. Schematics for such looped structures are presented in Fig. 3.9. Loops can be simple first order loops or complex higher order loops. Simple loops are defined as follows. Given that two different monomers of the chain make contact with each other and if there are no other contacts during circulation from one end of the contact towards the other end, it is identified as a first order loop $o = 1$. In Fig. 3.9, two such first order loops $o = 1$ are shown having size l_s and separation d_s along the polymer contour. Further, a second order loop, $o = 2$, necessarily embed one or more first order loop

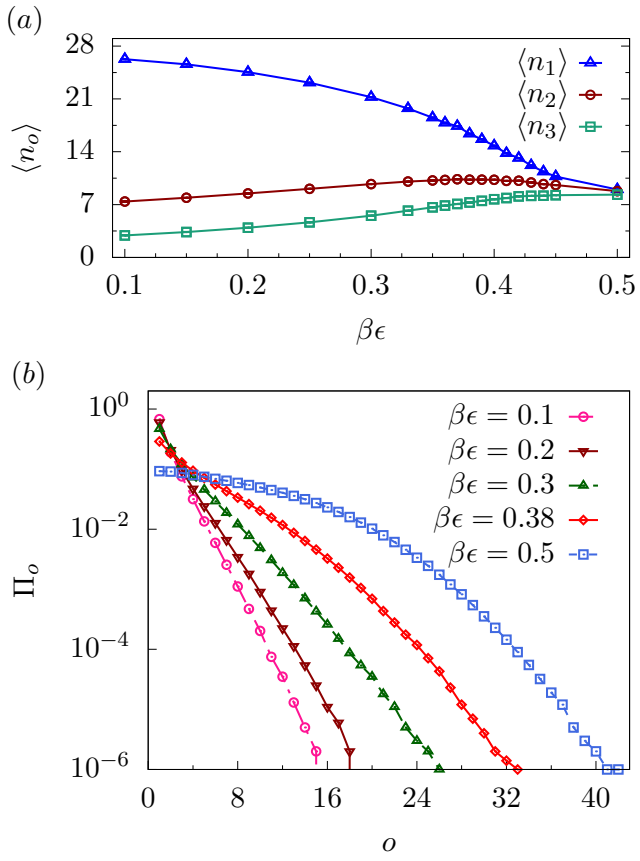


Figure 3.10: (a) The mean number of first, second and third order loops $\langle n_{1,2,3} \rangle$ are plotted with $\beta\epsilon$. (b) Probability Π_o of higher order loops are plotted on semi-log scale at various $\beta\epsilon$. Here, o represents the order of loops.

inside it. It means given a second order loop between two monomers, if one start circulation from one end of the loop and move towards its other end, one must encounter at least one first order loop. Three such examples are shown in Fig. 3.9 by $o = 2$. Further, a third order loop, $o = 3$, embeds at least one second order and one first order loop inside as shown in the figure. This definition can be extrapolated to the n-th order loop which embeds every lower order loops from $(n-1)$-th to $o = 1$. In comparison with previous chapter, there is subtle difference in the identification of loops. In previous chapter, cross-linkers were necessary for loop formation between monomers of self avoiding chain. In that context, all the contacts could not form a loop, because fraction of them are due to thermal motion of chain. Contrary

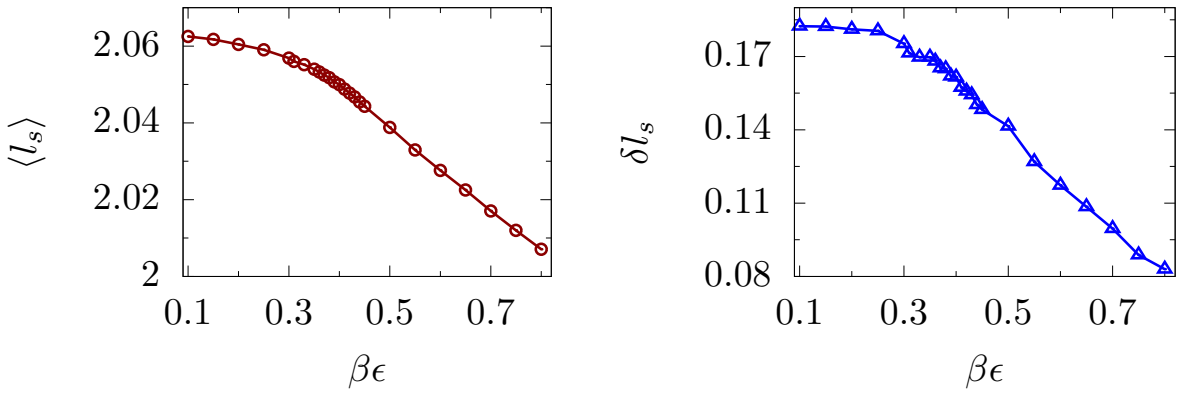


FIGURE 3.11: (a) The mean size of first order loops $\langle l_s \rangle$ is plotted with $\beta\epsilon$. (b) Corresponding fluctuation δl_s is plotted with $\beta\epsilon$.

to that, in the present case, any contact can be called a loop if it satisfies the criterion.

In Fig. 3.10 (a), the number of first, second and third order loops are plotted against $\beta\epsilon$. With increase in $\beta\epsilon$, the average number of first order loops $\langle n_1 \rangle$ decrease in monotonic fashion. This behavior is in contrast to previous chapter result which showed non- monotonic behavior across the phase transition. Further, mean number of second and third order loops, $\langle n_2 \rangle$ and $\langle n_2 \rangle$ respectively, show initial increase that saturates with increase in $\beta\epsilon$. In the globule phase, at $\beta\epsilon = 0.5$, $\langle n_2 \rangle$, $\langle n_3 \rangle$ remains very close to each other, slightly less than $\langle n_1 \rangle$. Despite its decreasing number, first order loops always dominate the loop structures.

Further, Fig. 3.10(b) shows probability Π_o of loops of o-th order on the semi-log scale, where $o = 1, 2, 3, ..., (N-1)$. The normalization $\Sigma_o \Pi_o = 1$ is followed in the numerical evaluation. The plot shows monotonous decrease in Π_o with respect to their order o. At small $\beta\epsilon$, Π_o decays very rapidly before the transition. As polymer folds, this decay slows down leading to relatively higher order loops. In the equlibrium globule at $\beta\epsilon = 0.5$, we find loops upto 42-th order.

Now, we analyze the average size of the first order loops $\langle l_s \rangle$, plotted in Fig. 3.11(a). The mean size of the first order loops decreases as polymer undergoes coil-globule transition. In the open state of polymer at small $\beta\epsilon$, relatively large simple loops can form compared to

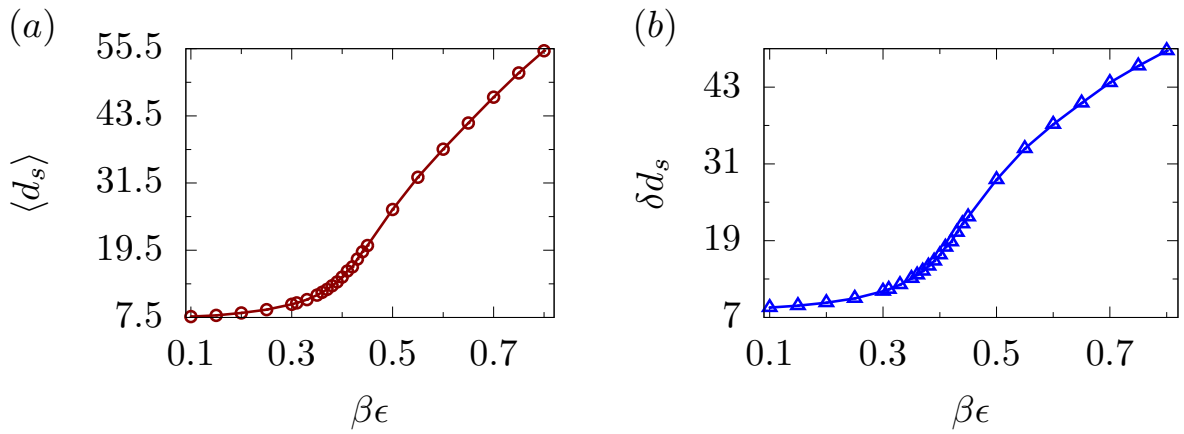


Figure 3.12: (a) The average separation between first order loops $\langle d_s \rangle$ is plotted against $\beta\epsilon$. (b) Fluctuation of mean separation between first order loops, δd_s is plotted against $\beta\epsilon$.

collapsed state in which higher order contacts impede the formation of simple large loops. Due to attractive interaction between the non-bonded monomers of the chain, the monomers separated by small distance along the chain contour seems to constitute majority of the first order loops causing small $\langle l_s \rangle$ across the span of $\beta\epsilon$. In Fig. 3.11(b), the fluctuation in the first order loop size, $\delta l_s = \sqrt{\langle l_s^2 \rangle - \langle l_s \rangle^2}$ is plotted against $\beta\epsilon$. Like $\langle l_s \rangle$, δl_s also decrease monotonically across coil-globule transition. Reduced fluctuations connotes that the loops are more stable in the globule phase. Further, in Fig. 3.12(a), we have plotted the mean separation $\langle d_s \rangle$ between the first order loops. With increase in $\beta\epsilon$, the mean separation $\langle d_s \rangle$ show monotonous increases. As polymer folds, due to formation of higher order loops, the separation between the first order loops increases. In Fig. 3.12(b), fluctuation in separation $\delta d_s = \sqrt{\langle d_s^2 \rangle - \langle d_s \rangle^2}$ as a function of $\beta\epsilon$ is plotted. It shows monotonic increase across the coil- globule transition. The magnitude of mean separation $\langle d_s \rangle$ and corresponding fluctuation δd_s are of same order.

3.3.4 Zippering

The attractive interaction between monomers of the chain can zipper its various segments. The quantification of zipper fraction on the chain is similar to previous chapter. It is given

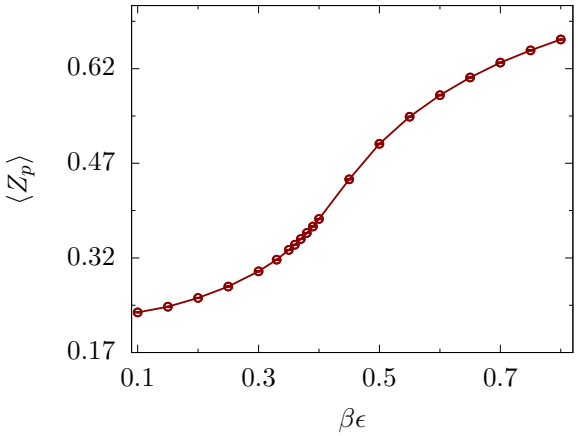


FIGURE 3.13: The zipper fraction, Z_p of the chain is plotted versus $\beta\epsilon$.

by, $Z_p = N_z/N$, where N_z is the total number of monomers participating in the zipper and N is the total number of monomers in the chain. In Fig. 3.13, we have plotted the mean zipper fraction Z_p with increasing $\beta\epsilon$. As polymer folds, Z_p increase monotonically in the sigmoidal fashion. In the coil phase at $\beta\epsilon = 0.1$, $Z_p = 0.23$, which increased to $Z_p = 0.35$ at the critical point and finally saturates to $Z_p = 0.67$ at $\beta\epsilon = 0.5$ in the globule phase.

3.3.5 Contact maps

We present average contact probability measure between various monomers of the chain in terms of contact maps. Such maps are shown in Fig. 3.14 at different $\beta\epsilon$. Here, averaging is performed over 10^6 equilibrium conformation. Criterion for contact formation is same as mentioned above, *i.e*, separation between the spatially close monomers must be less than 1.5σ. In the map, n_m denotes the monomer index from 0 to $N-1$, where N is the total number of monomers in the chain. Each plot is associated with a color code quantifying the contact probability. At small $\beta\epsilon$, only contour-wise neighboring monomers participate in the contact formation. As polymer keeps on folding, contour-wise separated monomers also begin to make contact with each other. Thus, at the critical point $\beta\epsilon = 0.4$, map shows percolation of contacts for the countour-wise distant monomers compared to $\beta\epsilon = 0.1$. Further, in the

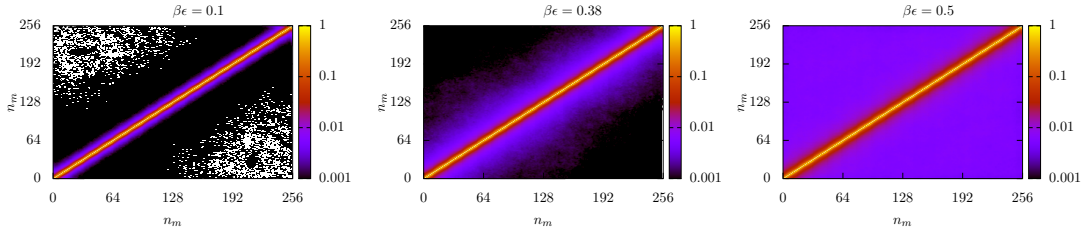


FIGURE 3.14: Left to right, contact map is plotted for $\beta\epsilon = 0.1, 0.38, 0.5$ respectively. n_m represents the monomer index along the chain. For color coding, log scale has been used.

globule phase at $\beta\epsilon = 0.5$, the contacts proliferate to all the monomers of the chain filling the entire map.

Our contact maps show uniform polymer organization across the coil- globule transition due to homogeneous nature of interaction between the monomers. Contrary to that, the Hi-C maps for various chromosomes exhibit checker board like pattern and topologically associating domains (TADs) [6, 84, 175, 192]. Within such domains, chromosome interacts more frequently compared to other parts of the chain, however checker board patters are due to cross- interaction between the domains. Previous studies have opted for heterogeneous sequence specific interaction along the chromatin with binder proteins which enabled them to capture similar territorial structures in the contact map [104]. In the following, we introduce a hetero- polymer model for chromatin to explore the effect of sequence specific interactions between monomers of chain.

First, we compartmentalize our homo-polymer chain with $N = 256$ in four segments such that each segment has 64 monomers. We have considered two kind of hetero-polymer models for chromatin. Schematics of first is shown in Fig. 3.15(a), in which each segment consists of monomers of certain type denoted by A, B or C. Each type is represented in different colors, i.e., type A monomers are shown in blue, type B in red and type C in green. The monomer index 0–63, 64–127, 128 to 191, and 192–255 denote segments consisting of monomer types A, B, C and B respectively. The heterogeneity is realized by tuning the interaction within and between different types, which is summarized in the following table.

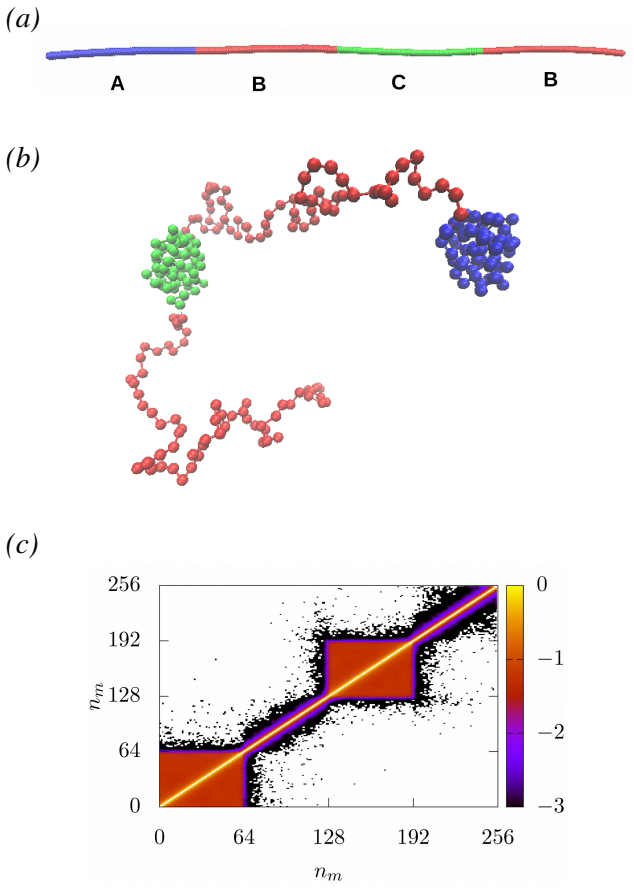


FIGURE 3.15: (a) Hetero-polymer consisting of total $N = 256$ monomers is compartmentalized in four sectors, each consisting 64 monomers. Monomers are represented by type A,B and C, which are shown in blue, red and green respectively. (b) Snapshot of the simulation is shown. (c) Contact map at $\beta\epsilon = 1.0$ is shown.

type	interaction
A-A	LJ
A-B	WCA
A-C	WCA
B-B	WCA
B-C	WCA
C-C	LJ

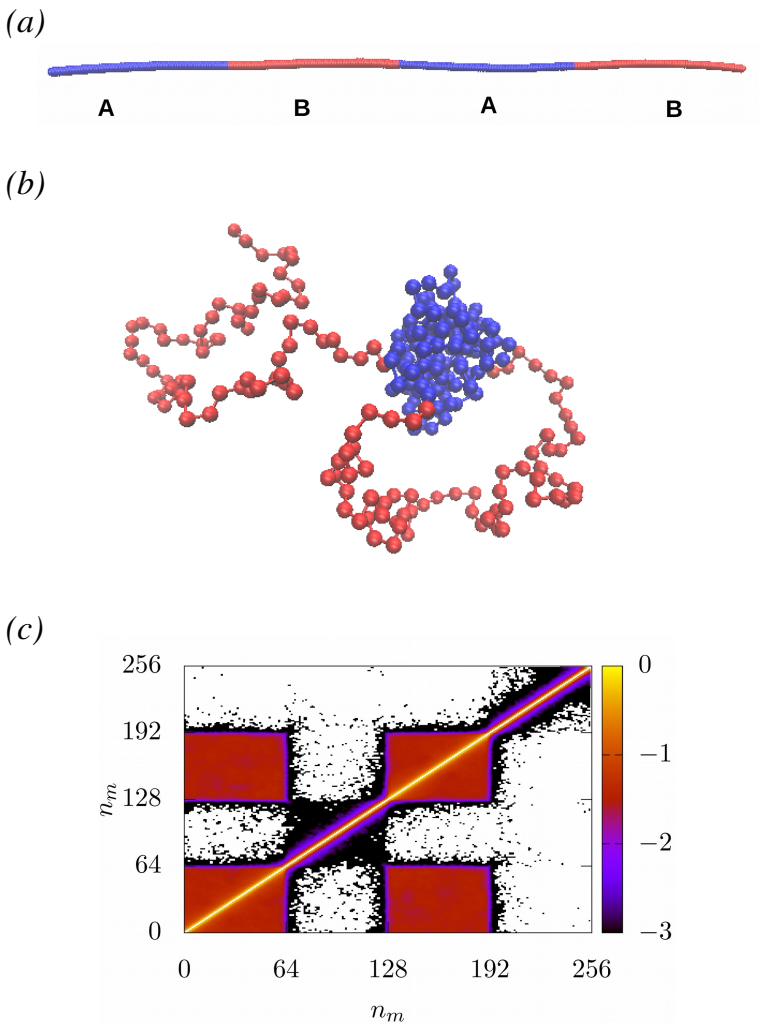


Figure 3.16: (a) Hetero-polymer consisting of total $N = 256$ monomers is compartmentalized in four sectors, each consisting 64 monomers. Monomers are represented by type A and B, which are shown in blue and red respectively. (b) Snapshot of the simulation is shown. (c) Contact map at $\beta\epsilon = 1.0$ is shown.

LJ stands for the Lennard-Jones potential with attractive tail, and WCA is abbreviation for the purely repulsive Weeks-Chandler-Anderson potential [157]. For the Lennard-Jones potential, we used the same form as mentioned in Eq. 3.2 with $\epsilon = k_B T$. The WCA potential

is defined as,

$$\begin{aligned} U_{WCA} &= 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6 + 0.25], \text{ for } r < 2^{1/6}\sigma \\ &= 0 \qquad \text{otherwise.} \end{aligned} \tag{3.8}$$

For WCA, we use the strength $\epsilon = k_B T$. Here, LJ potential provides attraction between monomers, and WCA potential imposes short range repulsion. All the other details of simulation are the same as in our homo-polymer model. At equilibrium, parts of the chain interacting through LJ potential, collapse. In Fig. 3.15(b), the snapshot of the simulation shows, type A shown in blue beads and type C shown in green beads collapse separately to local globules, separated by coils of B. To analyze the nature of contact formation between the monomers of the chain, we sampled the data over 10^5 equilibrium configurations which are separated from each other by $10^3\tau$ in time. Inside each globule, monomers interact frequently with each other compared to other parts of the chain. This gives rise the two separate domain in the contact map, one for blue monomers and other for green monomers. Apart from that, the other interactions are repulsive leading to small number of encounters between the beads of other colors including the inter-color monomer encounter (red). The well- separated rectangular regions of high contact probability are reminiscent of observed topological domains in bacteria [193] and eukaryotes [80].

In the second hetero- polymer model, we compartmentalized the chain in a different fashion, namely A,B,A,B as shown in Fig. 3.16(a). Again, Each sector consists of 64 monomers. Monomers of type A and B are colored in blue and red respectively. The interaction is summarized in following table.

type	interaction
A-A	LJ
A-B	WCA
B-B	WCA

Parameters for Lennard-Jones (LJ) and Weeks-Chandler-Anderson (WCA) potential are same as previous model. As a result of simulation, blue monomers in first and third sector collapsed into one globule while the red monomers form self avoiding chain segments that remain open 3.16(b). Due to collapse of monomers in first and third sector in one globule, three square patches on the contact map appeared 3.16(c). The first patches on the diagonal is due to frequent interaction of monomers in sector one (first blue sector) with index 0 to 63. The second patch along the diagonal is due to frequent interaction of monomers in sector three (second blue sector) with index 128 to 191. The patches away from diagonal region are due to cross interaction between blue monomers in first and third sector of the polymer. The red monomers do not have frequent encounter between themselves as well as with blue monomers, hence contacts between them are nearly absent in the map. The specificity of the interaction leads to the compartmentalization which is manifested very clearly in the contact map.

3.4 Conclusion

We have considered a biophysical model for chromatin fiber in terms of a self-avoiding flexible polymer with effective inter-segment attraction. In the cell, chromatin fibers are associated with proteins like NAPs, transcription factors etc, which form bridges between chromatin segments. To capture such effective attraction, we introduced an attractive tail between spatially non- neighboring monomers. Increasing the effective attraction between chromatin segments led to continuous coil- globule transition. Along the transition, local morphological changes of chromatin are analyzed in terms of contacts, loops and zippers. In the coil phase, sub- chain extension and contact probability captures the behavior of self-avoiding chain. However, at the critical point, they show results similar to average human chromosome behavior at various genomic separation [17]. In the collapsed state, typical equilibrium globule results are captured. The observed behavior of sub- chain extension and

contact probability are consistent with that in the previous chapter.

The number of first order loops show monotonic decrease across the coil-globule transition. Despite their decreasing number, the simple loops always dominate the loop topology in consistence with previous chapter. The mean size of first order loops and its fluctuation show monotonic decrease as the polymer folds. The separation between simple loops and its fluctuation show monotonic increase along the phase transition. The mean number of second and third order loops increases and saturates along the transition, a behavior consistent with previous chapter. In addition, with folding of polymer, the probability of complex higher order loops increase recapturing the behavior due to cross-linkers. The zipper fraction of chain increase along the coil- globule transition, being qualitatively consistent with previous chapter.

The contact maps show spreading of contacts as the polymer folds, however, the homo-polymer model does not exhibit any territorial organization due to homogeneous nature of interaction between chromatin segments. Capturing the sequence specific aspects of chromatin DNA-binding protein interaction, our hetero- polymer models show formation of domains like the topologically associated domains and the characteristic checker board patterns, typical to Hi-C maps of various chromosomes [80, 193].

4 Feather-boa model of bacterial chromosome

In bacterial cells a long circular DNA and associated proteins constitute the chromosome. The chromosomal structure is suspended in a crowded cytosolic fluid and forms a membrane-less organelle called the nucleoid occupying a central sub-volume of the cell [194]. The chromosome along with the cytosolic fluid and other components of the cell are confined by cell envelope which for a large class of bacteria like E. coli or B. subtilis has a cylindrical geometry. In E. coli bacteria, a 4.6 Mbp (1.6 mm) long negatively supercoiled circular DNA forms the nucleoid which occupies a nucleoid subvolume of the cylindrical confinement of the cell having diameter 0.8 μm and length 2-4 μm. The 1.6 mm long DNA has to fold about 10^3 times in order to fit inside the confinement [124, 126]. This huge compaction has to be concomitant with functional organization of chromosome because the chromosome has to express genes, regulate expression, and allow replication in order to survive [127].

There are at least four major physical and chemical mechanisms through which such compaction and organization are usually achieved [195, 196]. First, the cellular confinement itself. However, it is not all. A fact that can be appreciated by noting that the nucleoid only occupies a sub-volume of the cell, with a ratio of the nucleoid and cellular volume being less than one-fourth [126]. Second, the DNA is almost 5% negatively supercoiled, and this globally maintained, enzymatically controlled under-twisting results in warps and folds of various segments of the chain leading to the folded hair-braid like structures known as plectonemes [15]. Formation of plectonemes substantially decreases the effective size of the chromosome. Electrostatic zippering of DNA segments also contributes to compaction.

Third, simultaneous binding of nucleoid associated proteins (NAPs) at more than one location along the genome form loops and compacts the structure [125–127, 129, 197]. Impact of passive cross-linking has been discussed in the previous chapters. For the sake of completeness, we reiterate a few major points here. The NAPs are mainly of two categories: non- specific like HU which binds to DNA without any bias, and specific like FIS which binds to only a few target genetic sequences. The resultant loop formation leads to further compaction of the chromosome. There are dedicated proteins that convert these looped structures into topological associating domains [84, 127, 177, 198]. Further, structural modification of chromosome (SMC) proteins can lead to active loop extrusion. In human chromosomes, cohesin and CTCF protein complexes generate and maintain such loops [199]. In bacteria, the condensin SMC extrudes loops against barriers of RNA- polymerase or transcription machinery [86]. Electron microscopy experiments on lysed *E.coli* chromosomes showed a bell-shaped distribution of loops with the maximum close to 10-12 kbp [26]. This suggests a typical loop size of 10-12 kbp. With the advent of chromosome conformation capture (CCC) based techniques, it has been possible to directly probe the physiological interaction frequencies between different gene segments in the chromosome [144, 145]. These techniques allowed to form contact maps that exhibit probability of spatial contacts between various gene- location along the genome and one can see the presence of TADs in form of characteristic checker- board patterns on the 2-D map. Fourth, the depletion effect due to molecular crowding generated by cytosolic components can potentially compact the chromosome [200–202]. Within an equilibrium picture, the overall entropy of a system consisting of a polymer and molecular crowders increases by allowing more volume to the large degrees of freedom of the crowders, effectively compressing the polymer [203, 204]. Such entropic forces are expected to play an important role in chromosome compaction, and segregation of the chromosome from protein crowders [11, 205–209].

The high resolution live- cell imaging experiments over the last decade revealed an emergent helical organization of bacterial chromosome characterized by a few $(3-4)$ turns. It is

ubiquitous in rod-shaped bacteria [9, 12, 13]. They exhibit a definite pitch to length ratio. The helical shape of the nucleoid is closely related to the cylindrical cell- shape. Modification of the cell shape by impeding the cell wall formation led to a change in the nucleoid morphology [12]. Using an argument of optimal packing, it was shown that thick chains in a cylinder pack best as a helical object [210]. A polymer model with local loop structures while confined inside cylindrical confinement, showed emergent helical shape due to entropic forces in confinement [18]. The impact of confinement on the chromosome size and shape has been analyzed recently in growing bacteria [9, 10].

The effect of confinement on the size of polymer has long been appreciated in polymer physics [53, 211]. The blob picture due to de Gennes suggests that once the confining diameter D of the cylinder becomes smaller than the radius of gyration R_g of a self-avoiding polymer of length L, it starts to behave like a connected array of blobs of size D. Within the blob, the Flory scaling for free chain works, such that $D \sim n^{3/5}\, b$ assuming that each blob contains n segments of size b. The total number of blobs are N/n. Thus the effective chain extension inside the confinement increases as $(N/n)\, D \sim D^{-2/3}$ with decreasing D [53]. However, such a tube confined polymer does not display helical structures observed in bacterial chromosome [12, 13].

The helical shape emerges due to the interplay of the local loop morphologies of the chromosome, and the cylindrical confinement of the cell. In this chapter, we present the feather-boa model of chromosome, a model similar to the bottle- brush polymer [212–214], with polymeric loops attached to a backbone chain [18, 19]. Within this model, the entropic repulsion between the side- loops generate an effective bending rigidity to the backbone over micron length scales, much longer than the DNA persistence length. Further, the side-loops gives rise to an effective thickening to the backbone. The enhanced stiffness and thickening of the backbone can potentially lead to a helical shape to the cylinder- confined polymer [18].

In this chapter, within a given confinement, we study the impact of changing side loop size on the chromosome morphology. As has been mentioned earlier, the cytosolic environment is

fluidized by metabolic activity. We model this environment as a Langevin heat bath of free flowing fluid maintaining a constant effective temperature. The helical shape of the model chromosome emerges only beyond a threshold size of side loops.

4.1 Feather boa model : change in side loop size

Here we present a coarse-grained feather-boa model of the full chromosome. We define the feather-boa chain as a self-avoiding polymer consisting of n_b monomers in the backbone attached to side loops of n_s monomers. The linear chain of n_b monomers constitute the backbone of the feather boa chain. The real bacterial chromosome shows a distribution of loop sizes, however, with a maximum corresponding to a typical loop size [26]. Within our model, n_s monomers of side-loops stand for this typical size. The total number of monomers constituting the model chromosome is $N = n_b \cdot (1 + n_s)$. The bonds of the feather-boa chain are maintained by the harmonic interaction,

$$V_b = \frac{A}{2}(\mathbf{d_i} - \sigma \mathbf{u_i})^2, \tag{4.1}$$

where $\mathbf{d_i} = \mathbf{r_{i+1}} - \mathbf{r_i}$, with $\mathbf{r_i}$ denoting the position of i-th bead. The self-avoidance is incorporated via a short ranged WCA repulsion between the non-bonded monomers,

$$\beta V(r_{ij}) = 4\epsilon[(\sigma/r_{ij})^{12} - (\sigma/r_{ij})^6 + 0.25] \tag{4.2}$$

for $r_{ij} < 2^{\frac{1}{6}}\sigma$ and $\beta V(r_{ij}) = 0$ otherwise. Here $r_{ij} = |\mathbf{r_{i,j}}|$ with $\mathbf{r_{i,j}} = \mathbf{r_i} - \mathbf{r_j}$ the separation vector between i-th and j-th monomers [157]. The length and the energy scales of the problem are set by σ, and ϵ respectively, and together they set the time scale $\tau = \sigma\sqrt{\frac{m}{\epsilon}}$. We choose $A = 100\epsilon/\sigma^2$ to keep the bond- length fluctuations within 1%.

The cylindrical wall along with the top and bottom surfaces of the cylinder repel all the

monomers with short range interaction

$$\beta V_{wall} = 2\pi\epsilon \left[\frac{2}{5} \left(\frac{\sigma}{r_{iw}} \right)^{10} - \left(\frac{\sigma}{r_{iw}} \right)^{4} + \frac{3}{5} \right], \tag{4.3}$$

when $r_{iw} < \sigma$ and $\beta V_{wall} = 0$ otherwise. Here, r_{iw} is the separation of i-th monomer from the wall. The cylindrical confinement is realized with diameter $D = 29.5\sigma$ and length $L = 50.74\sigma$. We perform molecular dynamics simulations following velocity-Verlet algorithm in the presence of a Langevin heat bath. The Langevin thermostat is characterized by an isotropic friction constant $\gamma = 1/\tau$ fixing the temperature at $T = 1.0\epsilon/k_B$ as implemented by ESPReSo molecular dynamics package [161]. The numerical integrations are performed choosing time steps of size $\delta t = 0.005\tau$. The simulations are performed over 10^7 to 5×10^7 steps to equilibrate the feather-boa chain. For the analysis of equilibrium properties, data are collected over further 10^4 configurations separated from each other by $5 \times 10^3\tau$.

We fix the number of monomers in the backbone, $n_b = 200$ and vary the number of monomers in the side loops n_s as the tuning parameter. In Fig. 4.1, a few representative equilibrium configurations are plotted for the feather- boa chain with side loop sizes $n_s =$ 4, 8, 12, 16, 20, 24, 32, 36 and 40, which corresponds to total number of monomers $N =$ 1000, 1800, 2600, 3400, 4200, 5000, 6600, 7400 and 8200 respectively. The corresponding monomer densities are $\rho_m\sigma^3 = 0.03,\ 0.05,\ 0.07,\ 0.1,\ 0.12,\ 0.14,\ 0.19,\ 0.21$ and 0.24. The blue beads in Fig. 4.1 represent the backbone of the feather- boa chain and the green beads denote the side loops. For small n_s, the chain opts irregular shape, however, for $n_s \geq 16$, clear helical organization of backbone monomers emerges. The helical morphology of the backbone of feather-boa chain is quantified in the following.

4.1.1 Backbone helicity

The degree of helical organization of the chromosome can be quantified in terms of the tangent-tangent correlation function $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$ of unit bond vectors $\mathbf{u}_i$ along the backbone

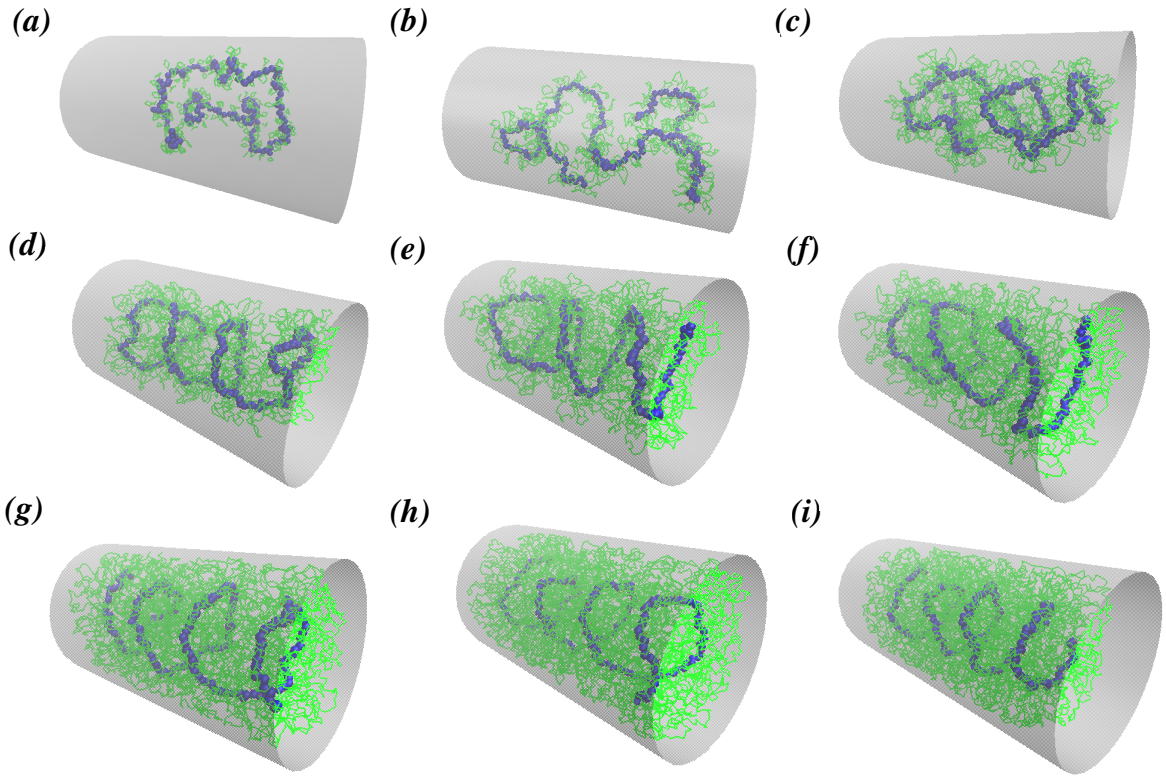


FIGURE 4.1: Snapshots of the simulation at various n_s are presented. The backbone monomers are shown in blue and side loops are in green.(a)-(i), n_s value correspond to 4, 8, 12, 16, 20, 24, 32, 36 and 40 in alphabetical order.

contour of the feather-boa chain separated by $s = |i - j|\sigma$ (Fig.4.2). First, we project the polymer configurations on the 2-dimensional radial plane of the cylinder denoted by xy-plane, assuming z is along the long axis of cylinder. The unit tangent vector $\mathbf{u}_i$ is calculated from the position vectors of this projected configuration of consecutive beads in the chain. We perform averaging over 10^4 equilibrated polymer configurations separated from each other by $5 \times 10^3\tau$. We compute $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$ for various side loop sizes n_s. For better visibility, we show the correlation functions in two separate plots. In Fig. 4.2(a), the correlation functions are shown for smaller n_s, and in Fig. 4.2(b) the same is shown for larger n_s. As we increase n_s beyond $n_s = 16$, the configuration starts to adopt a helical shape captured by periodic

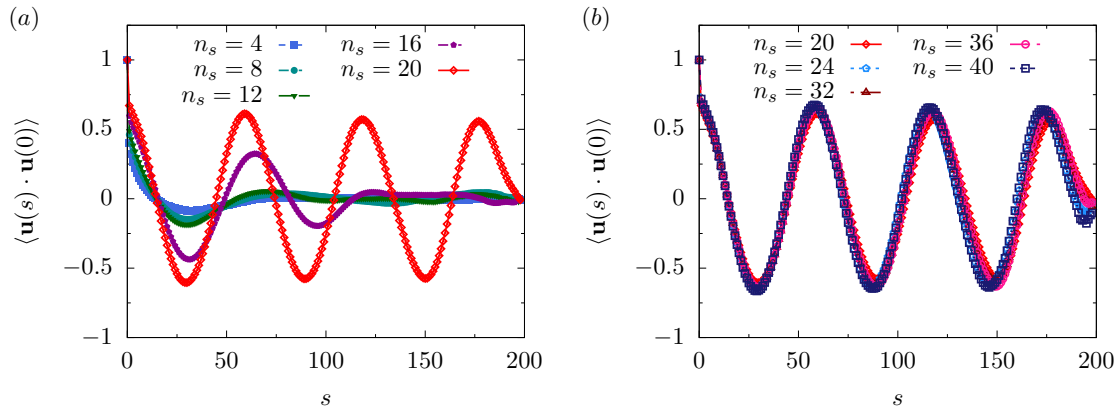


FIGURE 4.2: The tangent-tangent correlation, $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$ along the backbone of the chain is shown where s represents the separation along the backbone. (a) shows plot at smaller n_s compared to (b).

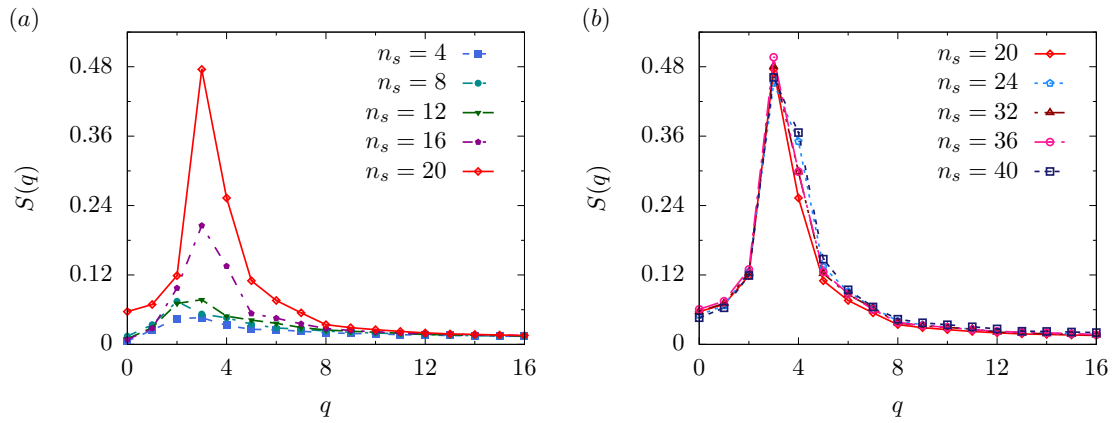


FIGURE 4.3: Structure factor $S(q)$ for $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$ is plotted against q. (a) shows the plot for smaller n_s values and (b) shows the same at larger n_s values.

oscillations in the correlation function $\mathbf{u}(s) \cdot \mathbf{u}(0)$ with separation s. The periodicity of the oscillation captures the helical pitch.

To quantify the number of turns in the backbone helix and the degree of helicity, we compute the structure factor $S(q)$ corresponding to tangent-tangent correlation functions $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$ at different n_s, which are shown in Fig. 4.3. The structure factor is evaluated by performing Fourier transform of the correlation function, $S(q) = \int_{-l_b/2}^{l_b/2} \langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle e^{isq} ds$. It

shows a peak at $q = q_p$ corresponding to the turn number $n_t = q_p$ as is displayed in Fig. 4.3(a) and (b). The amplitude of $S(q)$ at its maximum $q = q_p$ gives the degree of helicity, while the width of the peak is a measure of the statistical dispersion of the structure. In Fig. 4.3(a), increasing the side loop size (n_s), leads to a shift in the peak position q_p towards higher values. At small n_s (=8) the peak is at $q_p = 1$, which increases to $q_p = 3$ at $n_s = 16$. Following Fig. 4.3(a) and (b), we observe that for $n_s \geq 16$, the peak position remains at $q_p = 3$, which implies that the number of helical turns $n_t = 3$ remains unchanged. These three turns can be directly counted from the number of periodic oscillations in Fig.4.2 for $n_s \geq 16$. This $n_t = 3$ impliees a helical pitch $\lambda_{max} = l_b/n_t \approx 66.33\,\sigma$.

Note that the amplitude $S(q = q_p)$ increases with n_s in a monotonic fashion for small side loop sizes ($n_s \leq 20$) as shown in Fig. 4.3(a). At larger n_s, the difference in amplitudes diminishes and they almost overlap with each other (Fig. 4.3(b)). Thus the degree of helicity increases with n_s to saturate. The smallest value of n_s at which we observe a clear peak in $S(q)$ is at $n_s = 16$. In addition, a large difference between amplitudes $S(q_p)$ at $n_s = 12$ and 16 captures the emergence of backbone helicity near $n_s = 16$.

The above results establish that a threshold value of side loop size is necessary for the formation of backbone helix. With the given parameter values for confinement and backbone size, this value is $n_s = 16$. In the model section, it is mentioned that all the monomers in the feather- boa chain repel themselves with short range WCA potential. Due to this repulsion between the side loop monomers, a resistance towards bending emerges on the backbone of feather- boa chain, as a bent configuration leads to increase in the probability of overlap between two side-loops. Smaller side loops can not provide enough repulsion thus lacks in this emergent bending rigidity [18]. The effective bending rigidity becomes significant for longer side loops. Moreover, the side loops connected to backbone have to pack inside the cylinder impacting the conformational entropy of the feather-boa chain. The combination of emergent bending rigidity over short contour separation, and packing energy cost over longer contour separations lead to the observed helical organization of the feather-boa chain [18, 19].

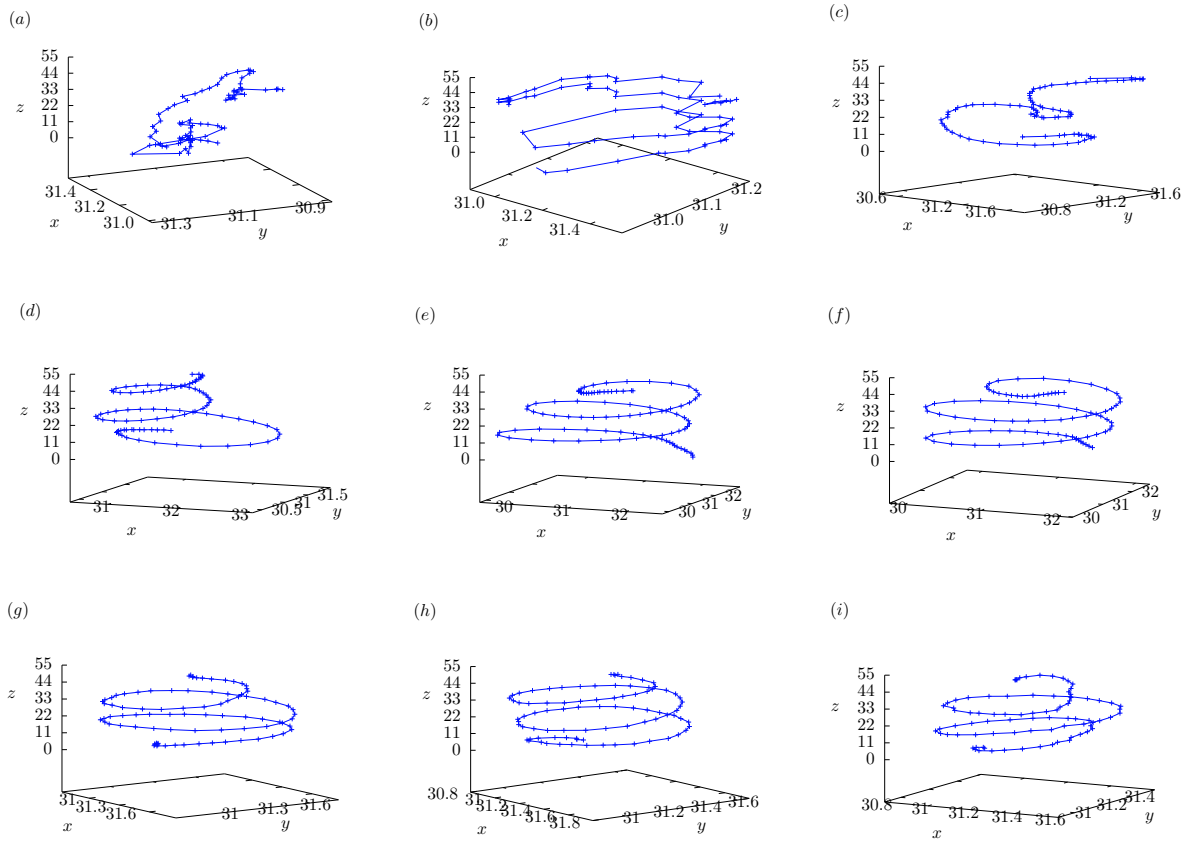


FIGURE 4.4: Center of mass for all the monomers are plotted at different n_s. Plots are at n_s value, 4 (a),8 (b), 12 (c), 16 (d), 20 (e), 24 (f), 32 (g), 36 (h) and 40 (i) respectively.

4.1.2 Center of mass organization

To further analyze the emergent morphology of the model chromosome, in this section, we consider the center of mass loci for all the monomers along the length of the cylindrical confinement. For this purpose, we divide the cylinder into small disk-like bins of width 0.5σ along its length which we assume to be along the z-direction. The xy-plane is perpendicular to the long axis of the cylinder. Unless stated otherwise, we will use this convention and binning procedure to compute the local properties of the feather-boa chain. The center of mass of all the monomers, irrespective of whether they correspond to the backbone or side-loops, is evaluated within each of these bins. Then, averaging is performed over the 10^4

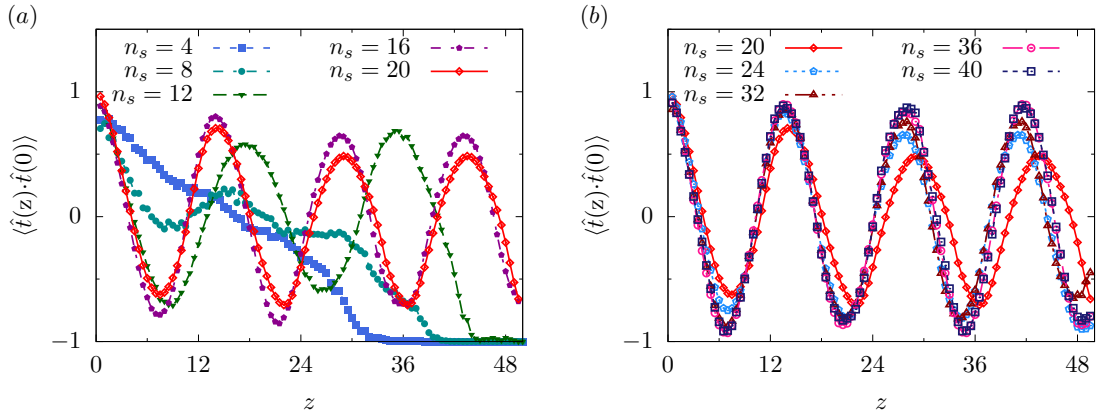


FIGURE 4.5: The correlation $\langle \hat{t}(z) \cdot \hat{t}(0) \rangle$ of the $xy-$ plane projected average center of mass orientation of monomers are shown along the length of the cylinder. (a) shows the plots for smaller n_s values compared to (b).

equilibrated configurations as stated above. The resultant center of mass loci for monomers of the feather- boa chain is shown in Fig. 4.4. The plots in (a)-(i) are at side loop size $n_s = 4, 8, 12, 16, 20, 24, 32, 36$ and 40, respectively, in alphabetical order. At small side loop size ($n_s \leq 8$), the average center of mass loci show irregular organization which are shown in Fig. 4.4 (a) and (b). With increase in n_s, the center of mass loci start to organize into a helical shape. At $n_s = 12$, Fig. 4.4 (c) shows a weak helical organization of monomer center of mass loci. For $n_s \geq 16$, helical organization of average center of mass loci emerges very clearly.

To further quantify this organization of average center of mass loci of monomers, we compute the tangent-tangent auto-correlation function along these loci. The center of mass loci of each polymer configuration is first projected on the $xy-$ plane. The tangent auto-correlation $\langle \hat{t}(z) \cdot \hat{t}(0) \rangle$ is computed along the projected loci, which is shown in Fig. 4.5. Here z-denotes different bins along the long axis of the cylinder. The tangent $\hat{t}(z)$ is a $2d$ unit vector calculated using center-of-mass loci in consecutive bins. The averaging $\langle \ldots \rangle$ is performed over all the equilibrated configurations. For better visibility, we present the correlations in two separate figures. Fig.4.5(a) shows plots at small side loop size and

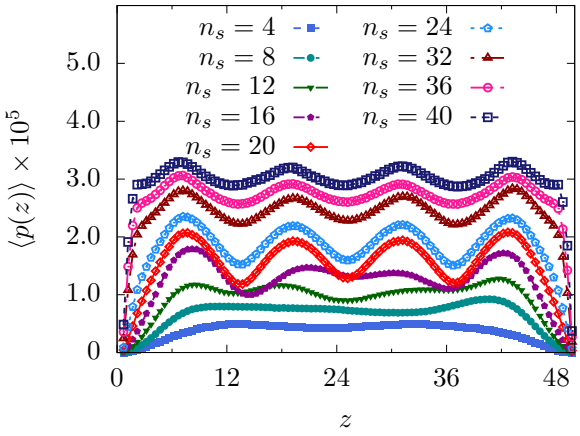


FIGURE 4.6: Probability density of monomers is plotted along long axis of cylinder at various n_s.

Fig.4.5(b) shows them at relatively large n_s. At small n_s, the tangent auto-correlation does not show any regular pattern which suggests random organization of the center of mass loci of the polymer. This can be verified visually by inspecting Fig. 4.4 (a) and (b). Increasing the side loop size to $n_s = 12$ leads to emergence of periodic oscillations in $\langle \hat{t}(z) \cdot \hat{t}(0) \rangle$ for the first time. Corresponding to this particular n_s value, the mean center of mass loci in Fig. 4.4 (c) shows a weak helical organization. Further increase in side loop size beyond $n_s \geq 16$ yields very clear periodic oscillations of the tangent auto-correlation function (Fig. 4.5 (a) and (b)). This nice periodic oscillation suggests helical organization of center of mass loci which can be observed in Fig. 4.4 (d-i).

4.1.3 Local monomer density

In this section, we analyze the spatial organization of the polymer in terms of the local monomer density along the cylinder. In Fig. 4.6, we show the variation of mean probability density $\langle p(z) \rangle$ of monomers belonging to the feather boa chain along the long axis of the cylinder. To evaluate this quantity, we used the bins along the long axis of the cylinder, as described in the previous section. The mean number density of monomers in each bin is

calculated. $\rho(z)$ represents this monomer density in a particular bin denoted by its position z. Assuming a particular bin has n' monomers in one configuration, then $\rho(z) = n'/\pi D^2 l'/4$. Here, diameter of the cylinder and each bin is $D = 29.5\sigma$ and the bin size is $l' = 0.5\sigma$. For any polymer configuration at a particular time, $\int_0^L \rho(z)dz = N$, where N is the total number of monomers in the feather- boa chain. Here we have $N = 8200$. We perform the averaging $\langle\rho(z)\rangle$ over equilibrated polymer configurations for all the bins along long axis of cylinder, and define the normalized probability density $\langle p(z)\rangle = \langle\rho(z)\rangle/N$. This yields the the normalization, $\int_0^L \langle p(z)\rangle dz = 1$. Like in previous sections, 10^4 equilibrated polymer configurations are used for the evaluation of $\langle p(z)\rangle$. Fig. 4.6 shows, at small side loop size ($n_s \leq 8$), $\langle p(z)\rangle$ has approximately constant value in the central region of cylinder, which decreases to 0 as we approach towards the two caps of the cylinder ($z = 0$ and $L = 50.74\sigma$). To increase configurational entropy, polymers avoid the wall. At smaller n_s values, away from the two caps, the density remains roughly uniform. At a larger n_s, a density modulation starts to appear. At further higher $n_s\,(\geq 20)$ we find nice periodic oscillations in $\langle p(z)\rangle$. A comparison with the discussion of helical organization in previous sections show that this onset of density modulation is associated with the emergence of helical organization of the model chromosome.

4.2 Conclusion

In this chapter, we have investigated the impact of local loop structures on the chromosomal organization of E. coli bacteria. We presented the feather-boa model, a coarse-grained model of the bacterial chromosome as a self-avoiding chain dressed with side loops. We used a cylindrical confinement to model the shape of rod-like bacterial cells, e.g., E. coli. Our study showed that the combined effect of confinement and side-loops can explain the emergent helical shape of bacterial nucleoid. The side-loops are known to produce an effective bending stiffness in the local scale, and thickening of the polymer that potentially impacts the relative

packing of the backbone [18]. Both the effects are expected to depend on the side loop size. For a given cylindrical confinement, we varied the size of side loops and analyzed the emergent shape of the confined feather-boa polymer. Our molecular dynamics simulations revealed, at small side loop sizes the feather- boa chain remains randomly organized. Beyond a certain threshold side loop size, the helical organization of backbone emerges spontaneously. This helical organization persists with further increase in side loop size. The smaller side loop sizes are unable to provide the necessary amount of bending rigidity and thickening of the backbone required to stabilize the helical morphology. Not only the backbone of the feather-boa chain, the local center of mass of monomers also shows the helical organization. The helicity is further characterized in terms of tangent correlation functions, and corresponding structure factors. The pitch of the helix, and the degree of helicity defined in terms of the height of the structure factor at its peak, increase to saturate with increasing side-loop size. Associated with the emergence of helicity, the monomer density shows periodic modulation along the long axis of the cell. The amplitude of oscillation increases with the monomer density brought about by increasing side-loop size. Here, we have shown the emergence of helicity in a linear feather-boa chain, although the real bacterial chromosomes are made up of circular DNA. Similar helical shape is observed even for a circular backbone chain of a feather-boa polymer. We return to this point in the following chapters.

This chapter has shown that the generic loop structures of the chromosome along with entropic repulsion, which are purely physical effects, can explain the emergent helical morphology of chromosomes confined to cylindrical cells, e.g., in E coli and B. subtilis. In the next two chapters, we study the impact of cytosolic crowding on chromosomal organization.

5 Impact of confinement and molecular crowders on chromosome organization

In the previous chapter, we introduced the feather- boa model of bacterial chromosome in a cylindrical confinement of rod-shaped cells like E. coli. We showed, beyond a certain threshold size of the side loops, the feather-boa chain adopts a helical organization spontaneously. In the current chapter, we study the impact of cytosloic crowding on such chromosomal morphology, and the relative organization of chromosome and crowders. For the sake of completeness, we start by describing the system under study.

In E. coli bacteria, a 4.6 Mbp (1.6 mm) long negatively supercoiled circular DNA and associated binding proteins constitute the chromosome. The chromosomal structure is suspended in a crowded cytosolic fluid and forms a membrane-less organelle called the nucleoid occupying a central sub-volume of the cell [194]. The chromosome along with the cytosolic fluid and other components of the cell are confined by a cell envelope which has a cylinder geometry. In a typical wild type *E.coli* cell, the cell envelope produces cylindrical confinement of diameter 0.8 μm, and length 2-4 μm. The 1.6 mm long DNA has to fold about 10^3 times in order to fit inside the confinement [12, 124, 126]. This huge compaction has to be concomitant with the functional organization of chromosome because the chromosome has to express genes, regulate the expression, and allow replication in order to survive [127].

There are at least four major physical and chemical mechanisms through which such compaction and organization are usually achieved [195, 196]. First, the confinement itself. However, it is not all. A fact that can be appreciated by noting that the nucleoid only occupies a sub-volume of the cell, with a ratio of the nucleoid and cellular volume being

$< \frac{1}{4}$ [126]. Second, the DNA is almost 5% negatively supercoiled, and this under-twisting results in warps and folds of various segments of the chain leading to the folded hair-braid like structures known as plectonemes [15]. Formation of plectonemes substantially decreases the effective size of the chromosome. Third, simultaneous binding of nucleoid associated proteins (NAPs) at more than one location along the genome [125–127, 129, 197]. Impact of such cross-linking has been discussed in the previous chapters. For the sake of completeness, we reiterate a few major points here. The NAPs are mainly of two categories: non- specefic like HU which binds to DNA without any bias, and specific like FIS which binds to only a few target genetic sequences. The resultant loop formation leads to further compactification of the chromosome. There are dedicated proteins which converts these looped structures into topological associating domains [84, 127, 177, 198]. Further, structural modification of chromosome (SMC) proteins can lead to active loop extrusion. In human chromosomes cohesin and ctcf generate and maintain such loops [199]. In bacteria, the condensin SMC extrudes loops against barriers of RNA- polymerase or transcription machinery [86]. Electron microscopy experiments on lysed *E.coli* chromosomes showed a bell-shaped distribution of loops with a maximum close to 10-12 kbp. With the advent of chromosome conformation capture (CCC) based techniques, it has been possible to directly probe the physiological interaction frequencies between different gene segments in the chromosome [144, 145]. These techniques allowed to form contact maps that exhibit probability of spatial contacts between various gene- locations along the genome and one can see the presence of TADs in form of characteristic checker- board patterns on the 2-D map. Fourth, the depletion effect due to molecular crowding generated by cytosolic components can potentially compact the chromosome [200–202]. Within an equilibrium picture, the overall entropy of a system consisting of a polymer and molecular crowders increases by allowing more volume to the large degrees of freedom of the crowders, effectively compressing the polymer [203, 204]. Such entropic forces are expected to play an important role in chromosome compaction, and segregation of the chromosome from protein crowders [11, 205–209]. In equilibrium, however, depletion

is known to drive the large molecules to a corner of the confining boundary. In contrast, bacterial nucleoid floats at the center of the cell with molecular crowders locating mostly near the cell envelop. Illustrating a reason for that is one of the goals of this chapter.

Recent experiments over the last decade showed an emergent helical organization of chromosome characterized by a few turns is ubiquitous in rod-shaped bacteria [9, 12, 13]. They exhibit a definite pitch to length ratio. The helical shape of the nucleoid is closely related to the cylindrical cell- shape. Modification of the cell shape by impeding the cell wall formation led to a change in the nucleoid morphology [12]. Using an argument of optimal packing, it was shown that thick chains in a cylinder pack best as a helical object [210]. The feather-boa model, a coarse-grained model of the chromosome as a bottle- brush like polymer, with polymeric loops attached to a backbone chain, has been proposed recently to describe observed chromosomal morphologies in bacteria [18, 19]. The entropic repulsion between side- loops provides an effective bending rigidity over micron length scales, much longer than the DNA persistence length. This, along with the monomer packing, leads to the emergent helical shape of the chromosome confined to a cylindrical volume of the bacteria [18]. The impact of confinement on the chromosome size and shape has been analyzed recently in growing bacteria that do not undergo cell division [9, 10].

Here, we assume the combined influence of confinement, and molecular crowding on chromosomal organization, utilizing the feather-boa model of the chromosome where the side-loops are assumed to be stabilized by proteins [18]. We focus on the question as to how the cytosolic crowding impacts the chromosomal organization and relative positioning of crowders with respect to the chromosome inside a cylindrical cell. Note that the poly-dispersed and the highly crowded cytosolic environment remains fluidized by the strong metabolic activity [40]. This allows us to model the cytosolic environment as a Langevin heat bath maintaining an effective temperature.

In the current chapter, we assume non-additive crowders, that repel against the chromosome chain and confinement boundaries, but otherwise are considered non-interacting. We

show, the center of mass of the chain along the cell length is organized in a helical fashion. Intersetingly, the local center of mass of crowders along the cell- length adopt a helical trajectory occupying positions complementary to the helical location of the chromosomal helix. Associated with this, the crowders and monomers show density modulation along the cell- length, with out-of-phase periodicities. This observation is reminiscent of a similar complementary density modulation of ribosomes with respect to the chromosome in *E.coli*, observed in recent experiments [11].

To probe deeper into the physical effects of molecular crowding on the formation of the helicoid morphology of chromosomes, and their size, we adopt a simulation-wise more tractable model in which we further coarse-grain the feather-boa chain. The effect of the side loops is mimicked by an additional Gaussian- core repulsion between the monomers of the backbone. Under the cylindrical confinement, such polymers are also known to show emergent helical morphology [10, 18, 19]. In Ref. [10], using a single crowder- size, longitudinal segregation of crowders and chromosome has been shown. The crowders were expelled towards the two caps of the cylinder, generating a compression that maintained a helicoid shape of the nucleoid, even in the longest cells.

In the current chapter, we fix the length of the cylinder, the density of crowders, and probe the impact of change in crowder size. We probe its impact on the relative organization of crowders and the model chromosome. We show that, at small crowder size, the chromosome spans the whole cell, and crowders remain uniformly distributed across its volume. As we increase the crowder size, the chromosome starts to adopt a helical morphology, and crowders organize into a complementary helix with respect to the chain. This happens via transverse segregation between the chromosome and crowders. At a larger crowder size, due to strong entropic repulsion, the chain expels the crowders longitudinally, to both the ends of the cylinder in a symmetric fashion. This leads to full longitudinal segregation at large enough crowder size. The full feather- boa model, and the effective coarse-grained model shares this property of reciprocal helical organization of the local center of mass of the crowders with

respect to the chromosome, along the cell length.

In the following, we present the details of our simulation study in two sections. In the first section, the all-atom molecular dynamics simulation of the feather boa polymer in the presence of crowders is presented. In the following section, we present the coarse-grained model of the chromosome. We summarize the results in the respective sections.

5.1 Full feather boa model

The chromosome at its full length is modeled as feather-boa polymer. We define a feather-boa chain as a flexible linear polymer that consists of n_b monomers and every monomer in the chain has a ring polymer attached to it which consists of n_s monomers. The linear chain of n_b monomers constitute the backbone of the feather boa chain. The ring polymers attached to the backbone chain represent the presence of loops in the chromosome which is observed in experiments [26]. The total number of monomers in the chain equals $n_b(1 + n_s)$. The consecutive monomers of the feather-boa chain (backbone as well as side loops) are bound by harmonic interaction,

$$V_b = \frac{A}{2}(\mathbf{d_i} - \sigma \mathbf{u_i})^2, \tag{5.1}$$

where $\mathbf{d_i} = \mathbf{r_{i+1}} - \mathbf{r_i}$, $\mathbf{r_i}$ is position of i-th bead, and $\mathbf{u_i} = \frac{\mathbf{d_i}}{|\mathbf{d_i}|}$ is the local tangent vector to the chain. One of the monomer of a ring polymer is attached to a single backbone monomer with same harmonic potential. All the monomers of the feather boa chain belonging to backbone as well as side loops, repel each other with short range repulsion given by WCA interaction,

$$\beta V(r_{ij}) = 4\epsilon[(\sigma/r_{ij})^{12} - (\sigma/r_{ij})^6 + 0.25], \tag{5.2}$$

for $r_{ij} < 2^{\frac{1}{6}}\sigma$ and $\beta V(r_{ij}) = 0$ otherwise. r_{ij} is the separation between i-th and j-th monomer and σ is the monomer diameter [157]. The length and the energy scale are set by σ, ϵ, and together they set the time scale $\tau = \sigma\sqrt{\frac{m}{\epsilon}}$. We chose parameters $A = 100\epsilon/\sigma^2$. The number of monomers in the backbone $n_b = 200$ and number of monomers in each side loop $n_s = 40$, which leads to total number of monomers in the feather-boa chain equal to 8200. The cylindrical confinement is realized with parameters, $D = 29.5\sigma$ and $L = 50.74\sigma$, where D and L is the diameter and length of the cylinder receptively. The cylinder wall along with top and bottom surface (the caps at two ends of cylinder) repel all the monomers with short range interaction,

$$\beta V_{wall}(r_{iw}) = 2\pi\epsilon\left[\left(\frac{2}{5}\right)\left(\frac{\sigma}{r_{iw}}\right)^{10} - \left(\frac{\sigma}{r_{iw}}\right)^{4} + \frac{3}{5}\right], \tag{5.3}$$

when $r_{iw} < \sigma$, and $\beta V_{wall}(r_{iw}) = 0$ otherwise. Here, r_{iw} is the separation of the monomer from the wall. This set of parameters yield the monomer packing fraction equal to approximately 0.24. Note that the details of the feather- boa chain and confinement are consistent with the previous chapter. With the above parameters, molecular dynamics simulations are performed using the velocity-Verlet scheme in the presence of a Langevin thermostat characterized by isotropic friction coefficient $\gamma = 1/\tau$ fixing the temperature $T = 1.0\epsilon/k_B$. This entire scheme is implemented by ESPResSo molecular dynamics package [161]. We conduct the numerical integration using step-size $\delta t = 0.005\tau$.

To perform the simulation, first, a feather boa chain in a very large cylinder is equilibrated in the heat bath. Then, very slowly the cylinder diameter and length are reduced and equilibration is performed at each intermediate cylindrical confinement until target confinement is achieved. After attainment of target confinement, the simulation is performed for further 10^7-5×10^7 steps to equilibrate the system. For the analysis of equilibrium properties, the data is collected over the next 10^4 configurations with the gap of $5 \times 10^3\tau$ in time.

The interplay between local effective stiffness in the backbone due to repulsion among

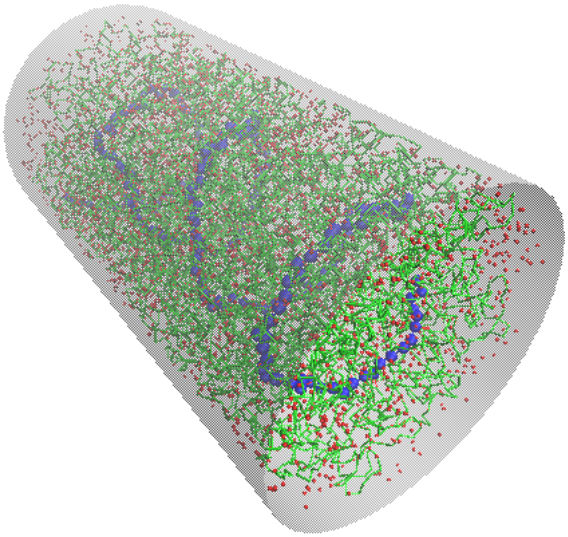


FIGURE 5.1: The equilibrium configuration of the feather-boa chain with backbone length $l_b = 200\sigma$ and side loop size $l_s = 40\sigma$. The backbone is shown in thick blue color and side loops are shown in thin green line. The crowders are shown by red dots.

neighboring side loops and intrachain packing of monomers in the cylinder gave rise to the helical organization of the backbone under cylindrical confinement [18]. In the previous chapter, we have studied the impact of changing the side loop size of a feather- boa chain on the local morphology. Here we fix the side loop size at $n_s = 40$, and introduce the so-called non-additive crowders inside the cylinder along with feather boa chain to capture the generic impact of cytosolic crowding on chromosome organization.

5.1.1 Feather boa chain in a cylinder: a constant density of crowders

We introduce $N_c = 3000$ non-additive crowders in the system. The crowders do not interact between themselves, however, repel the monomers of feather-boa chain with WCA potential $V(r_{ij})$ and confinement walls with potential $V_{wall}(r_{iw})$. With the above-mentioned parameters, we perform the molecular dynamics simulation.

In Fig. 5.1, a snapshot of equilibrium configuration is shown. The monomers in the backbone of the feather-boa chain are represented by blue beads and the side loops are

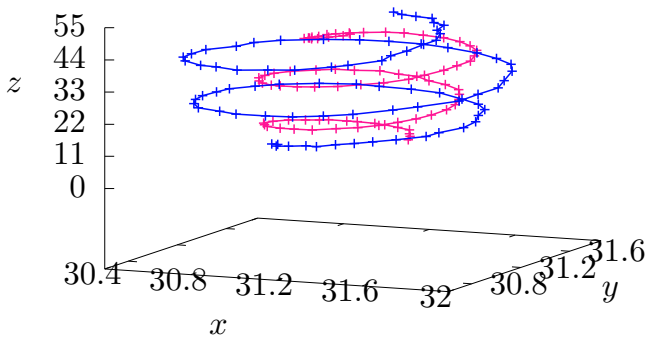


FIGURE 5.2: Complimentary helical organization of center-of-mass loci for monomers and crowders are shown in blue and pink respectively. z- denote the long axis of cylinder and xy represent a radial plane.

shown by green beads. The crowders are denoted by red dots. The snapshot clearly shows the helical organization of the backbone in the presence of crowders. As discussed in the previous section, this kind of helical organization of the backbone is observed in the absence of crowders [18], perhaps the introduction of crowders with small density does not impact the helical organization. Further, it seems the crowders are distributed homogeneously throughout the cylinder, however in the following we probe any possible local organization of them relative to model chromosome.

In Fig. 5.2, the mean center of mass orientation for model chromosome and the crowders are shown in blue and pink lines respectively. Averaging for the respective center of mass loci is performed over 10^4 equilibrium configurations separated from each other by $5 \times 10^3 \tau$. Similar to previous chapter, the bin size used for the local center of mass computation is 0.5σ. The two loci show the complementary helical orientation of center-of-mass for monomers and the crowders along the long axis (z) of the cylinder. To have a better view of the respective center-of-mass for chromosome and the crowders, we show the scatter plot

in Fig. 5.3(a). Along the long axis of the cylinder, around various xy-planes recognized by their z values, we consider bins of size 0.5σ. In each bin, we calculate the center-of-mass of all the monomers, and all the crowders. Afterward, we collect the respective center of mass co-ordinates for monomers and the crowders over 10^4 configurations which are represented by blue dots for monomers and red dots for crowders. In Fig. 5.3(a), scatter plots are shown at $z = 25\sigma, 28\sigma, 32\sigma$ along the long axis of cylinder. The arrows represent the separation vector between monomer and crowder center-of-masses in that particular xy- plane. The separation vector rotates around each other along the z-axis of the cylinder. It suggests that the crowders organize themselves into the helical domain complementary to the feather-boa backbone.

To quantify the complimentary helicity, we compute the tangent-tangent correlation which is shown in Fig. 5.3(b). In Fig. 5.2, we have already shown the time average center of mass loci for monomers and the crowders. We project the respective center of mass loci for the monomers and the crowders on a $xy-$plane. Now, we calculate the auto-correlation $\langle\hat{t}(z)\cdot\hat{t}(0)\rangle$ for corresponding unit vectors along that plane projected loci. The autocorrelation function for monomer center of mass is shown in blue $\circ$. The periodic oscillation confirms the helical organization of center-of-mass of monomers with a helical pitch $\approx 14.5\sigma$. Next we compute the cross-correlation function $\langle\hat{t}_c(z)\cdot\hat{t}(0)\rangle$ between plane projected center-of-mass loci of monomers and the crowders. The cross-correlation function represented by green $\square$ shows out-of-phase periodic oscillation relative to $\langle\hat{t}(z)\cdot\hat{t}(0)\rangle$ with the same helical pitch. This establishes the fact that the center of mass orientations for monomers and the crowders are organized in complementary helical fashion with respect to each other.

In Fig. 5.3(c), we have plotted the probability density profiles for the monomers and crowders along the long axis of the cylinder in blue $\circ$ and red $\square$ respectively. Details of the computation are defined in the previous chapter. They show out-of-phase density oscillation along the long axis of the cylinder. Near the two ends of the cylinder, their density profiles show opposite behavior, *i.e*, the crowders have the largest density while the monomer density

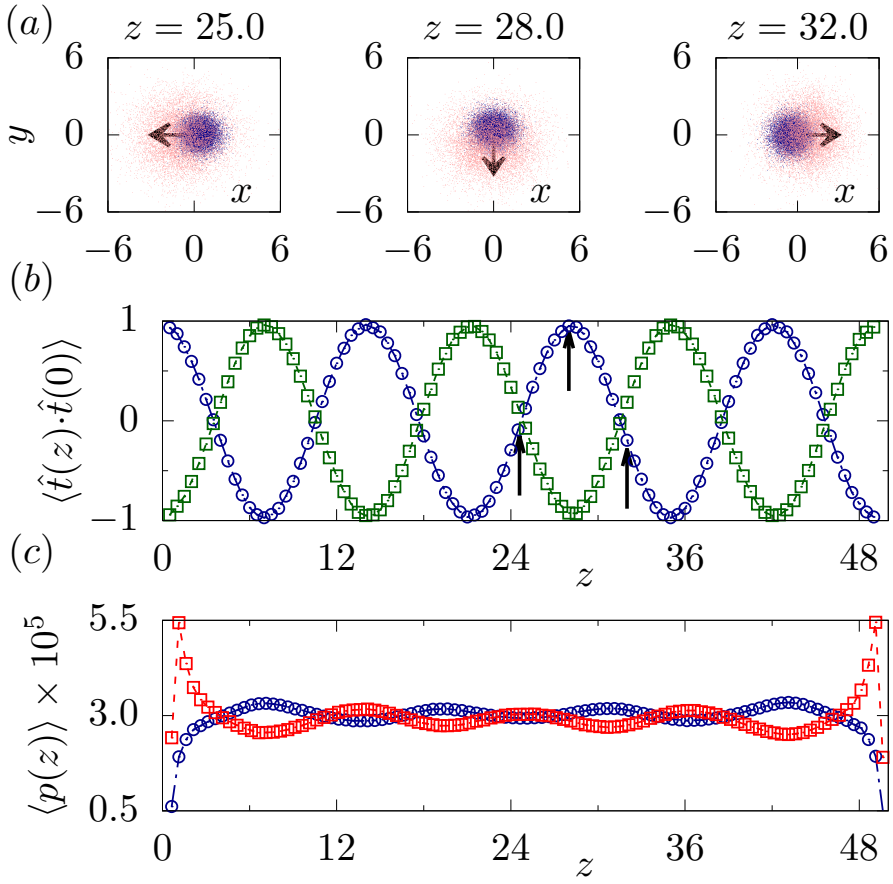


FIGURE 5.3: (a) Scatter plots of center-of-mass of monomers and crowders are shown in blue and red dots respectively. The respective center-of-masses are computed at different $xy-$planes along the long axis (z) of the cylinder, within a bin of size 0.5σ over 10^4 equilibrium configurations. The arrow denotes relative orientation of crowder- centre of mass with respect to that of monomers. It rotates along cell length. (b) The auto-correlation $\langle \hat{t}(z) \cdot \hat{t}(0) \rangle$ of the xy plane- projected center of mass orientation of monomers (blue $\circ$), and its cross-correlation with that of crowder center of mass $\langle \hat{t}(z) \cdot \hat{t}(0) \rangle$ (green $\square$) vary periodically with the length of cylinder. The periodicity gives the helical pitch $\approx 14.5\sigma$. The arrows on the $\langle \hat{t}(z) \cdot \hat{t}(0) \rangle$ graph denote the longitudinal positions at which the scatter plots are shown in (a). (c) Probability density of the monomers ($\circ$) and crowders ($\square$) along the long axis of the cylinder.

depletes to almost zero. This implies the crowders like to wet the wall similar to confined fluid and monomers of the chain like to stay away from the wall. A similar kind of out-of-phase modulation for chromosome and ribosome density has been observed in a recent experiment in rod-shaped bacteria like *E. coli* [11].

Together, these results show that the emergent helicity of chromosomal morphology not only is robust with respect to the introduction of crowders but moreover imprints itself on the spatial distribution of the crowders. In the next section, we consider a more computationally tractable coarse-grained model for feather-boa chain asking the similar morphological questions.

5.2 The coarse-grained model

After quantifying the organizational and morphological details of model feather-boa chromosome in presence of molecular crowders, we further explore the impact of the degree of cytosolic crowding on chromosome morphology. To make this study computationally more tractable, we coarse-grain the feather-boa chain to an effective chain replacing the impact of side-loops by Gaussian core repulsion between the monomers of the backbone [9, 10, 18]. Strength and interaction range of Gaussian interaction depends on the radius of gyration of side-loops. This additional repulsion between the monomers of backbone provides effective soft thickening of monomers that allows the chain to behave like a soft tube yielding effective bending stiffness along the backbone. Form of Gaussian-core interaction is given by,

$$V_{gc}(r_{ij}) = a \exp[-r_{ij}^2/2\Sigma^2], \tag{5.4}$$

with $\Sigma^2 = 2R_g^2$. Here, r_{ij} is the separation between $i\text{-}th$ and the $j\text{-}th$ monomer, and R_g is the radius of gyration of each side loop. $R_g = cn_s^{3/5}\sigma$ with c = 0.323, a number consistent with earlier simulations is assumed [18, 215]. Note that, cutoff value beyond which Gaussian core repulsion vanishes is considered as four times the width of potential. This is valid for all the

Gaussian core potentials mentioned in this entire thesis. In this particular case, width is Σ.

We model *E.coli* chromosome with a circular feather-boa chain having n_b monomers in the backbone, and to each monomer, a side loop consisting of n_s monomers is attached. The backbone along with side loops is assumed to be realized by about 4×10^4 monomers where each monomer has diameter $\sigma = 0.04$ mm or 115 bp. We have considered each side loop consisting of $n_s = 62$ beads, representing 7.2 kbp of DNA that converts to 2.4 mm. The backbone consists of $n_b = 636$ beads which amounts to length 24.8 mm. Hence, the total chain length enumerates to $l = (n_b + n_b \cdot n_s)\sigma = 1.6mm$, as for bacterial chromosome. Here, we realize the presence of side loops with Gaussian core potential with following details.

The coarse-grained chain is considered inside a cylinder of the length of approximately 12.2 μm and diameter 1 μm. Such a long cell can be grown inside a confining channel under specific biochemical and genetic control [9]. This large cell provides augmented possibilities for the organization of model chromosomes and crowders. In the units of simulation, the length of such cylinder is $L = 320\sigma$ and diameter $D = 26.67\sigma$. The side loops has radius of gyration $R_g = cn_s^{3/5}\sigma = 0.14D$ which is very small compared to confining cylinder diameter. Hence, the effective interaction between the side-loops can be modeled as that in bulk. The strength of Gaussian repulsion between the interacting polymers depends on their topology. For example, the two very long linear chains in bulk has repulsive strength 2 k_BT [215]. On the other hand, two circular chains repel themselves with strength in the range 2-6 k_BT. Inside the cell, the chromosome has loops as well as plectoneme like structures [15]. Hence, to incorporate the possibilities of both, we have considered an intermediate value of $a = 3k_BT$. The monomers of the coarse-grained chain repel the confining walls of the cylinder with short-range potential V_{wall} as described for the full feather-boa chain in the previous section. In addition to V_{wall}, Gaussian core contribution $V_{gc}(r_{iw})$ mimics the effective repulsion of side loops with cylinder walls. Form of $V_{gc}(r_{iw})$ is given as,

$$V_{gc}(r_{iw}) = \frac{a}{2} \exp[-r_{iw}^2/2\Sigma^2]. \tag{5.5}$$

Here width $\Sigma = R_g$, and strength $a/2$ is considered. r_{iw} represent the separation of a monomer with the walls of the cylinder. Further, the connectivity between the neighboring monomers are maintained by a finitely extensible non-linear elastic (FENE) potential

$$V_F(r_{i+1,i}) = -(k/2)\ln[1-(r_{i+1,i}/R)^2], \tag{5.6}$$

where $k = 30$ and $R = 1.5$ [216]. Self-avoidance between monomers is captured by WCA potential $V(r_{ij})$ mentioned in the previous section. To summarize the model, the neighboring monomers are bound by FENE potential and the monomers (bonded as well as non-bonded) in the interaction range repel themselves with WCA and soft core Gaussian repulsion $(V(r_{ij}) + V_{gc}(r_{ij}))$. Further, the monomers of this coarse-grained chain repel the confining wall with combined short-range potential $V_{wall}(r_{iw})$ and soft core Gaussian repulsion $V_{gc}(r_{iw})$.

To investigate the impact of cytosolic crowding on chromosome morphology, we introduce the non-additive crowders inside the confinement along with the model chromosome. The crowders are non-interacting between themselves and interact repulsively with the monomers of the chain and the confining walls of the cylinder. The repulsive interaction between crowders and the monomers of the chain has two contributions. First, short-range WCA interaction given by $V(r_{ij})$. Second, soft core Gaussian repulsion with following form,

$$V^c_{gc}(r_{ij}) = a\exp[-r^2_{ij}/2\Sigma^2_c]. \tag{5.7}$$

Here, r_{ij} is the separation between i-th crowder and j-th monomer. a and Σ_c denote strength and width of interaction. Σ_c can be considered as a measure of crowder size. Potential $V^c_{gc}(r_{ij})$ incorporates the effective interaction between side loops and the polymeric crowders at the coarse-grained scale. We fix, $a = 3k_BT$ same as above. The repulsive interaction of crowders with the wall also has two components. First is, short-range repulsion $V_{wall}(r_{iw})$ as mentioned in the previous section. Second, soft core Gaussian interaction $V_{gc}(r_{iw})$, capturing

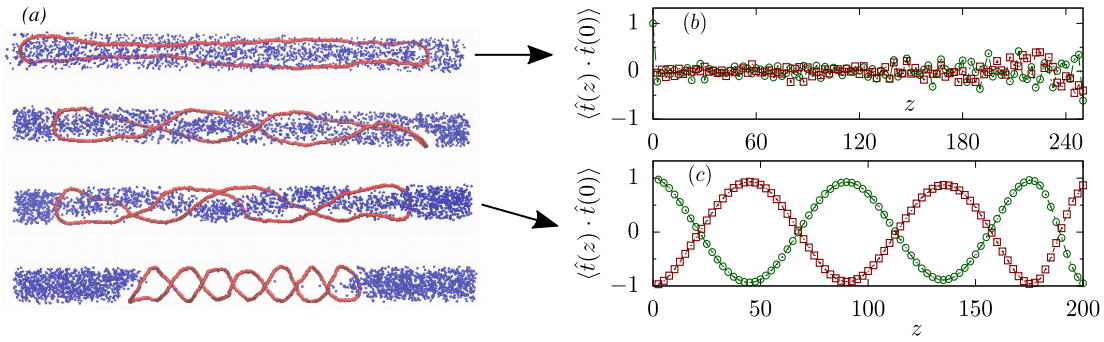


Figure 5.4: (a) Snapshots of the simulations are shown at various Σ_c. Top to bottom Σ_c denotes 2.0σ, 3.0σ, 3.5σ, and 4.0σ respectively. Cylinder length $L_z = 12D$ is considered. (b)-(c) The morphologies are analyzed using the projected tangent-tangent correlation $\langle \hat{t}(z) \cdot \hat{t}(0) \rangle$ in xy plane of the monomers (green $\circ$), and cross-correlation $\langle \hat{t}_c(z) \cdot \hat{t}(0) \rangle$ between monomers and crowders (brown $\square$) are shown along the length of the cylinder. (b) Orientations are uncorrelated at $\Sigma_c = 2.0\sigma$. (c) Complementary helicity is observed in terms of the out of phase oscillations of the two correlations at $\Sigma_c = 3.5\sigma$.

the effect of polymeric crowders. To investigate the impact of this effectively large crowder size, we consider a smaller number of crowders $N_c = 2000$. Molecular dynamics simulation is performed using velocity-Verlet algorithm with step size $\delta t = 0.01\,\tau$. The temperature of the heat bath is maintained by Langevin thermostat which fixes the temperature $T = 1.0\epsilon/k_B$ [161]. As a proxy for varying the degree of cytosolic crowding, we chose the size of crowder Σ_c. This governs to 0-th order the strength of induced depletion interaction [217]. In the following, we thoroughly investigate the impact of crowder size Σ_c on local morphology and relative organization of model chromosome and the crowders.

5.2.1 Impact of crowder size on organization

In Fig.5.4(a) snapshots of the simulation are shown at various crowders size Σ_c. Top to bottom, as we increase Σ_c, an apparent change in the relative organization of the model chromosome and the crowders is observed. When Σ_c is small, the crowders freely move and are distributed inside the cylindrical confinement in a homogeneous fashion. Due to the large cylinder length $L = 12D$, the chromosome opens up completely. As we increase the Σ_c,

the crowder and the monomers of the chain start to fell the repulsion, thus begin to avoid each other. Around $\Sigma_c = 3\sigma$, due to repulsion, the spatial separation between the crowders and the chromosome becomes apparent. The crowders start to push the model chromosome towards the helicoid shape. At $\Sigma_c = 3.5\sigma$, a clear complementary helical organization of chromosome and the crowders is observed as shown in Fig.5.4(a). Further increase of Σ_c leads to more compression of the chromosome due to crowders, thus increasing the local monomeric density. This augmented local density of monomers expels the crowders towards cap regions of the cylinder leading to longitudinal separation of monomers and the crowders along the long axis of the cylinder. Consequently, the chromosome occupies the central region and gets compressed by crowders which localize around both the cap regions.

5.2.2 Complimentary helicity

Around intermediate values of Σ_c, the complementary helical organization of crowders and the chromosome is observed. The helicity for monomer center-of-mass orientation is quantified in terms of tangent-tangent correlation function $\langle \hat{t}(z) \cdot \hat{t}(0) \rangle$ along the plane projected average center of mass loci of monomers of the chain. First, the average center of mass loci of all the monomers is computed. The averaging for computation of center of mass is performed over 10^4 equilibrium configurations separated from each other by $10^3\tau$. Then the resulting mean loci are projected on the xy-plane assuming z being the long axis of the cylinder. Afterwards computation of $\langle \hat{t}(z) \cdot \hat{t}(0) \rangle$ is performed in the xy-plane. The computation scheme is similar to the previous section. Further, we calculate the mean center of mass loci for crowders similar to monomers. Then, cross-correlation $\langle \hat{t}_c(z) \cdot \hat{t}(0) \rangle$ of plane projected center of mass loci for the crowders and the monomers is computed. The two plots at different Σ_c are shown in Fig. 5.4(b)-(c). At small Σ_c, we observe the correlation between plane projected monomer center of mass loci (green $\circ$) and cross-correlation between plane projected center of mass loci of monomers and the crowders (brown $\square$) do not show any structures (see Fig. 5.4(b)). At intermediate value of crowder size $\Sigma_c = 3.5\sigma$, we see out

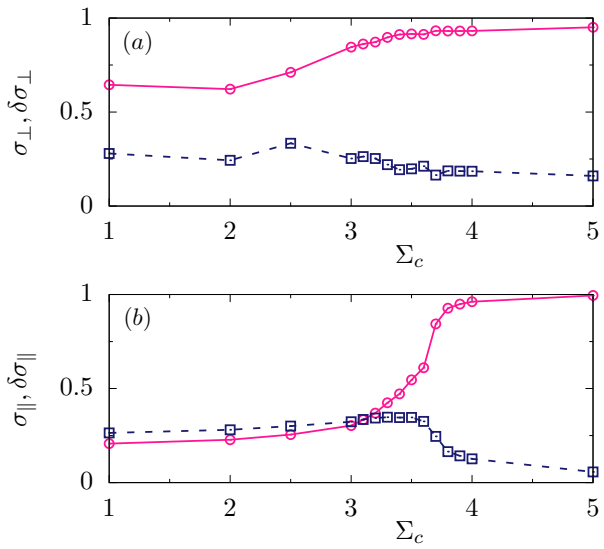


Figure 5.5: (a) Transverse spatial segregation $\sigma_\perp$ ($\circ$) and the corresponding root mean squared fluctuations $\delta\sigma_\perp$ ($\square$) as a function of crowder size Σ_c. (b) Change in the longitudinal separation $\sigma_\parallel$ ($\circ$) and the corresponding root mean squared fluctuations $\delta\sigma_\parallel$ ($\square$) with Σ_c.

of phase oscillation with the same periodicity between the two correlation functions which implies the fact that center of mass for monomers and the crowders are organized in complementary helical fashion (see Fig. 5.4(c)). The two completely different behavior in Fig. 5.4 (b) and (c) illustrates the impact of crowder size on the organization of the chromosome and the crowders. First of all, a minimal crowder size ($\Sigma_c = 3\sigma$) is required to generate enough repulsion between crowders and the chromosome to nucleate the formation of a complementary helical organization of two. Further increase of crowder size, $\Sigma_c \geq 4\sigma$, leads to full longitudinal separation of crowders and the chromosome.

5.2.3 Transverse and longitudinal segregation

As we have observed, at small Σ_c, crowders were homogeneously distributed and the chromosome remains structure-less. With the increase of Σ_c to intermediate values, the crowders start to push the chain leading to transverse (radial) segregation of crowders and the chromosome that results in the complementary helical organization of two. Further increase in

Σ_c repels the chain strongly, hence augmenting the local monomeric density of chromosome. Due to this, at large crowder size, we observe longitudinal segregation of crowders and the chromosome along the long axis of the cylinder. The transverse and the longitudinal spatial segregation are quantified in terms of the following order parameter.

$$\sigma_{\parallel,\perp} = \left\langle \left| \frac{\rho_m - \rho_c}{\rho_m + \rho_c} \right| \right\rangle .$$

The quantities $\sigma_{\parallel,\perp}$ are the mean values of the expression on the right side measured over respective bins. The cylinder is binned into transverse and longitudinal bins. During the transverse binning, the cylinder is binned into co-axial 10 cylindrical shells. In the longitudinal binning, similar to the previous section, we bin the cylinder into 120 consecutive disks along the long axis of the cylinder. $\rho_{m,c}$ represents the local density of the monomers and the crowders within each bin respectively. The local densities are normalized by the respective total number of monomers N_m and the crowders N_c.

To quantify the transverse spatial segregation of monomers and the crowders, we restrict our calculation to the region of the cylinder where the chromosome resides. In that particular spatial region, we compute $\sigma_\perp$ and average it over 10^4 well-separated equilibrium configurations. Similarly, the longitudinal separation between crowders and chromosome is quantified by computing $\sigma_\parallel$ and averaging it over same number of configurations. In Fig. 5.5, we have presented the change of $\sigma_{\parallel,\perp}$ with respect to crowder size Σ_c. In the disordered phase (low Σ_c), $\sigma_{\parallel,\perp}$ have small value, and with increase in Σ_c, as segregation sets in, the respective order parameters monotonically increase. The respective fluctuations (root means square deviations) are defined as, $\delta\sigma_{\parallel,\perp} = \sqrt{\langle\sigma^2_{\parallel,\perp}\rangle - \langle\sigma_{\parallel,\perp}\rangle^2}$. They show little increase around the cross-over points between the two different segregation regimes. The radial segregation sets in around $\Sigma_c = 2\sigma$, which occurs before the onset of longitudinal segregation around $\Sigma_c = 3\sigma$. At last, beyond $\Sigma_c = 4\sigma$, full macro-phase longitudinal segregation of chromosome and the crowders is achieved.

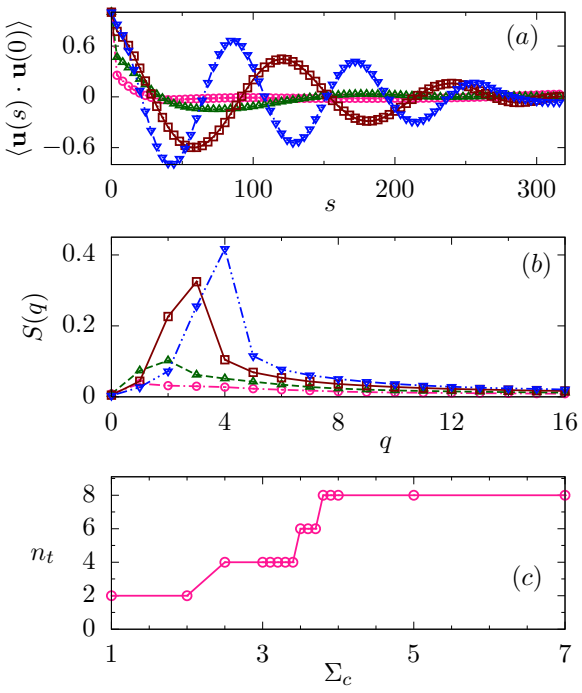


Figure 5.6: (a) The tangent-tangent correlation functions along the contour of the chain s evaluated at $\Sigma_c = 2\sigma(\circ)$,$3\sigma(\triangle)$, 3.5σ $(\square)$ and 4σ $(\triangledown)$ respectively. (b) The corresponding structure factors evaluated at the same values of Σ_c, showing peaks at $q = q_p$. (c) The number of turns $n_t = 2q_p$ increases with Σ_c.

5.2.4 Change in helicity

An increase in crowder size causes the helical morphology of the chromosome. The change in helicity is associated with two phenomena, *e.g*, spatial separation of crowders and the chromosome, and the resultant compression of the chromosome by crowders. Both these effects get augmented as we increase the Σ_c. We quantify the helical orientation of the chromosome in terms of tangent-tangent correlation function $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$ along the backbone of chain. The separation along the chain contour is represented by s. We chose half the chain length $n_b\sigma/2$, the longest separation along the long axis of cylinder for the computation of $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$. To enumerate the number of turns in the helical chromosome, plane projected tangent vectors along the backbone chain $\mathbf{u}(\mathbf{s})$ have been used. At small Σ_c, the chromosome

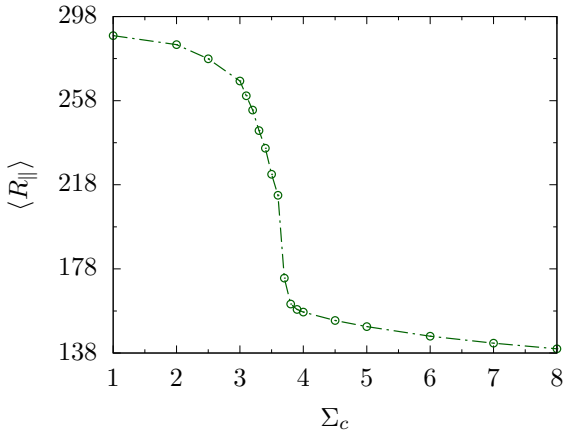


FIGURE 5.7: Extension of chromosome along the long axis of cylinder $R_\parallel$ is plotted against Σ_c.

remains open and the tangent-tangent correlation does not show any structures as shown in Fig. 5.6(a). With increase in Σ_c, the helicity in the chromosome sets in which is manifested in terms of periodic oscillations in the $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$. With further increase in Σ_c, the oscillation in terms of amplitude and the number grows. The periodicity of the oscillation determines the pitch of the chromosome helix. We compute the structure factor as, $S(q) = \frac{1}{2\Pi} \int_0^{l_h} \langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle e^{i\mathbf{s}\cdot\mathbf{q}} ds$. Here, $l_h = n_b\sigma/2$ denote half the chromosome length and q is reciprocal to length. $S(q)$ shows a clear peak at $q = q_p$ that determines the total number of turns in the helix by expression $n_t = 2q_p$ (Fig. 5.6(b)). Factor of 2 is due to fact that we have calculated the tangent-tangent correlation for half the chain length ($n_b\sigma/2$) of circular chromosome. In Fig. 5.6(c), we have shown the increase in n_t as Σ_c increases. In the complimentary helical organization, we have 6 turns in the chromosome helix which increase to 8 in the full-longitudinally segregated regime.

5.2.5 Chromosome extension along long axis

In Fig. 5.7, we have shown the variation of chromosome extension along the long axis of the cylinder while increasing the crowder size. As we have discussed earlier, increasing crowder size Σ_c repels the chain with increasing strength pushing it to the center of the cylinder from

both ends. Consequently, the extension of the model chromosome along the long axis of the cylinder $R_{\|}$, monotonically decreases with an increase in Σ_c. Initially, at small crowder size ($\Sigma_c \leq 3\sigma$) before the onset of formation of helicoid morphology of chromosome, the reduction in size $R_{\|}$ is small with the increase in Σ_c. In the regime where complimentary helix of crowders and chromosome form, $3\sigma < \Sigma_c < 4\sigma$, the reduction in chromosome size $R_{\|}$ occurs very sharply with increase in Σ_c. With further increase in Σ_c, decrease in $R_{\|}$ slows down.

5.3 Conclusion

In this chapter, we considered a coarse-grained model of the bacterial chromosome as a feather-boa polymer. Under the cylinder confinement typical to an E.coli cell, we introduced ideal gas like crowders to investigate the impact of cytosolic crowding on the organization of chromosome. Crowders are non-interacting between themselves, however, they repel the feather- boa chain with short-range interaction. Our molecular dynamics simulation revealed such crowders do not impact the helical morphology of feather- boa chain which is reported in the previous chapter. In fact, the center-of-mass orientation of crowders and chromosome adopts a complementary helical organization. Due to this, their corresponding local densities show out-of-phase periodic oscillation along the long axis of the cylinder. Such behavior is similar to chromosomal and ribosomal density modulation observed in rod-shaped bacteria, e.g., E.coli [11]. In addition, crowders and chromosomes show the opposite behavior at the cell ends. The crowders wet the two caps of the cylinder however monomer density vanishes there.

The cytosolic environment is populated with proteins (crowders) of various sizes. To systematically investigate the impact of crowder size on morphology and size of the chromosome, we considered a computationally more tractable model. We coarse-grained a circular feather- boa polymer by replacing the effects of side loops with Gaussian- core repulsion

between monomers of its backbone. In the presence of small crowders, such a chromosome opens up completely which is allowed by the long cylindrical cell. Such crowders remain homogeneously distributed in the cellular volume. In the range of intermediate crowder size, transverse (radial) segregation of chromosomes and crowders lead to the complementary helical organization of two. Further increase in crowder size causes longitudinal segregation along the long axis that leads to complete expulsion of crowders by helical chromosome towards opposite cap regions of the cylinder in a symmetric fashion. Consequently, crowders compress the chromosome to nucleoid like subvolume in the center of the cell. With the increase in crowder size, chromosome extension along the long axis of the cylinder decreases that lead to an increase in the number of helical turns in the chromosome morphology. The current study illustrates the interesting physics due to interplay between crowding, confinement and chromosome morphology.

In the next chapter, we will investigate the impact of crowder density on relative organization and size of model chromosome.

6 Impact of crowder density on the morphology of bacterial chromosome

6.1 Introduction

In the previous chapter, we studied the impact of various sizes of molecular crowders in the morphology of bacterial chromosome in rod-shaped bacteria. As we have seen, the smallest crowders can penetrate all the cellular space irrespective of whether it is occupied by chromosome or not. However, for larger crowders, with sizes larger than the *mesh-size* in the chromosomal region, the molecular crowders cannot penetrate such zones. Thus they undergo radial or longitudinal segregation with respect to the chromosome. As a result the crowders further compacts the chromosome, and show interesting relative organization inside the bacterial cell. In this chapter, we focus on the impact of the variation of crowder density on the chromosome organization.

As has been discussed before, the bacterial chromosome of E. coli comprises of a 4.6 Mbp (1.6 mm) long negatively supercolied circular DNA and associated proteins. It is suspended in a crowded cytosol to form the nucleoid, a membrane-less organelle, occupying a central sub-volume of the cell [194]. The cell envelop, which has a sphero- cylindrical geometry confines the chromosome along with the cytosolic fluid and other components of the cell. A typical wild type E. coli cell has a diameter of 0.8 μm and length of 2-4 μm. The long DNA chain has to undergo a large compaction of minimum 10^3 fold. This requires a functional organization as the DNA segments must be available for gene expression, regulation, and

DNA replication [12, 124, 126, 127]. In the previous chapter, we have discussed about the four major physical and chemical mechanisms through which such compaction and organization is usually achieved [195, 196]. The confinement due to cell envelope brings in a large degree of compaction. However, this is not sufficient, as can be appreciated by recognizing that the nucleoid only occupies $\frac{1}{4}$-th of the available cell volume [126]. Second, the negative supercoiling of DNA yields folded hair-braid like structures known as plectonemes [15], which greatly reduces the overall chromosome size [15]. Third, the impact of DNA- associated proteins. For example, cross- linking of DNA segments by nucleoid associated proteins (NAPs) brings in contour-wise separated genomic sequences in spatial proximit [125–127, 129, 197]. In addition, active loop extrusion by structural modification of chromosome (SMC) proteins contributes to formation of large chromosomal loops [86]. Such loop formation leads to further compaction of the chromosome. These loops are observed in electron microscopy experiments [26], and using chromosome confirmation capture techniques and it Hi-C variants [144, 145]. Fourth, and the reason which is mostly relevant to the focus of this chapter, is the depletion effect due to molecular crowders in the cytosol [200–202]. Earlier simulation and experimental studies consisting of molecular crowders and polymer showed, compression of the polymer allowing more space to crowders [203, 204]. Within an equilibrium picture, this increases the overall entropy of the system. In equilibrium, depletion forces drive the big molecules, e.g., the polymer towards the boundary of confinement. In contrast, bacterial nucleoid floats at the center of the cell with molecular crowders locating mostly near the cell envelop. In a living cell the molecular crowders are produced in a region around the chromosome, thereby generating a density gradient maintaining the central location of the nucleoid.

Experiments during the last decade showed a helical organization of the chromosome with a few turn to be ubiquitous in live bacteria with cylindrical geometry [13]. The feather-boa model of chromosome, with polymeric loops attached to a backbone chain, describes chromosomal shape and size observed in bacteria [18, 19]. As has been shown recently, the

presence of molecular crowders further compresses the chromosome to maintain its helicoid shape even in long filamentous cells [9, 10].

In the previous chapter, we have investigated the combined influence of confinement, and molecular crowding on chromosomal organization, utilizing both the full feather-boa model of chromosome, and an effective coarse-grained model. We studied in detail the impact of crowder size on the relative organization, keeping the number density of crowders unchanged. While the small crowders permeate the whole cellular space without impacting the chromosome appreciably, at intermediate crowder size, we observed a remarkable complementary helicity of crowder organization with respect to the helicoid shape of the chromosome. To be consistent with the previous chapter, here we use the same cylindrical confinement and molecular crowding to investigate the impact of crowder density, instead of the crowder size, on the relative organization of chromosome and crowders. We stick to the model of non-additive crowders that repel against the model chromosome and confinement boundaries, but otherwise are considered an ideal gas.

The main results of this chapter are as follows. Within the full feather-boa model, the chromosome is organized in a helical fashion along the long axis of the cylinder. For a range of crowder densities, the crowder center of mass occupies spatial positions complimentary to the local center of mass of chromosome, along the cell length. This is the complimentary helical organization of the chrosomome and crowders. Associated with this, the crowders and monomers show density modulations along the cell- length, with out- of- phase periodicities. This observation is reminiscent of a similar out-of-phase density modulation of ribosomes with respect to the chromosome in E. coli bacteria [11]. Further, to incorporate the polymeric nature of the relatively longer protein crowders, we use an additional Gaussian- core repulsion between the crowders and monomers and confinement. This additional repulsion enforces an excess degree of segregation between the crowders and chromosome, and an additional compressive force is generated on the chromosome causing further deformation.

To investigate the physical impact of the variation of crowder density on the size and shape

of the chromosome in quantitative detail, we use a further coarse- graining of the chromosome model. This is computationally less expensive in numerical simulations than the full feather-boa chain. In this, the side loops of the feather-boa model are replaced by an additional Gaussian- core repulsion between the monomers of backbone. As has been shown before, this effective model captures the emergent helical morphology of the chromosome when confined to a cylindrical cell [10, 18, 19]. In the previous chapter, changing the crowder size, we have shown an onset of spatial segregation between crowders and chromosome. In an intermediate range of crowder size, the transverse segregation of chromosome and crowders led to a complimentary helical organization of the two. Further increase in crowder size led to longitudinal segregation of the crowders and chromosome, with crowders localizing near the caps compressing the chromosome to the center of the cell.

In this chapter, we fix the length and radius of the cylinder, and the size of the crowders to a value at which the complimentary helicity is observed in the previous chapter, to probe the impact of the change in crowder density. We investigate the resultant change in the organization of crowders and the chromosome. At small density of crowders, in a long enough cell, the chromosome opens up and the crowders get uniformly distributed across cylindrical volume. As we increase the crowder density, the chromosome starts to adopt a more compressed helical morphology, and crowders start to organize into a complimentary helix with respect to the chromosome. This happens via a transverse segregation between the chromosome and crowders. Further increase in crowder density leads to a strong compression of the chromosome via a longitudinal segregation of the chromosome and crowders. In this regime, the chromosome expels crowders to both the ends of cylinder along its long axis in a symmetric fashion. The full feather-boa chain and the effective coarse grained model with intermediate crowder densities share the property of reciprocal helical organization of the local center of mass of the crowders with respect to the chromosome, along the cell length. We characterize the change in helicity in terms of tangent- autocorrelation, winding number, turning number and the number of kinks.

In the following we present the models that we consider and results derived from them. In Sec. 6.2 a full feather-boa chain is considered in the presence of non-additive crowders that repel against the chain and the confinement with a local Weeks- Chandler- Anderson (WCA) potential. Sec. 6.3 describes the impact of an additional Gaussian core repulsion associated with the crowders, to incorporate the polymeric nature of the protein crowders up to the simplest approximation. In Sec.6.4 we consider the coarse- grained chain to investigate the impact of changing crowder density in greater detail. Finally we conclude summarizing our main results in Sec. 6.5.

6.2 Feather boa model : change in crowder density

Here we consider the feather boa polymer under cylindrical confinement in the presence of cytosolic crowders. The chain properties and confinement details are the same as mentioned in the previous chapters. The feather-boa chain consists of a backbone of 200 monomers, attached to side- loops of 40 monomers. This makes backbone length $l_b = 200\sigma$ and side loop length $l_s = 40\sigma$. The connectivity between the neighboring monomers is maintained by harmonic potential,

$$V_b = \frac{A}{2}(\mathbf{d_i} - \sigma \mathbf{u_i})^2, \tag{6.1}$$

where $\mathbf{d_i} = \mathbf{r_{i+1}} - \mathbf{r_i}$, $\mathbf{r_i}$ is the position of i-th bead, and $\mathbf{u_i} = \frac{\mathbf{d_i}}{|\mathbf{d_i}|}$ is the local tangent vector to the chain. All the non-bonded monomers repel each- other with the short range WCA potential,

$$\beta V(r_{ij}) = 4\epsilon[(\sigma/r_{ij})^{12} - (\sigma/r_{ij})^6 + 0.25] \tag{6.2}$$

for $r_{ij} < 2^{\frac{1}{6}}\sigma$ and $\beta V(r_{ij}) = 0$ otherwise. r_{ij} is the separation between i-th and j-th monomer [157]. The length and the energy scales are set by σ, and ϵ respectively, and

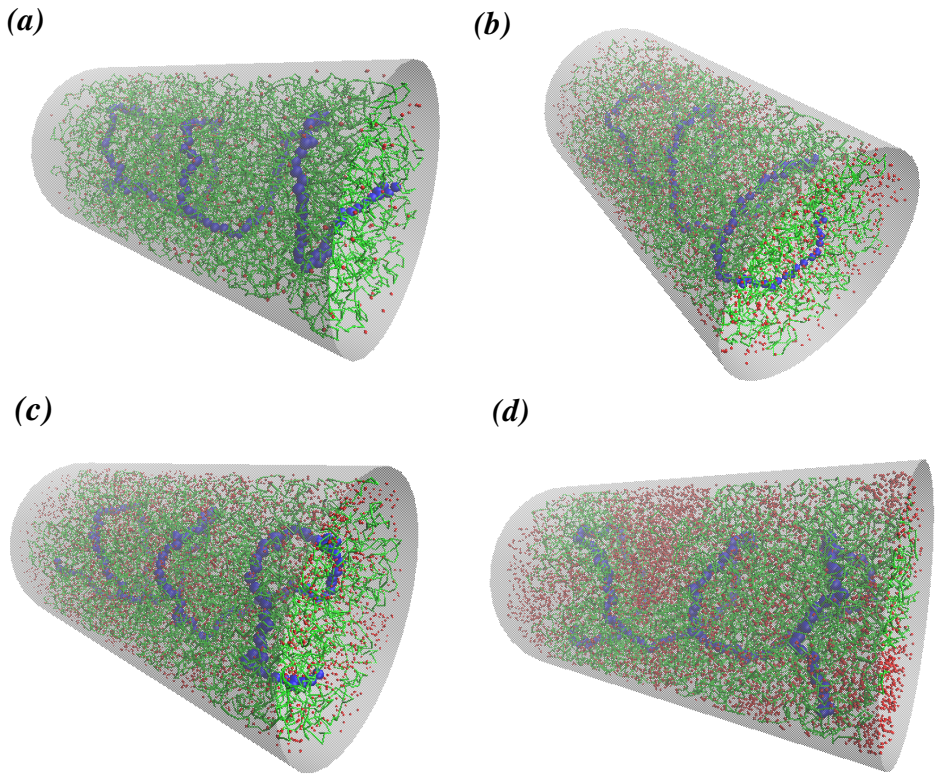


FIGURE 6.1: Snapshots of simulation are shown at various ρ_c. The backbone is shown in thick blue colour and side loops are shown in thin green line. The crowders are shown by red dots. (a), (b), (c), and (d) represents configurations at $\rho_c = 0.01, 0.09, 0.17$ and 0.24 respectively.

together they set the time scale $\tau = \sigma\sqrt{\frac{m}{\epsilon}}$. We choose $A = 100\epsilon/\sigma^2$ to keep the bond-length fluctuations within 1%.

The cylinder wall along with top and bottom surface (the caps at two ends of cylinder) repel all the monomers with short range interaction

$$\beta V_{wall} = 2\pi\epsilon \left[\frac{2}{5}\left(\frac{\sigma}{r_{iw}}\right)^{10} - \left(\frac{\sigma}{r_{iw}}\right)^{4} + \frac{3}{5}\right], \tag{6.3}$$

when $r_{iw} < \sigma$ and 0 otherwise. Here, r_{iw} is the separation of monomer from the wall. The cylindrical confinement is realized with parameters, $D = 29.5\sigma$ and $L = 50.74\sigma$, where D and L is the diameter and length of the cylinder respectively. These set of parameters yields

the monomer number density equal to approximately $\rho_m\sigma^3 = 0.24$. The form of all the potential used in this section is consistent with previous chapters, that led to emergence of spontaneous helicity in the cylinder- confined feather-boa model of chromosome.

To study the impact of molecular crowding due to cytosol, we consider the so-called non-additive crowders along with fethar-boa chain. The crowders do not interact between themselves but repel the monomers of the chain with short range WCA interaction $V(r_{ij})$, where r_{ij} is the separation between i-th monomer and j-th crowder. Like monomers of the chain, the wall-crowder interaction is short ranged repulsion given by $V_{wall}(r_{iw})$. We change the number density, $\rho_c = 4N_c/(\pi D^2 L)$ of the crowders by changing the total number of crowders N_c keeping the geometry unchanged.

We perform molecular dynamics simulations following velocity-Verlet algorithm in presence of Langevin heat bath. The Langevin thermostat is characterised by an isotropic friction constant $\gamma = 1/\tau$ fixing the temperature at $T = 1.0\epsilon/k_B$ as implemented by ESPReSo molecular dynamics package [161]. For numerical integration is performed choosing time steps of step-size $\delta t = 0.005\tau$. The simulations are performed for 10^7 to 5×10^7 steps to equilibrate the feather-boa chain in absence of crowders. After that, the crowders are slowly introduced and equilibration is performed over further 10^7 to 5×10^7 steps. For the analysis of equilibrium properties, data are collected over further 10^4 configurations which are separated from each other by $5 \times 10^3\tau$.

In Fig. 6.1, a few representative equilibrium configurations are plotted for a system with number of crowders $N_c = 500$, 3000, 6000 and 8200 which corresponds to number density $\rho_c\sigma^3 = 0.01, 0.09, 0.17$ and 0.24 in (a), (b), (c), and (d) respectively. The blue beads representing the backbone of the feather-boa chain shows spontaneous organization into helix at all the crowder densities considered here, as in Ref. [18]. The helical morphology of the backbone of feather-boa chain is quantified in the following section.

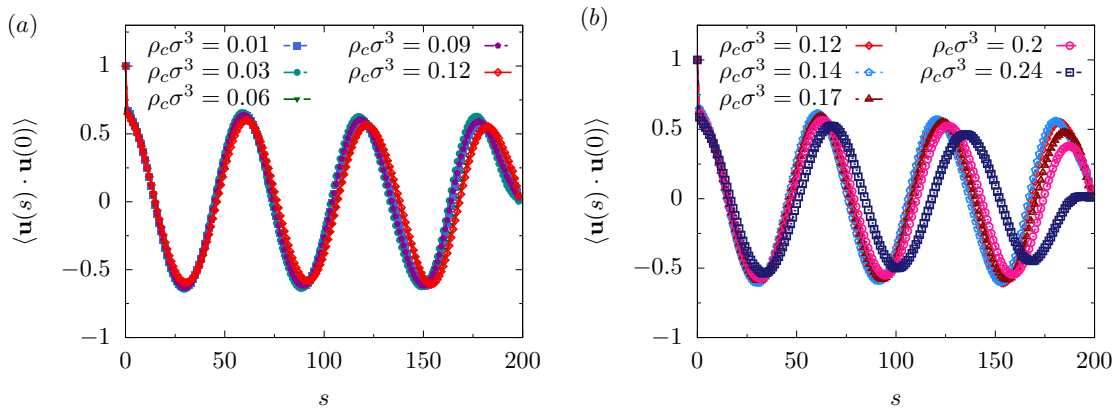


FIGURE 6.2: The tangent-tangent correlation, $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$ along the backbone of the chain is shown where s represents the separation along the backbone. (a) shows plot at smaller ρ_c compared to (b).

6.2.1 Backbone helicity

The degree of helical morphology of the chromosome is quantified in terms of tangent-tangent correlation function $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$ of unit bond vectors along the backbone contour of the feather-boa chain (Fig.6.2). The separation between contour segments along the backbone is denoted by $s = |i - j|\sigma$, with $i, j = 0, 1, 2...199$. As mentioned above, the averaging is performed over 10^4 equilibrium configurations separated from each other by $5 \times 10^3\tau$. To count the number of turns we use the radially projected tangent vector $\mathbf{u}(s)$. We compute $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$ at various number densities ρ_c of crowders. For better visibility, we show them in two plots. In Fig. 6.2(a), the correlation functions are shown for smaller ρ_c, and in Fig. 6.2(b) the same is shown for larger ρ_c. Throughout the span of $\rho_c\sigma^3$ from 0.01 to 0.24, we find periodic oscillations of correlation functions indicating helical morphology of the backbone. The relative deformation at the largest density is long- lived and is maintained by enhanced kinematic barrier.

In Fig. 6.3, structure factors $S(q)$ corresponding to tangent-tangent correlations $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$ are shown at different ρ_c. The structure factor is evaluated by performing Fourier transform of correlation function, $S(q) = \int_{-l_b/2}^{l_b/2} \langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle e^{isq} ds$. It shows peak at $q = q_p$

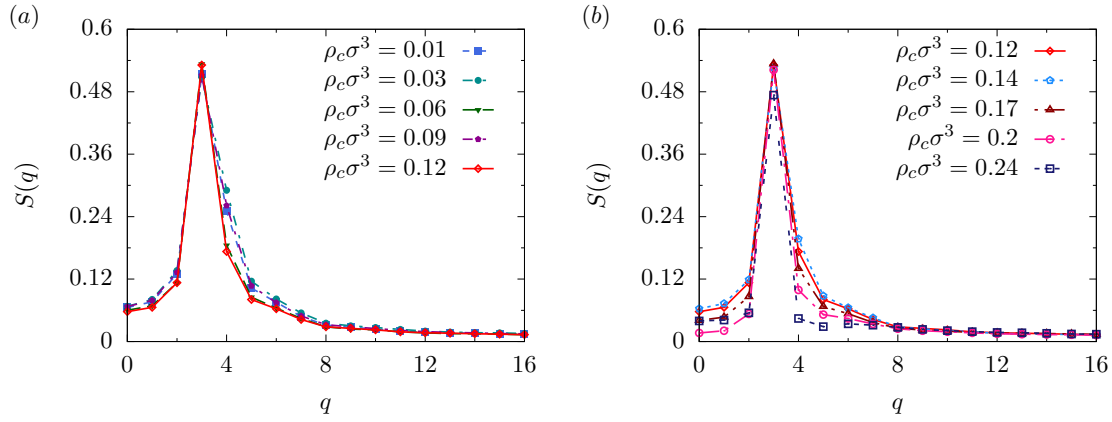


FIGURE 6.3: Structure factor $S(q)$ for $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$ is plotted against q. (a) shows the plot for smaller ρ_c values and (b) shows the same at larger ρ_c values.

corresponding to the turn number $n_t = q_p$ at all ρ_c values shown in Fig. 6.3(a) and (b). All of them shows $n_t = 3$ leading to a helical pitch $\lambda_{max} = l_b/n_t$. The three turns could be directly counted from the number of periodic oscillations in Fig.6.2. The amplitude of $S(q)$ at its maximum $q = q_p$ represents the relative measure of the degree of helicity, while the width of the peak is a measure of the statistical dispersion of the structure. The plots in Fig. 6.3(a) and (b), show almost same amplitude and width which implies that the backbone helicity does not change with changing density of crowders ρ_c. As we show later, in Sec. 6.3, this feature depends on the size of the crowders. Consideration of polymeric nature of the protein crowders, indeed, brings about significant modification of the chromosomal organization with crowder density.

6.2.2 Center of mass organization

In this section, we consider the center of mass organization for all the monomers and crowders along the length of the cylindrical confinement. For this purpose, we divide the cylinder into small disk-like bins of height 0.5σ along it length. Center of mass of the monomers and crowders are evaluated within these bins, and averaging is performed over 10^4 equilibrium configurations. The z-direction represents the long axis of the cylinder, while the xy-plane is

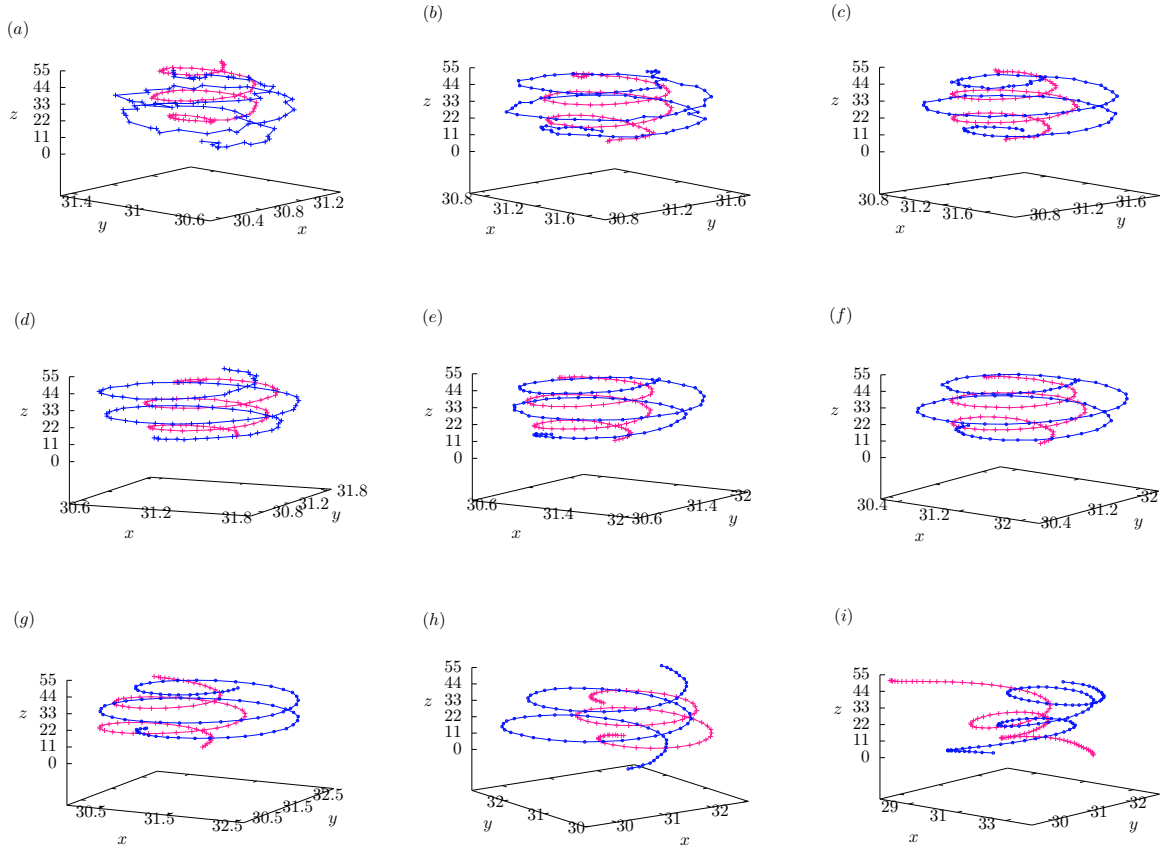


FIGURE 6.4: Center of mass for all the monomers and crowders are shown in blue and pink respectively. Different plots are at $\rho_c\sigma^3$ value, 0.01 (a), 0.03 (b), 0.06 (c), 0.07 (d), 0.12 (e), 0.14 (f), 0.17 (g), 0.2 (h) and 0.24 (i) respectively.

in the radial direction. In Fig. 6.4, the mean center of mass position of monomers is shown in blue and that for crowders is shown in pink. The plots in (a)-(i) are at crowder number densities $\rho_c\sigma^3 = 0.01, 0.03, 0.06, 0.07, 0.12, 0.14, 0.17, 0.2$ and 0.24, respectively. They organize spontaneously into complimentary helical fashion. For the $\rho_c\sigma^3 \leq 0.2$, the complimentary helical arrangement for monomers and the crowder center of mass appears very nice. This nice organization gets disrupted partially at the highest density, $\rho_c = 0.24$, because of large kinematic barrier.

Next, we quantify the complimentary helical organization of the monomer and crowder center of mass in terms of tangent-tangent correlation along the respective mean center

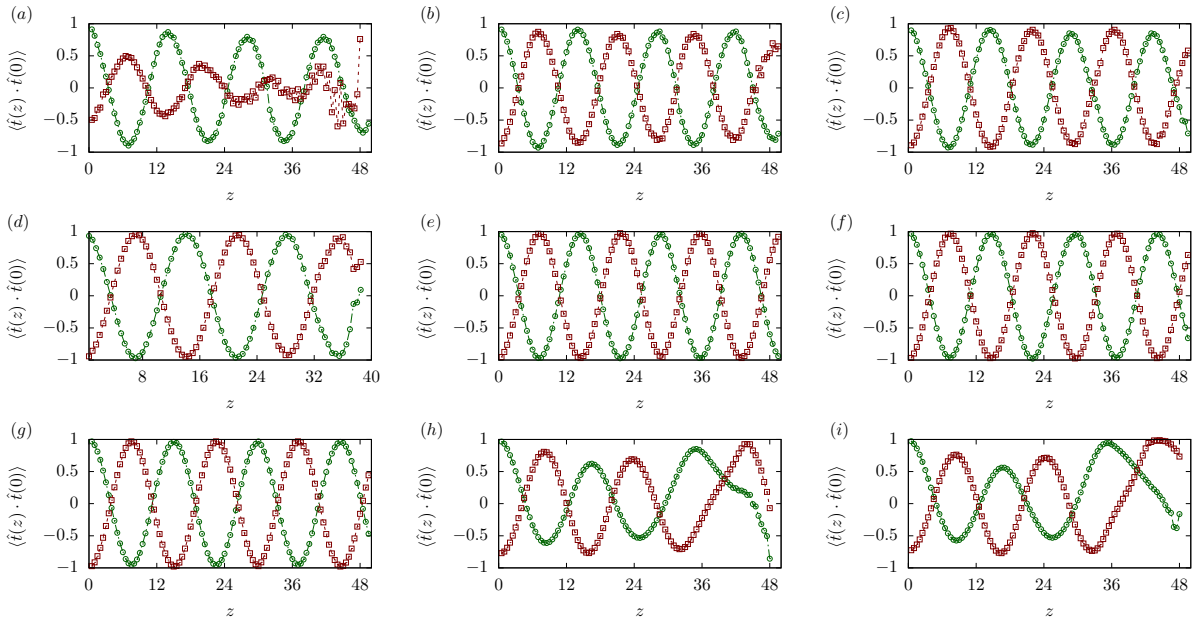


FIGURE 6.5: The correlation $\langle \hat{t}(z) \cdot \hat{t}(0) \rangle$ of the $xy-$ plane projected average center of mass orientation of monomers (green ○), and it's cross-correlation $\langle \hat{t_c}(z) \cdot \hat{t}(0) \rangle$ with that of crowder center of mass loci (brown □) is plotted. Different plots are at $\rho_c \sigma^3$ value, 0.01 (a), 0.03 (b), 0.06 (c), 0.07 (d), 0.12 (e), 0.14 (f), 0.17 (g), 0.2 (h) and 0.24 (i) respectively.

of mass loci, and the cross-correlation of monomer center of mass with crowder center of mass loci. For this purpose, we project the two sets of center of mass loci onto the xy-plane. We use these to compute the tangent vectors $\hat{t}(z)$ and the corresponding tangent-tangent correlation function $\langle \hat{t}(z) \cdot \hat{t}(0) \rangle$. z here represents the location of unit vector along the long axis of the cylinder. For the monomers, the auto-correlation function is shown in Fig. 6.5(a)-(i) denotes by green ○. This shows periodic oscillations capturing the helical turns. Next, we consider the relative orientation vector of crowder center of mass with respect to monomer center of mass $\hat{t_c}(z)$ at different z-planes. We calculate the cross- correlation $\langle \hat{t_c}(z) \cdot \hat{t}(0) \rangle$, where $\hat{t}(0)$ denotes the unit vector between the monomer center of mass positions. The cross-correlations at different crowder densities are shown in Fig. 6.5(a)-(i) denoted by brown □. The cross-correlations show out-of-phase oscillations compared to the tangent auto-correlation functions in monomers. However, they share the same periodicity. This

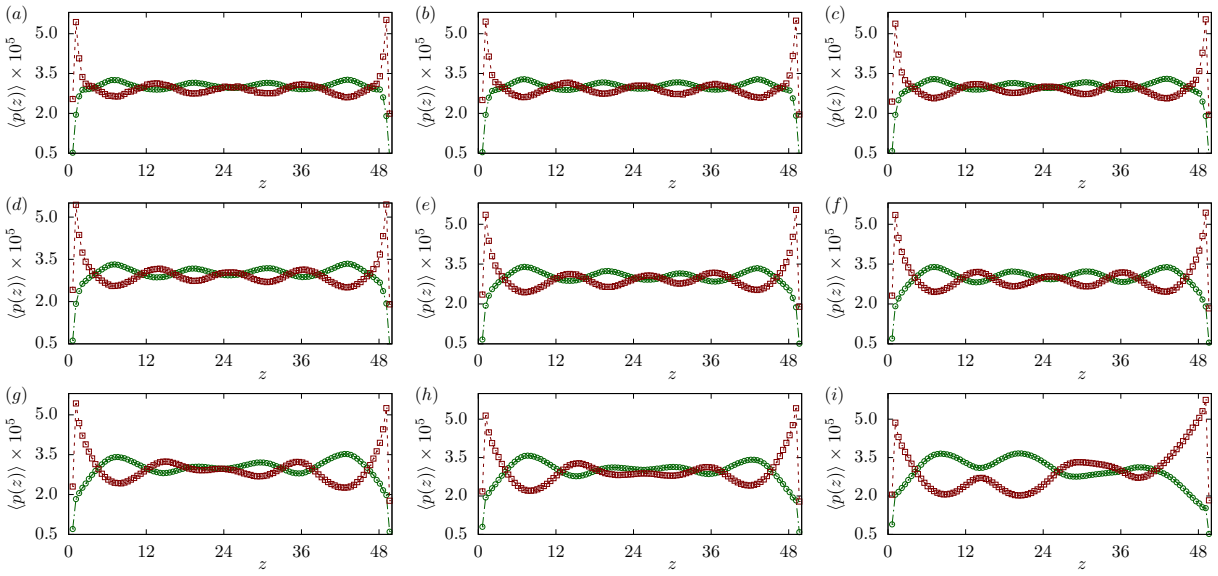


FIGURE 6.6: Probability density $p(z)$ is plotted along the z axis of cylinder. The density for monomers are denoted by monomers (crowders) green ○ (brown □). (a)-(i), $\rho_c\sigma^3$ value corresponds to $0.01, 0.03, 0.06, 0.07, 0.12, 0.14, 0.17, 0.2$ and 0.24 in alphabetical order.

clearly establishes the fact that the center of mass loci for monomers and the crowders are organized in a helical fashion complimentary to each other.

6.2.3 local probability density

Further, their local probability densities along the cylinder length shows periodic density modulations. They share the same periodicity but oscillates out of phase, similar to the behavior of tangent correlations established above. The crowders wet the two caps of the cylinder. This is shown in terms of the maximum (minimum) in crowder (monomer) density at the two caps at $z = 0, L$. The connectivity between the monomers, and possibility of the polymer to explore more conformations if it stays away from the wall leads to the low monomer density at the two caps. On the other hand, the hydrostatic pressure pushes a layer of crowders on the wall, a behavior shared by confined simple liquids. The out- of-phase oscillation of the two probability densities are observed at all densities of crowders examined here.

In this section we have investigated the impact of crowders on the morphology of full feather-boa model of chromosome and relative organization of crowders with respect to chromosome. The crowders have short range repulsion with chromosome which didn't lead to much deviation of chromosome morphology from the helical organization, with increasing crowder density.

In the following section we consider the polymeric nature of the protein crowders within the simplest approximation, and examine the effect of increasing their density on the feather-boa chain.

6.3 Feather-boa chain : crowders with additional Gaussian- core repulsion

The effective interaction between the center of mass of long polymers has a simple description in terms of an approximately Gaussian repulsion [215]. Although monomers of self-avoiding chains cannot cross each- other, the center of mass can, by undergoing conformational changes while crossing each other. This effective potential or free energy is entirely due to entropy of the participating chains. The strength of this potential is around $2k_BT$, and the range is defined by the radius of gyration of the polymers R_g. Instead of linear chains, if one considers circular polymers, the strength of effective repulsion can be few times larger [218]. To incorporate the polymer nature of protein crowders, in this section, we assume an additional repulsion between the monomers and crowders:

$$V_{gc}(r_{ij}) = \mathcal{E}_0 \exp[-r_{ij}^2/2\Sigma_c^2]. \tag{6.4}$$

Here the separation between particles is r_{ij}. The parameters $\mathcal{E}_0$ and Σ_c denote the strength and range of the Gaussian core repulsion. We use $\Sigma_c = \sigma$, and consider $\mathcal{E}_0 = 2k_BT$ and $3k_BT$. Similarly an additional Gaussian core repulsion as in Eq.(6.4) is used between the crowders and the walls, but with half the range. In all other terms, the polymer, crowder

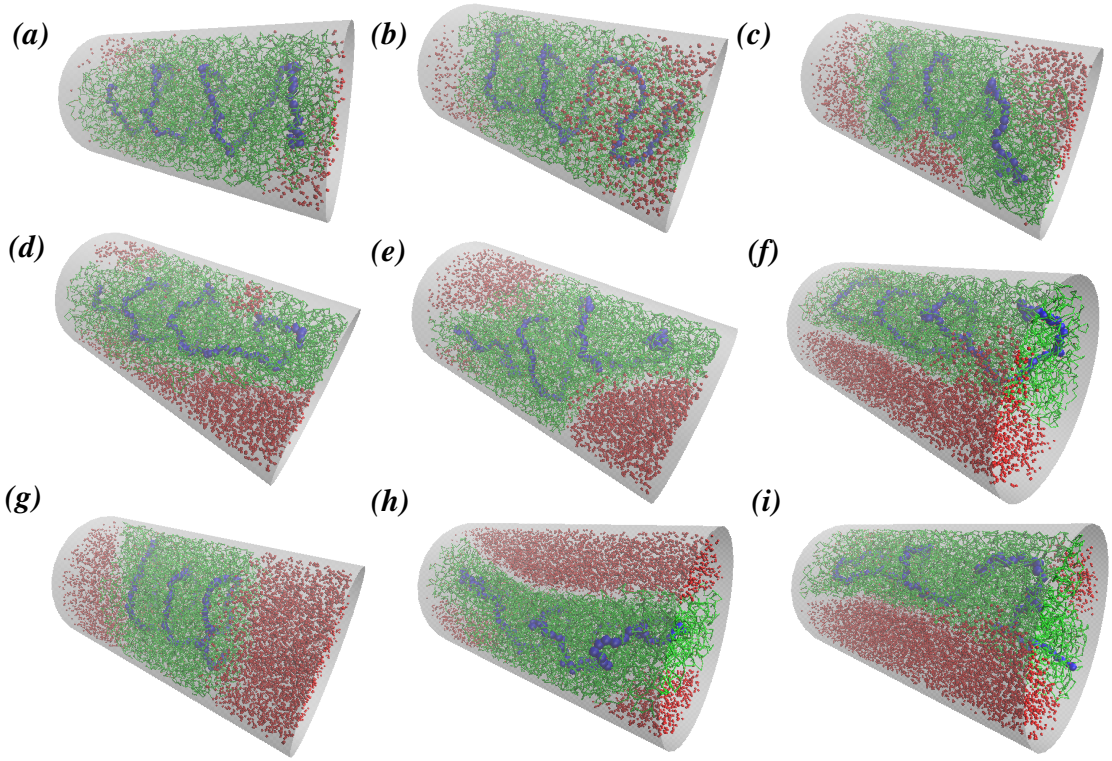


FIGURE 6.7: Snapshots of the simulation at various parameter values are presented. The backbone monomers are shown in blue, side loops are in greed and the crowders by the red dots. For Gaussian core interaction, $\mathcal{E}_0 = 2k_BT$, $\Sigma = \sigma$ are used. (a)-(i), $\rho_c\sigma^3$ value corresponds to $0.01, 0.03, 0.06, 0.07, 0.12, 0.14, 0.17, 0.2$ and 0.24 in alphabetical order.

and confinement model remains exactly the same as in the previous section. We use this model to explore the impact of changing crowder density ρ_c (Figures6.7, and 6.8). In the figures, the backbone of the feather-boa chain is denoted in blue, and the side loops in green. The crowders are shown in red. Clearly, in contrast to the results in the previous section, the relatively larger size of the polymeric crowders show local segregation between chromosome and crowders.

In Fig. 6.7, the snapshots of the simulation with $\mathcal{E}_0 = 2k_BT$ at various crowder density ρ_c are presented. Similar plots with $\mathcal{E}_0 = 3k_BT$ are presented in Fig. 6.8. In this case, we observe significant change in organization with increase in crowder density. The chromosome and

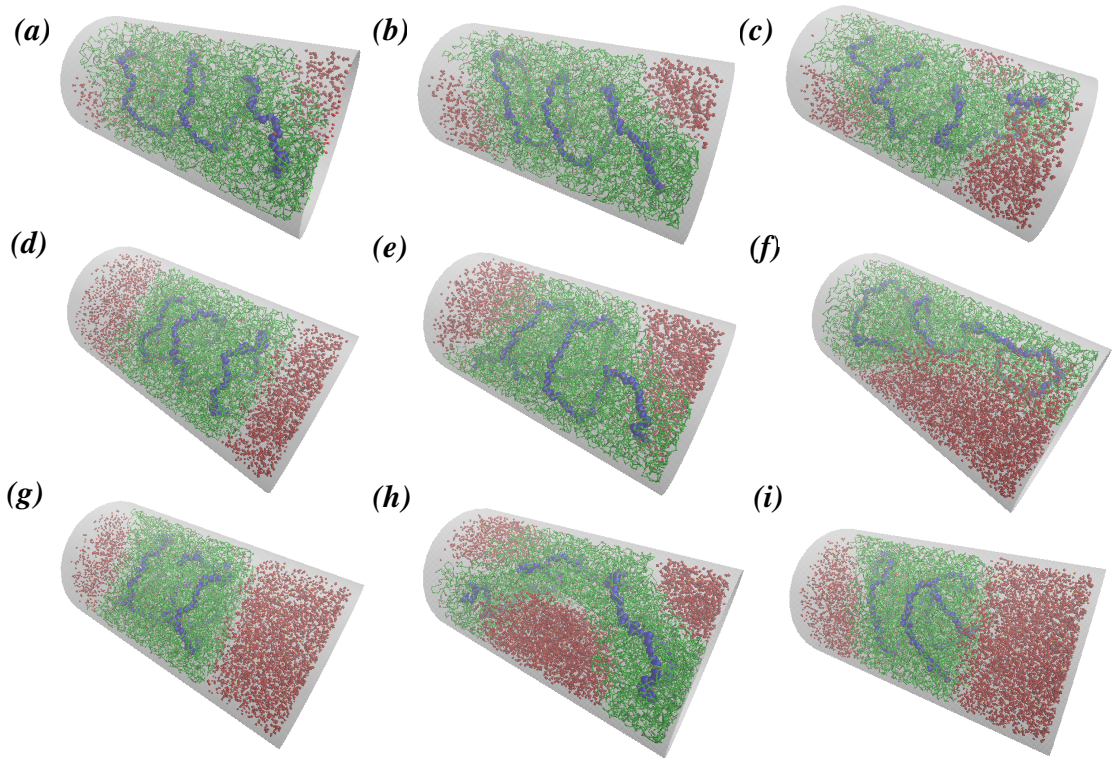


Figure 6.8: Snapshots of the simulation at various parameter values are presented like Fig. 6.7. For Gaussian core interaction, $\mathcal{E}_0 = 3k_BT$, $\Sigma = \sigma$ are used. (a)-(i), $\rho_c\sigma^3$ value corresponds to $0.01, 0.03, 0.06, 0.07, 0.12, 0.14, 0.17, 0.2$ and 0.24 in alphabetical order.

crowders show mostly radial (transverse) segregation at lower densities, which transforms into mostly longitudinal segregation as the density of crowders increase. A broad range of densities show *coexistence* of both kind of segregations carrying strong hysteresis. In this regime, the actual organization observed depends on the initial condition. Also, we find organization in which crowders and chromosome get segregated in a partially longitudinal and partially transverse manner.

In the following we investigate the impact of increasing crowder density in a more systematic manner, using a computationally more tractable coarse-grained model of the chromosome.

6.4 Coarse-grained model

Here we consider a computationally more tractable model, in which we coarse-grain the the feather-boa chain to an effective chain and investigate the impact of density of cytosolic molecular crowders. The coarse-graining of the feather-boa chain is performed by replacing the side loops with an additional Gaussian core repulsion between the monomers of the backbone [10], as described in the previous chapter. Strength and interaction range for this Gaussian interaction depends on topology of chain and radius of gyration of side-loops [215, 218]. The introduction of soft core Gaussian potential effectively thickens the backbone to a soft tube, incorporating approximately the repulsive cloud of monomers due to side-loops. The form of the Gaussian-core repulsion between backbone monomers is given by, $V_{gc}(r_{ij}) = \mathcal{E}_0 \exp[-r_{ij}^2/2\Sigma^2]$, with $\Sigma^2 = 2R_g^2$. Here, r_{ij} denotes the separation between $i\text{-}th$ and the $j\text{-}th$ monomer and R_g is the radius of gyration of each side loop, $R_g = cn_s^{3/5}\sigma$ with c = 0.323 and n_s denoting the number of beads in the Feather-boa side loop [18].

The 4.6 Mbp (1.6 mm) circular genome of E. coli can be modeled as a circular feather-boa chain with the backbone consisting of n_b monomers, each of which is connected to a side loop of n_s monomers. A use of $n_b = 636$ and $n_s = 62$ gives a total chain length 39494σ. Comparing it with the E. coli DNA gives an effective bead diameter $\sigma \approx 116\,\text{bp} = 0.04\,\mu\text{m}$. Recent experiments utilized filamentous E. coli cells that grow up to about $30\,\mu$m to investigate the impact of confinement [9]. Here we consider the circular model chromosome inside a cylindrical confinement corresponding to a cell- length of $L = 12.8\mu\text{m} = 320\sigma$ and diameter $D = 1\mu\text{m} = 26.67\sigma$. The side loops has radius of gyration $R_g = cn_s^{3/5}\sigma = 0.14D$ in terms of this cylinder diameter.

As described before, the effective interaction between self- avoiding polymers can be described approximately in terms of a Gaussian core repulsion. However, the repulsion strength depends on whether the chain is open or close, as well as on their topology. For long linear chains in bulk, the repulsion strength is $2k_BT$, while for circular chains it varies from 2-6 k_BT.

We fix an intermediate value for strength $\mathcal{E}_0 = 3k_BT$ to account for loops as well as linear chain- like plectoneme structures. In the interaction between the backbone monomers and walls, we consider the thickening due to side- loops. Thus in addition to V_{wall} we use a Gaussian core repulsion between the backbone monomers and walls, $V_{gc}(r_{iw}) = \frac{\mathcal{E}_0}{2}\exp[-r_{iw}^2/2\Sigma^2]$ with $\Sigma^2 = R_g^2$, i.e., half the strength and range with respect to the additional repulsion between the backbone monomers due to side- loops. Here, r_{iw} is the separation of the monomer with the walls of the cylinder.

As before, we use non-additive crowders that does not interact between themselves, but repel the chromosome with short range WCA potential and an additional Gaussian-core potential, $V_{gc}^c(r_{ij}) = \mathcal{E}_0\exp[-r_{ij}^2/2\Sigma_c^2]$, where r_{ij} is the separation between interacting crowder and the monomer of the chain. We use $\mathcal{E}_0 = 3k_BT$ and $\Sigma_c = 3.5\,\sigma$ and vary the number density of crowders $\rho_c = 4N_c/(\pi D^2 L)$. We vary the density $\rho_c\sigma^3$ over an order of magnitude from 0.002 to 0.04 by using the number of crowders N_c = 200, 400, 800, 1200, 1600, 2000, 2400, 2800, 3200, 3600, 4000, 4800, 5600, 6400, 7200. In the rest of this section we use the $\rho_c D^3$ as a dimensionless measure of density.

We use molecular dynamics simulations performed using velocity-Verlet algorithm with step size $\delta t = 0.01\,\tau$. The temperature of the heat bath is maintained by a Langevin thermostat which fixes the temperature $T = 1.0\epsilon/k_B$, as implemented by the ESPResSo package [161]. In the simulations, the configurations are equilibrated over $10^7\tau$ and data is gathered over next 2×10^4 configurations at a separation of $10^3\tau$.

6.4.1 Impact of crowder densities on organization

In Fig. 6.9, snapshots of the equilibrated configurations are shown at various crowder densities. The red line represents the coarse-grained polymer and the non-additive crowders are shown by blue beads. Fig. 6.9(a), shows the configurations at lower crowder densities. The first snapshot from the top, at $\rho_c D^3 = 42$, displays an open conformation of the chromosome. Due to small density ($N_c = 400$), the crowders get homogeneously distributed throughout

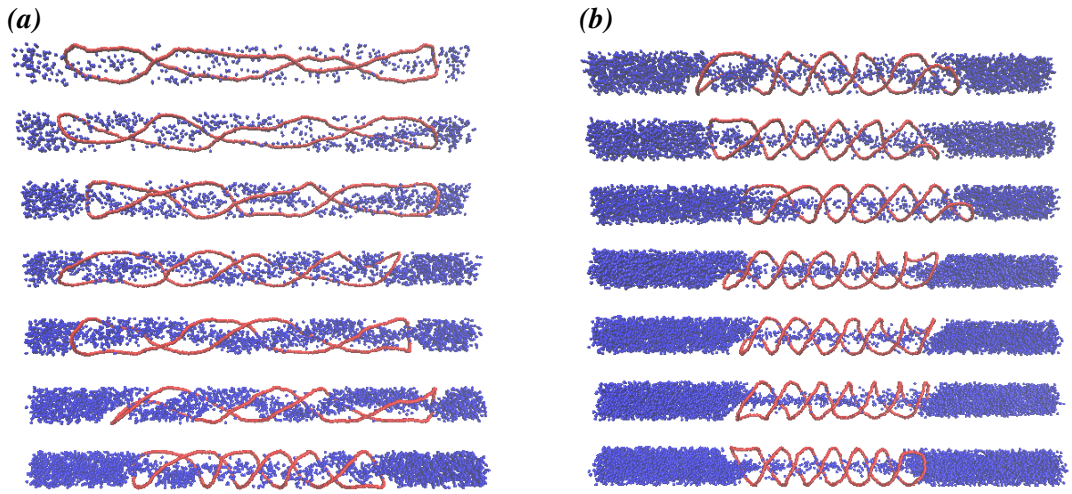


FIGURE 6.9: Snapshots of the simulation is presented. The chromosome is shown in red line and crowders in blue dots. (a) Top to bottom, snapshots are shown at crowder densities $\rho_c D^3 = 42, 85, 127, 170, 212, 255$ and 297 respectively. (b) Top to bottom, snapshots are shown at densities $\rho_c D^3 = 339, 382, 424, 509, 594, 679$ and 764 respectively.

the cylindrical cell and the chromosome opens up. At higher density, $\rho_c D^3 = 85$ and 127 ($N_c = 800$ and 1200) the figures show a onset of a weak helical organization of chromosome. At $\rho_c D^3 = 170$, 212 and 255 the helical organization gets more pronounced, and we find a remarkable longitudinal and radial segregation of chromosome and crowders. The longitudinal segregation locating a higher density of crowders near the cylinder caps provide a compressing pressure to stabilize the helical organization of the chromosome. The radially segregated crowders at the mid- section of the cell organize into a distinctive helical shape, running complementary to the chromosomal helix. Note that the complementary helicity observed at density $\rho_c D^3 = 212$ correspond to the same observation at $\Sigma_c = 3.5\sigma$ discussed in the previous chapter. At further higher density, the chromosome gets more and more compressed as the segregation becomes predominantly longitudinal. Note that the crossover in segregation from a predominantly radial to a predominantly longitudinal with increasing crowder density is in agreement with the prediction in Fig.6.14. Unlike in the previous chapter, where we observed complete longitudinal segregation of crowders from the chromosomal

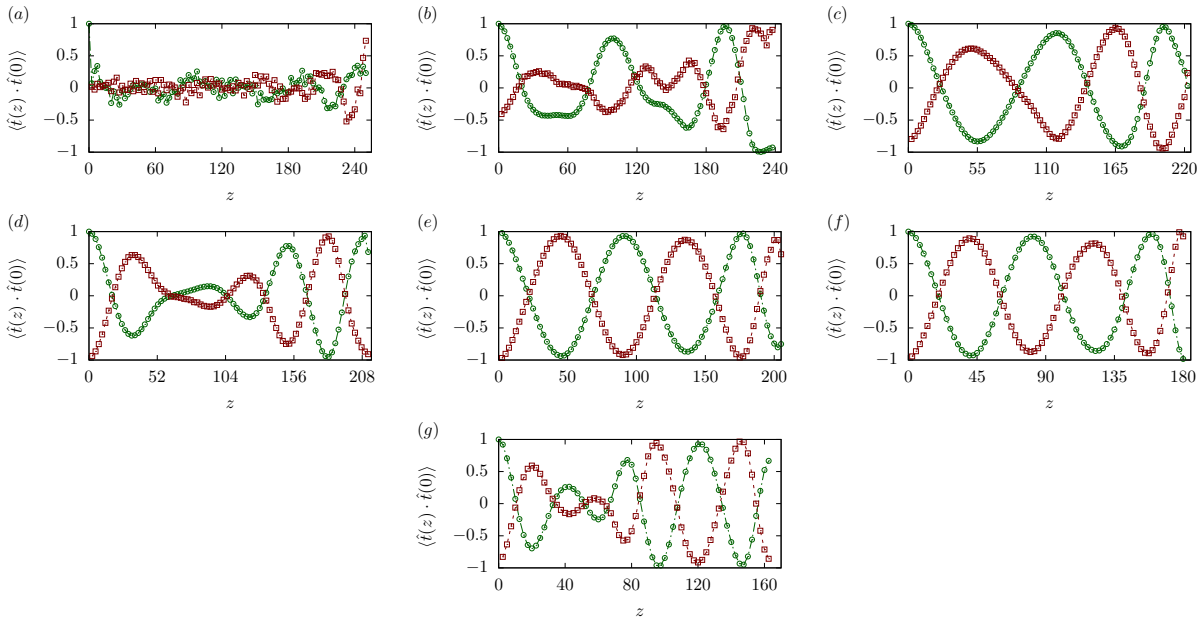


FIGURE 6.10: The correlation $\langle \hat{t}(z) \cdot \hat{t}(0) \rangle$ of the xy-plane projected average center of mass orientation of monomers (green $\circ$), and it's cross-correlation $\langle \hat{t_c}(z) \cdot \hat{t}(0) \rangle$ with that of crowder center of mass loci (brown $\square$) is plotted. Different plots are at $\rho_c D^3$ value, 42 (a), 85 (b), 127 (c), 170 (d), 212 (e), 255 (f), 339 (g) respectively.

region, small number of crowders remain localized along a line along the long axis.

6.4.2 Complimentary helicity

In the intermediate range of crowder density ($127 \leq \rho_c D^3 \leq 339$), we found complimentary helical organization of crowders and the chromosome. In this section, we quantify such organization. For this purpose we consider disk-shaped bins along the long axis of the cylinder, the z-axis. We find the center of mass position of the monomers and crowders in each bin projected onto the radial xy-plane. The separation vector of these center of mass positions in consecutive bins are used to obtain local tangent vectors $\hat{t}(z)$ for monomers and $\hat{t}_c(z)$ for crowders. We use them to calculate the autocorrelation function $\langle \hat{t}(z) \cdot \hat{t}(0) \rangle$, and the cross-correlation function $\langle \hat{t_c}(z) \cdot \hat{t}(0) \rangle$. The averaging in the correlations are performed using 2×10^4 equilibrated configurations. These correlations are shown in Fig. 6.10.

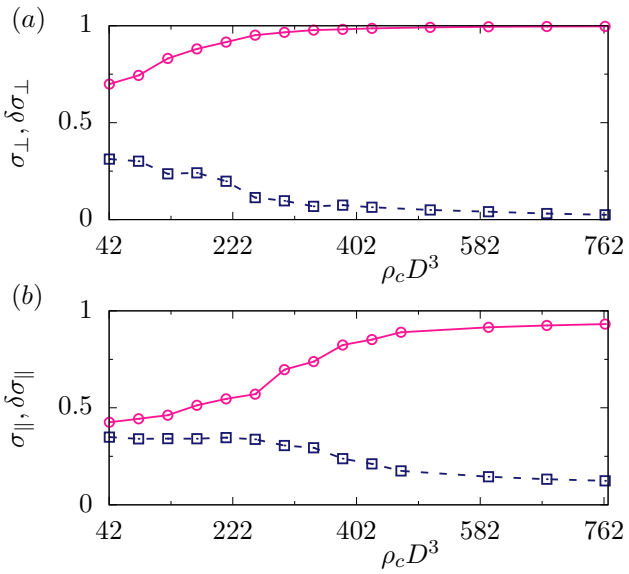


FIGURE 6.11: (a) Measure of transverse spatial segregation of chromosome and crowders $\sigma_\perp$ is shown in pink $\circ$ along with corresponding root mean squared fluctuation $\delta\sigma_\perp$ (blue $\square$) as a function of $\rho_c D^3$. (b) Measure of longitudinal spatial segregation of chromosome and crowders $\sigma_\parallel$ is shown in pink $\circ$ along with corresponding root mean squared fluctuation $\delta\sigma_\parallel$ (blue $\square$) as a function of $\rho_c D^3$.

Fig. 6.10(a) show lack of structure at $\rho_c D^3 = 42$. At higher densities periodic oscillations in the two correlations start to appear. These are most pronounced at $\rho_c D^3 = 212,\ 255$ as shown in Fig. 6.10(e),(f). The out of phase oscillations in the two correlations in Fig. 6.10(c)-(g) characterizes complementary helical organizations of the chromosome and crowders. Increase of density beyond $\rho_c D^3 = 339$, leads to full longitudinal segregation of crowders and the chromosome, where the complimentary helicity disappears.

6.4.3 Transverse and longitudinal segregation

As we have observed, at small $\rho_c D^3$, the crowders were homogeneously distributed and chromosome opens up. At higher densities we find radial, followed by a longitudinal segregation of the crowders and chromosome. These can be quantified in terms of the following measure:

$$\sigma_{\parallel,\perp} = \left\langle \left| \frac{\rho_m - \rho_c}{\rho_m + \rho_c} \right| \right\rangle .$$

To compute the expressions $\sigma_{\parallel,\perp}$, the cylinder is binned into longitudinal and transverse sections. The binning procedure is the same as described in previous chapter. In transverse binning, the cylindrical volume is binned into 10 co-axial cylindrical shells. In longitudinal binning, we bin the cylinder into 120 consecutive disks along the long axis of cylinder. $\rho_{m,c}$ represent the local densities of the monomers and the crowders within each bin respectively. The local densities are normalized by respective total number of monomers N_m and the crowders N_c. For the computation of transverse segregation of monomers and crowders, we restrict our calculation to the central region of cylinder where the chromosome is present. In each bin, the quantity $\sigma_\perp$ is calculated and averaging is performed over 2×10^4 equilibrium configurations. The longitudinal segregation is quantified using $\sigma_\parallel$ in each longitudinal bin and averaging it over the same number of equilibrium configurations.

In Fig. 6.11, we present the variation of $\sigma_{\parallel,\perp}$ and there fluctuations $\delta\sigma_{\parallel,\perp} = \sqrt{\langle\sigma^2_{\parallel,\perp}\rangle - \langle\sigma_{\parallel,\perp}\rangle^2}$, with crowder density $\rho_c D^3$. The transverse segregation sets in at a low density $\rho_c D^3 \approx 85$ as can be seen by following $\sigma_\perp$, which saturates by $\rho_c D^3 \approx 297$. On the other hand, the longitudinal segregation quantified by $\sigma_\parallel$ start to set in around $\rho_c D^3 \approx 297$ and saturates at a further higher density $\rho_c D^3 \approx 509$. As $\sigma_{\parallel,\perp}$ increases, their fluctuations $\delta\sigma_{\parallel,\perp}$ reduces to finally saturate as the segregation progresses.

6.4.4 Chromosome extension

Finally, we plot the chromosome extension along the cell length $\langle R_\parallel \rangle$ with increasing $\rho_c D^3$ in Fig. 6.12. The extension in a configuration is calculated in terms of the longest separation between monomers along the length of the cylinder. The averaging is performed over 2×10^4 equilibrium configurations. The longitudinal segregation of the crowders from chromosome is present all through, although it becomes the dominant mechanism of segregation at higher densities, $\rho_c D^3 > 297$. The compression due crowders segregated at the caps of the cylinder, reduces $\langle R_\parallel \rangle$ with increasing $\rho_c D^3$. In Eq.(6.15) we derived an approximate closed- form expression for $\langle R_\parallel \rangle$. Its dependence on crowder density has a form $\langle R_\parallel \rangle \approx \mathcal{A}[\sqrt{b^2 + \mathcal{B}\rho_c D^3} -$

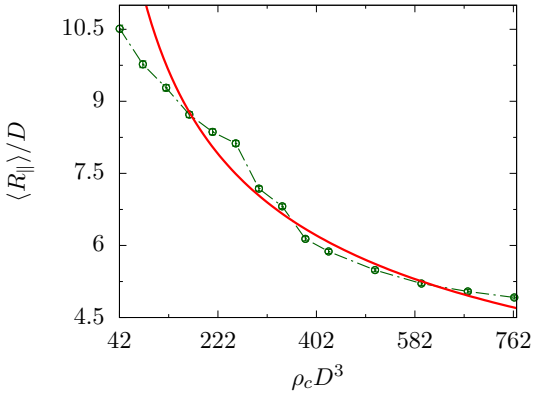


FIGURE 6.12: The mean extension of chromosome along cylinder length $\langle R_{\parallel}\rangle$ is plotted against $\rho_c D^3$. The fit to $\langle R_{\parallel}\rangle \approx \mathcal{A}[\sqrt{b^2+\mathcal{B}\rho_c D^3}-b]/\rho_c D^3$ (Eq. 6.15) for $\rho_c D^3 \gtrsim 297$, with $\mathcal{A} = 376.83$, $\mathcal{B} = 106.95$ and $b = 33.44$, is shown by the red line.

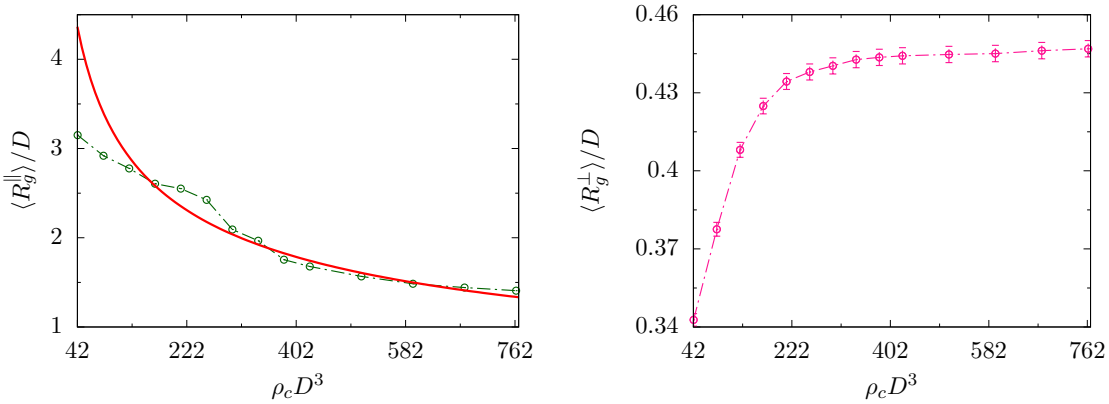


FIGURE 6.13: Components of radius of gyration of chromosome is plotted against $\rho_c D^3$. $\langle R_g^{\parallel}\rangle$ is along long axis of cylinder and $\langle R_g^{\perp}\rangle$ is along radial direction. The fit to $\langle R_g^{\parallel}\rangle \approx \mathcal{A}[\sqrt{b^2+\mathcal{B}\rho_c D^3}-b]/\rho_c D^3$ (Eq. 6.15) for $\rho_c D^3 \gtrsim 297$, with $\mathcal{A} = 214.27$, $\mathcal{B} = 24.98$ and $b = 11.59$, is shown by the red line.

$b]/\rho_c D^3$ (Eq.6.15 in the following section), a function which is fitted to the simulation data in Fig. 6.12. This captures the behavior of the data after longitudinal segregation. The change in slope at $\rho_c D^3$ corresponds to the crossover from predominantly radial segregation at lower density to longitudinal segregation at higher densities.

6.4.5 Mean field description

Here we present a mean field description of the dependence of chromosome size on crowder density, allowing the longitudinal and transverse segregation of the polymer and crowders. Under longitudinal segregation the polymer gets squeezed along the cell length with the extension $R_\parallel$ decreasing with crowder density. When the polymer gets segregated from the crowder in the radial direction, it polymer covers a fraction αD with $\alpha < 1$ of the diameter D of the cylindrical confinement.

Treating the chromosome as a self- avoiding chain in a cylindrical volume, its free energy may be written down in terms of the de Gennes' blob picture. Let the polymer occupy a cylindrical space of diameter αD. It consists of N/g blobs where each blob contains g monomers out of N available in the chain. Assuming the extension $R_\parallel$, the free energy of the chain is expressed as

$$\beta F_c = A_1 \frac{R_\parallel^2}{(N/g)\alpha^2 D^2} + B_1 \frac{\alpha D (N/g)^2}{R_\parallel}.$$

The three-dimensional Flory scaling is maintained within a blob, $g \sim (\alpha D)^{5/3}$. Using this in the above expression we get,

$$\begin{aligned} \beta F_c &= A \frac{R_\parallel^2}{N(\alpha D)^{1/3}} + B \frac{N^2}{(\alpha D)^{7/3} R_\parallel}, \\ &= \frac{a}{2} \frac{\bar{R}_\parallel{}^2}{\alpha^{1/3}} + \frac{b}{\bar{R}_\parallel \alpha^{7/3}}, \end{aligned} \tag{6.5}$$

where in writing the second line we used dimensionless extension $\bar{R}_\parallel := R_\parallel/D$, $a = 2AD^{5/3}/N$

and $b = BN^2/D^{10/3}$. In a system with constant diameter D and chain length N these dimensionless parameters control the chain properties.

The volume occupied by chromosome is $C_d(\alpha D)^2 R_\parallel$, where the geometrical pre-factor $C_d = \pi/4$ for cylinders and $C_d = 1$ for rectangular parallelepiped. Thus the volume available to crowders

$$V_d = C_d D^2 (L - \alpha^2 R_\parallel) = C_d D^3 (\bar{L} - \alpha^2 \bar{R}_\parallel). \tag{6.6}$$

This relation works in the limit of crowder size $\Sigma_c \ll R_\parallel, \alpha D$. For finite Σ_c the space occupied by the chromosome will have its impact, $V_c = C_d(\alpha D + \Sigma_c)^2(R_\parallel + \Sigma_c)$. As a result the free volume modifies to

$$\begin{aligned} V_d &= C_d D^2 L - C_d(\alpha D + \Sigma_c)^2 (R_\parallel + \Sigma_c) \\ &= C_d D^3 \left[\bar{L} - (\alpha + \bar{\Sigma}_c)^2 (\bar{R}_\parallel + \bar{\Sigma}_c)\right]. \end{aligned} \tag{6.7}$$

This relation is valid only when $R_\parallel + \Sigma_c < L$ and $\alpha + \bar{\Sigma}_c < 1$, otherwise $R_\parallel = L$ ($\alpha = 1$). The free energy of the non-additive crowders of size Σ_c can be written as

$$\beta F_d = N_d \left[\ln \frac{N_d \Sigma_c^3}{V_d} - 1\right], \tag{6.8}$$

where, $N_d = \rho_c C_d D^2 L$. Let us express all lengths in the units of diameter D such that

$$\beta F_d = C_d \bar{L} \bar{\rho}_c \left[\ln \frac{\bar{L}\ \bar{\rho}_c \bar{\Sigma}_c^3}{\bar{L} - (\alpha + \bar{\Sigma}_c)^2 (\bar{R}_\parallel + \bar{\Sigma}_c)} - 1\right] \tag{6.9}$$

where, $\bar{L} = L/D$, $\bar{\rho}_c = \rho_c D^3$, $\bar{\Sigma}_c = \Sigma_c/D$ and $\bar{R}_\parallel = R_\parallel/D$.

For longitudinal segregation, we minimize the total free energy $F_\ell = F_c + F_d$ with respect

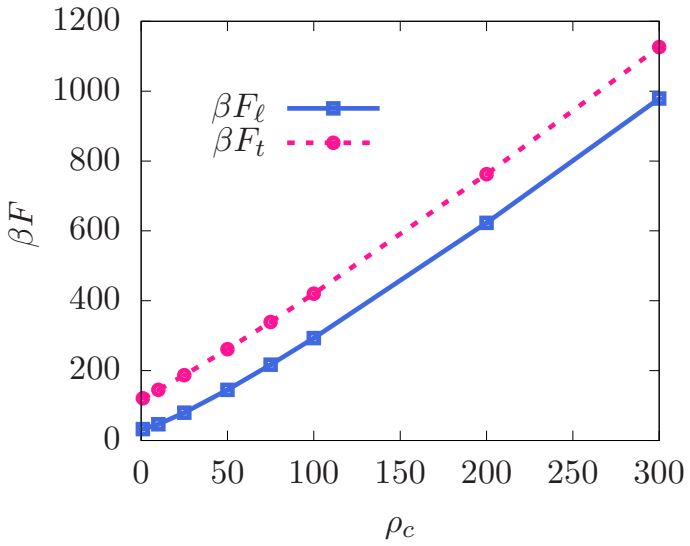


FIGURE 6.14: Comparison of free energies βF_ℓ and βF_t for longitudinal and transverse segregation, respectively. We use dimensionless parameters $a = 1.55$, $b = 82.47$, $c = 0.53$ for a cell of $L = 12.0\,D$.

to $\bar{R}_\parallel$ to obtain the force balance equation,

$$\frac{\partial \beta F_\ell}{\partial R_\parallel} = a\frac{\bar{R}_\parallel}{\alpha^{1/3}} - \frac{b}{\alpha^{7/3}{\bar{R}_\parallel}^2} + C_d\bar{L}\bar{\rho}_c\frac{(\alpha+\bar{\Sigma}_c)^2}{[\bar{L}-(\alpha+\bar{\Sigma}_c)^2(\bar{R}_\parallel+\bar{\Sigma}_c)]} = 0. \tag{6.10}$$

Similarly, for transverse segregation, we minimize $F_t = F_c + F_d$ with respect to α to obtain,

$$\frac{\partial \beta F_t}{\partial \alpha} = -\frac{1}{6}a{\bar{R}_\parallel}^2\alpha^{-4/3} - \frac{7b}{3\bar{R}_\parallel}\alpha^{-10/3} + 2\,C_d\bar{L}\bar{\rho}_c\frac{(\bar{R}_\parallel+\bar{\Sigma}_c)(\alpha+\bar{\Sigma}_c)}{[\bar{L}-(\alpha+\bar{\Sigma}_c)^2(\bar{R}_\parallel+\bar{\Sigma}_c)]} = 0. \tag{6.11}$$

In the limit of $\bar{\Sigma}_c \ll \bar{R}_\parallel, \alpha$ The above two relations Eq.(6.10) and (6.11) get independent of Σ_c. In general, equations (6.10) and (6.11) are to be solved simultaneously to obtain the chromosome size in the longitudinal and transverse directions.

To simplify the exposition, we set $(\alpha + \bar{\Sigma}_c) = 1$ in Equation (6.10), and $\bar{R}_\parallel + \bar{\Sigma}_c = \bar{L}$ in Equation (6.11), separately, and define the corresponding free energies as longitudinal and radial free energies, and compare. Thus for a given $\bar{\rho}_c$, we can use the above equations to determine $\bar{R}_\parallel(\bar{\rho}_c)$ and $\alpha(\bar{\rho}_c)$ respectively, and obtain the free energies in the longitudinal and

radial sectors,

$$\beta F_\ell(\rho_c) = \frac{a}{2}\bar{R_\parallel}^2 + \frac{b}{\bar{R_\parallel}} + C_d\bar{L}\,\bar{\rho}_c\,\left[\ln\frac{\bar{L}\,\bar{\rho}_c\bar{\Sigma}_c^3}{\bar{L}-(\bar{R_\parallel}+\bar{\Sigma}_c)} - 1\right] \tag{6.12}$$

and

$$\beta F_t(\rho_c) = \frac{a}{2}\frac{\bar{L}^2}{\alpha^{1/3}} + \frac{b}{\bar{L}\alpha^{7/3}} + C_d\bar{L}\,\bar{\rho}_c\,\left[\ln\frac{\bar{\rho}_c\bar{\Sigma}_c^3}{1-(\alpha+\bar{\Sigma}_c)^2} - 1\right], \tag{6.13}$$

respectively. A comparison between them can be used to determine if the segregation would take place in a longitudinal or transverse fashion. For example, using dimensionless parameters $a = 1.55$, $b = 82.47$, $c = 0.53$ for a cell of $L = 12.0\,D$, as in the coarse-grained simulations in Fig.6.13. The plots of the two free energies βF_ℓ, βF_t as a function of ρ_c minimized for $R_\parallel$ and α respectively, are shown in Fig.6.14. Within the parameter regimes considered, $F_\ell < F_t$ suggesting a longitudinal segregation between chromosome and crowders.

In the system showing longitudinal segregation of crowders and chromosome, one can find a closed form expression of the decreasing chromosome size with crowder density by performing the following approximation. In the limit of large N, $N_d \gg 1$, neglecting $a \sim 1/N$, Eq.(6.10) leads to a quadratic equation $C_d\bar{L}\bar{\rho}_c\bar{R_\parallel}^2 + b\bar{R_\parallel} - b(\bar{L}-\bar{\Sigma}_c) = 0$. This equation has a closed form solution

$$\bar{R_\parallel} \approx \left[\sqrt{b^2 + 4bC_d\bar{L}(\bar{L}-\bar{\Sigma}_c)\bar{\rho}_c} - b\right]/2C_d\bar{L}\bar{\rho}_c. \tag{6.14}$$

Assuming $\bar{L} \gg \bar{\Sigma}_c$, the relation simplifies to

$$\bar{R_\parallel} \approx \left[\sqrt{b^2 + 4bC_d\bar{L}^2\bar{\rho}_c} - b\right]/2C_d\bar{L}\bar{\rho}_c. \tag{6.15}$$

This relation describes the variation of extensions $R_\parallel$ and $R_\parallel^g$ as depicted in Fig.s 6.12 and 6.13 fairly well.

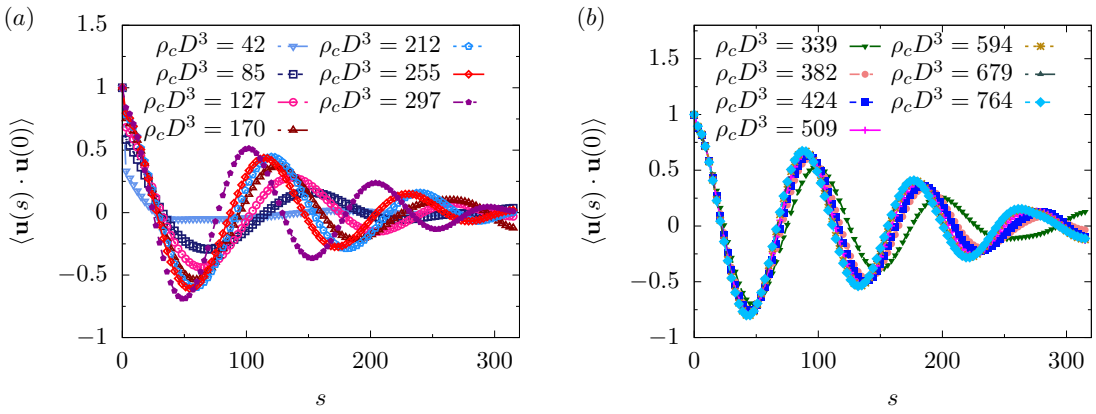


FIGURE 6.15: The tangent-tangent correlation, $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$ along the chain is shown where s represents the separation along the chain contour. (a) shows plot at smaller $\rho_c D^3$ compared to (b).

In the limit of both long L or large N_d, this relation simplifies further to $R_\parallel \sim \rho^{-1/2}$, indicating how the size of a nucleoid-like sub-volume is maintained by the crowders. Further, in the limit $\rho_c \to 0$, the first two terms in Eq.(6.10), along with $\alpha = 1$, leads to the relation

$$(a\bar{R}_\parallel{}^3 - b)(\bar{L} - \bar{R}_\parallel) = 0,$$

due to the first two terms. This gives $\bar{R}_\parallel = (b/a)^{1/3}$ leading to $R_\parallel \sim ND^{-2/3}$ the de-Gennes' scaling form for long cylinders with $L > R_\parallel$, and $R_\parallel = L$ otherwise.

6.4.6 Change in helicity

To better quantify the change in helicity of the chromosome with increasing crowder density, we use the tangent vectors on the chromosome projected onto the transverse plane $\mathbf{u}(s)$, where s denotes a position along the contour. We calculate the correlation function $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$ averaged over the equilibrium configurations. The correlations are plotted up to the longest separation on the circular chain, i.e., half the chain length $n_b\sigma/2$ (Fig.6.15). For better visibility, we plot the relatively low density $\rho_c D^3$ correlations in Fig.6.15(a). It

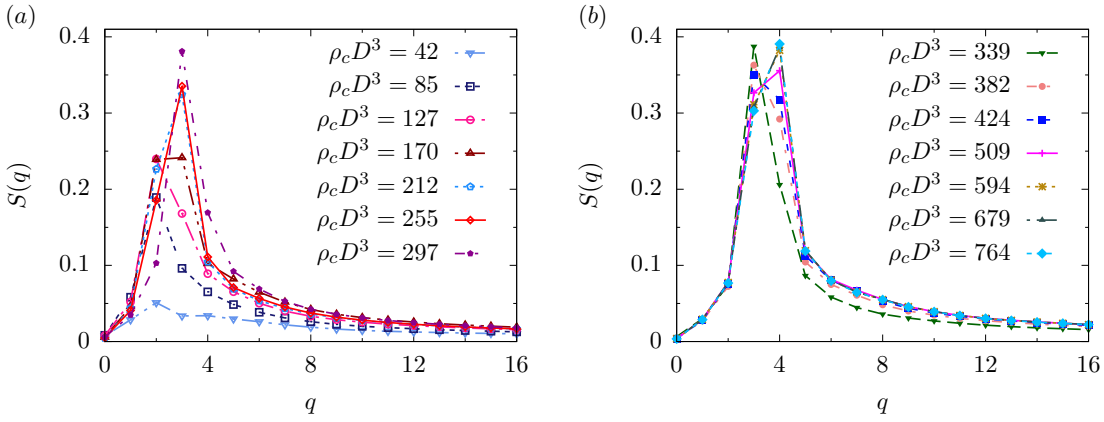


FIGURE 6.16: The structure factor $S(q)$ corresponding to $\langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle$ is plotted. (a) shows plot at smaller $\rho_c D^3$ compared to (b).

clearly shows the reduction in the oscillation period (helical pitch), and increase in the amplitude of oscillation (helicity) at higher density. However, beyond $\rho_c D^3 = 339$, the behavior saturates (see Fig.6.15(b)).

Performing Fourier transform of the correlation function, we obtain the structure factor $S(q) = \int_0^{l_b/2} \langle \mathbf{u}(s) \cdot \mathbf{u}(0) \rangle e^{isq} ds$ (Fig.6.16). The location of the maximum determines the total number of turns $n_t = 2q_p$ for the circular chromosome. As is evident from Fig.6.16(a), the number of turns n_t increases with density $\rho_c D^3$, and the degree of helicity measured by $S(q_p)$ also increases monotonically with density. The relative change of helicity with density slows down as the crowder density increases above $\rho_c D^3 = 339$. In Fig.6.17 we show the variation of n_t with $\rho_c D^3$.

6.4.7 Winding number, turning number and kinks

The helical organization can be quantified in an alternative manner using the turning number. For this purpose we consider the two- dimensional (2d) projection of the polymer conformation on the radial plane of the cylindrical cell. The winding number can be defined

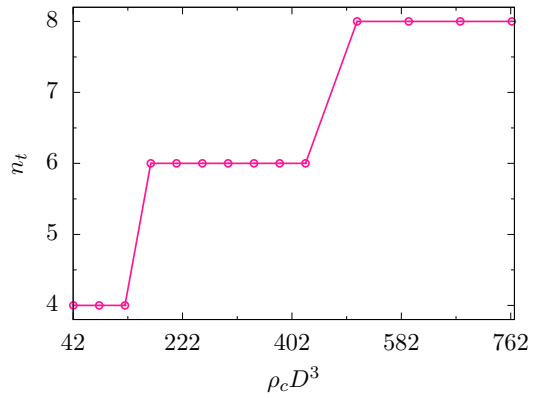


FIGURE 6.17: The number of helical turns $n_t = 2q_p$ is plotted as a function of $\rho_c D^3$.

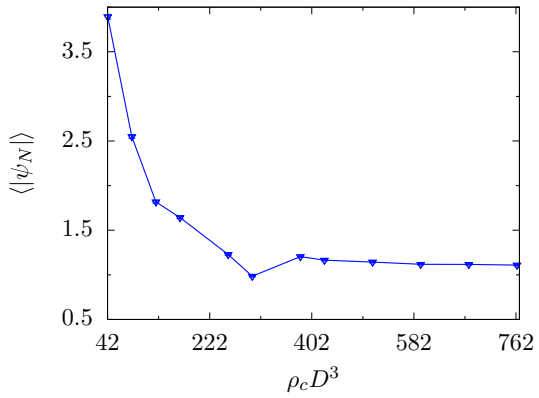


FIGURE 6.18: Mean of absolute winding number $\langle|\psi_N|\rangle$ is plotted against $\rho_c D^3$.

as [219],

$$\psi_N = \frac{1}{2\pi}\sum_{i=0}^{N-1}[\theta_{i+1} - \theta_i], \tag{6.16}$$

where, the angle θ_i denotes the angle subtended by the i-th bond with x-axis, such that the unit tangent vector on the bond can be expressed as $\mathbf{u}_i = (\cos\theta_i, \sin\theta_i)$. In the above expression $\Delta\theta_i = [\theta_{i+1} - \theta_i]$ gives the increment in angle between the consecutive bonds. The quantity ψ_N can be both positive or negative depending on whether the local turns

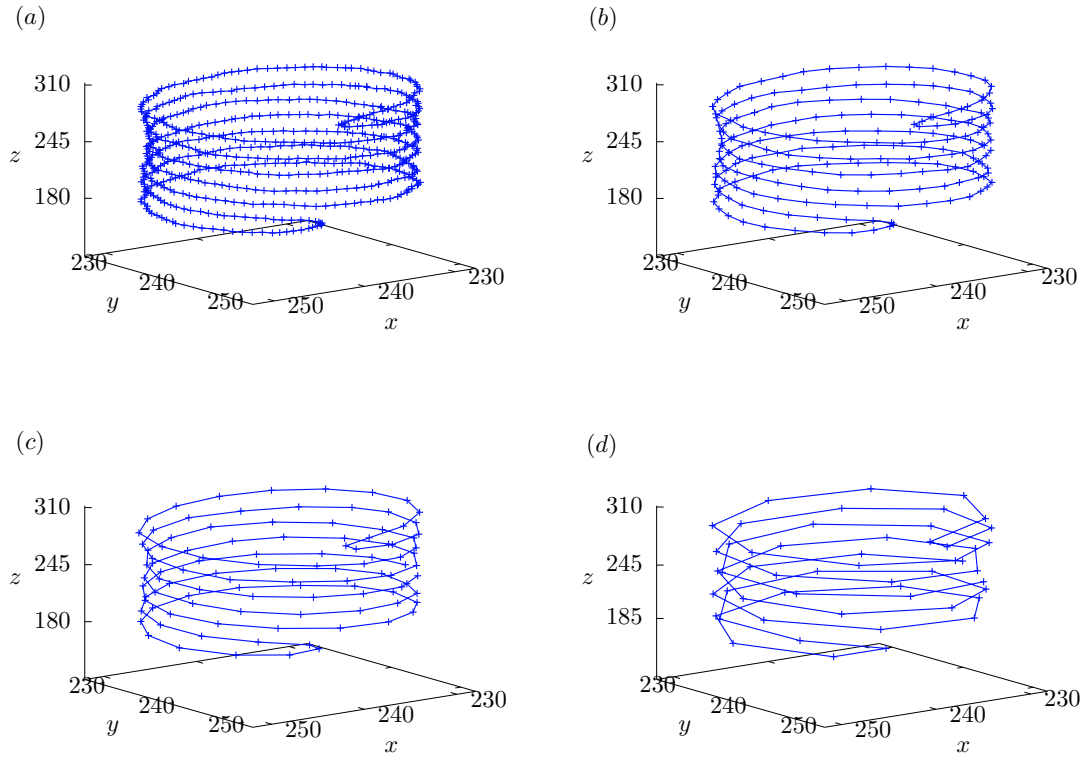


FIGURE 6.19: Chain configuration is shown at $\rho_c D^3 = 764$. (a) shows configuration in absence of coarse-graining. (b)-(d) represents configuration for coarse-grain size $n_{seg} = 3, 6$ and 12.

appear in anti- clockwise or clockwise directions. This accounts for both the gradual change in angle, and steep change due to kinks. If along the chain the curve changes from clockwise to anti-clockwise winding, the overall winding number may turn out to be vanishingly small even after several turns. For example, the plot of the mean absolute value of winding number $\langle|\psi_N|\rangle$ as a function of crowder density $\rho_c D^3$ in Fig. 6.18 shows significantly smaller values with respect to n_t plotted in Fig. 6.17. Note that the winding number ψ_N is a real number, unlike the integer n_t obtained in the previous section by analyzing tangent correlations.

A closer look at instantaneous helical configurations of the circular chain reveals presence of kinks stabilized by confinement and crowder density. Two most prominent kinks are clearly observed near the two caps of the cylinder at $z = 0$, L in Fig.6.19(a). The kinks

can be quantified by large change in angle between consecutive bonds $|\Delta\theta_i| = |\theta_{i+1} - \theta_i|$. In semiflexible polymers having persistence lengths significantly longer than bond- length, such kinks between consecutive bonds remain suppressed. In our model, however, the effective bending rigidity arises due to Gaussian core repulsion between backbone monomers. Although it maintains an overall smooth structure, local large angle fluctuations are not completely suppressed, and its amount depends on the crowder- density. At lower crowder density one observes higher bond-angle fluctuations.

Thus it takes extra care to extract the mean turning number from the definition of winding number in Eq.(6.16) within our model. We use a local coarse- graining to determine the relative importance of kinks and turning number in the overall structure. In Fig.6.19(b)-(d) the coarse- grained chain- conformation is shown using coarse- graining over $n_{seg} = 3$, 6, and 12 beads. The N-bead chain is partitioned into N/n_{seg} segments, and the center of mass position of each segment along with bonds connecting them are plotted. Such a coarse-graining allows us to integrate over fluctuations at the shortest length scales. However, as can be seen from Fig.6.19(d), beyond a point, too large a segment size n_{seg} obstructs identification of the true morphology. Thus an intermediate coarse- graining over $n_{seg} = 6$ appears to be best suited for the analysis of turning numbers and significant kinks.

We first present the analysis of average number of kinks $\langle\zeta_N\rangle$ with varying n_{seg} in Fig. 6.20. For the original chain conformation, as well as the coarse- grained conformation, we define a kink by large change in angle between consecutive bonds $|\Delta\theta_i| > \pi/2$. In Fig. 6.20, we show the density dependence of $\langle\zeta_N\rangle$ for various choice of $n_{seg} = 1, 2, \ldots, 6$. The averaging is done over 10^4 independent configurations. With increasing density $\rho_c D^3$, the number of kinks decrease. The measure over all bonds of the original chain shows very large estimate of $\langle\zeta_N\rangle$ at small density, reflecting the large bond angle fluctuations. Increasing n_{seg}, we find that the results saturate as n_{seg} approaches 6. This captures the kinks which survives integration of fluctuations over the shortest length scales, and thus is a characteristic of the model chromosome. At the largest crowder density, $\langle\zeta_N\rangle$ saturates to 2 (Fig. 6.20).

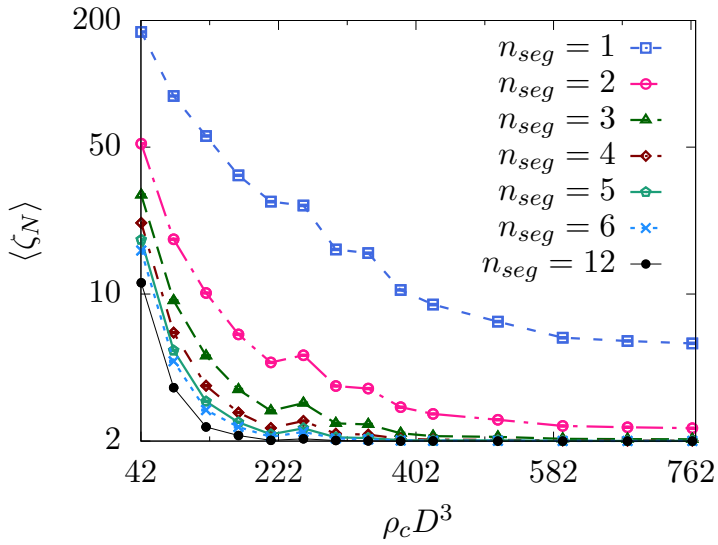


FIGURE 6.20: Average number of kinks is shown as a function of $\rho_c D^3$ in the semi-log plot for different coarse- graining size n_{seg}.

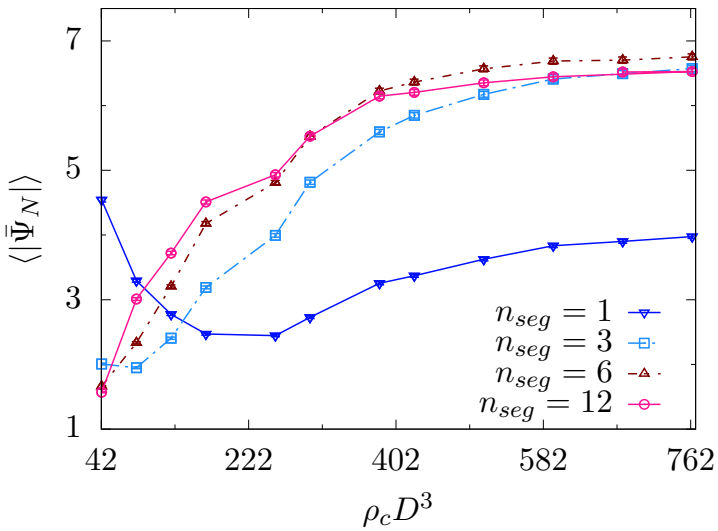


FIGURE 6.21: Mean absolute turning number $\langle |\bar{\Psi}_N| \rangle$ is plotted against $\rho_c D^3$ at different coarse-gaining size.

In calculating the turning number we account for the kinks. A direct application of Eq.(6.16) for winding number in configurations containing kinks suppresses the total turning number, by changing a clockwise turn into anti- clockwise. We use this expression to define a turning number

$$\bar{\Psi}_N = \frac{1}{2\pi} \sum_{i=0}^{N-2} \text{sign} \times \Delta\theta_i, \tag{6.17}$$

where, the sign function switches between ± 1 each time a kink is encountered. Fig.6.21 shows the increase in the turning number amplitude $\langle|\bar{\Psi}_N|\rangle$ as a function of crowder density $\rho_c D^3$ for various coarse- graining segment size n_{seg}. The results stop changing by $n_{seg} = 6$. The turning number increases with density to saturate to $\langle|\bar{\Psi}_N|\rangle \approx 6$ at the highest densities. A comparison of this with the total number of turns n_t obtained from tangent correlation and structure factor, obtained in previous section shows discrepancy. In Fig. 6.17, the plot of n_t shows saturation to $n_t = 8$, whereas here we see saturation of $\langle|\bar{\Psi}_N|\rangle$ to ≈ 6. The discrepancy gets resolved once the saturation value for the number of kinks $\langle\zeta_N\rangle = 2$ is accounted for, leading to $n_t \approx \langle|\bar{\Psi}_N|\rangle + \langle\zeta_N\rangle$. At high density, the number of kinks and turning number together accounts for the turn number n_t obtained from tangent- tangent correlation. However, at low densities, the kink number remains significantly high due to survival of strong bond- angle fluctuations, and thus $\langle|\bar{\Psi}_N|\rangle + \langle\zeta_N\rangle$ remains large with respect to n_t. The above analysis establishes the relative importance of turning number and kinks in the helical morphology of the model chromosome.

6.5 Conclusion and outlook

In this chapter we have studied the impact of changing crowder density on chromosomal organization. We began by considering a feather- boa model of chromosome and non-additive crowders inside a cylindrical confinement, to model the chromosome in bacterial cells like

E. coli. The chromosome and crowder center of mass showed complementary helical organization along the cell length. This is related to out- of- phase density modulations of the chromosome and crowders, a behavior recently observed in the relative organization of ribosomes and chromosome in live E. coli [11]. However, due to the non-additive nature of crowders with short- ranged WCA repulsion against the chromosome and walls, change in crowder density does not produce any significant change in relative organization. Use of an additional Gaussian core repulsion in the crowder- chromosome and crowder- wall repulsion, to incorporate the polymeric nature of protein crowders, led to a stronger density dependence of chromosomal organization. With increasing crowder density, we found a change from predominantly radial to a predominantly longitudinal segregation of the crowders and chromosome.

To investigate this dependence in quantitative detail, we used a computationally more tractable model of the chromosome, replacing the side- loops of the feather-boa model by additional Gaussian core repulsion between the monomers of the backbone. To model bacterial chromosome, we used a circular backbone in this section. As before, we used an additional Gaussian core repulsion in the crowder- polymer and crowder- wall interaction. We found a predominantly radial crowder- chromosome segregation at small crowder- densities, which crossed over to a predominantly longitudinal segregation between crowders and polymer at high densities. The longitudinal segregation compresses the chromosome into a central nucleoid- like structure. We presented a mean field description of the change in chromosome size with crowder density. The radial segregation is associated with a complementary helical organization of the chromosome and crowders. The helical organization was characterized using a tangent- tangent correlation function. An analysis using the winding number, allowed us to quantify the continuous variation of turning number and kinks associated with the change in helical configurations with crowder density. Our predictions are amenable to experimental verifications in live bacteria, and possible soft matter experiments involving polymers and colloidal crowders in confinement.

We have used a minimal polymer-based model of chromosome to capture the main physical processes involved in chromosome organization in bacteria. However, many other physical and chemical specifics can potentially impact their organization. This includes– (1) In bacterial cells, the chromosome has a distribution of loops which reorganize in time during the cellular processes such as replication, gene regulation, segregation etc [3, 25, 26, 97]. The assumption of equispaced side loops of fixed size in the feather- boa model chromosome presents an approximation to this distribution. (2) Bacterial chromosomes are organized into Mb-sized macrodomains, $\sim$ 40-300 kb chromosomal interaction domains (CIDs) and $\sim$ 10 kb supercoil domains or topological domain [25, 26, 144, 193, 220–223]. The relationship between such domains are not clearly established, however they are believed to be organized in a hierarchial fashion. Within each domain, various DNA- loci interact more frequently compared to DNA- loci of other domains. In E.coli, four macrodomains are observed- Ori, Ter, Left and Right, with two less-structured DNA regions. MatP dimers bind to *matS* (13 bp long) sites in the Ter macrodomain forming DNA- loops and compact the domain [224, 225]. Similar class of proteins may organize other macrodomains. Further, Hi-C analysis of C.crescentus and E.coli chromosome revealed the presence of $\sim$ 23 and $\sim$ 31 CIDs respectively [144, 193]. It is often observed that highly expressed genes play an important role in the formation of CID boundaries. Furthermore, the macrodomains and CIDs contain supercoiled domains within it [25, 96]. In *E.coli*, 400 supercoiled domains of $\sim$ 10 kb have been found [26]. Such supercoiled domains are believed to be regulated by the DNA associated proteins and gene expression. This kind of hierarchical organization of chromosome has not been considered in our simple model. (3) In most of the rod shaped bacterial cells, Ori and Ter domains are localized in the specific spatial regions. In C.crescentus, it is achieved by polar anchoring of ParB:*parS* complex with PopZ proteins which localize along the cell poles and provide help in DNA segregation [226–229]. In E.coli SMC protein complex MukBEF is required to maintain the left-ori-right conformation of chromosome [230]. Such specific mechanisms were not considered in our model. (4) Transcription and translation are active processes using energy

from ATP hydrolysis [1]. Consequently, the active force generation can impact the dynamical organization of chromosome. Apart from that, the rate of transcription of highly expressed genes, length of transcript and variation in local concetration of NAPs directly determine the formation of chromosomal domains. Our simple model lack such details. (5) In E.coli cell, 4288 different types of proteins are observed [231]. In addition to that ribosomes, plasmids, different form of RNA, and cytoskeletal filaments (Ftsz, Mreb, Filp etc.) are found which can produce significant variance in the crowder size distribution [37–39, 41–43, 232, 233]. Unlike this, in each of our study, we have assumed a single crowder size. Note that, small crowders can produce deplection effect, however the large crowders undergo phase segregation with respect to the chromosome. (6) The coupled translation and insertion of proteins into the cell membrane by signal recognition particle SRP and Sec translocase lead to transertion which are known to tether some stretch of DNA to the cell membrane [62, 234]. We did not consider such associations. (7) In many bacteria, partitioning complex $parABS$ facilitate the chromosome segregation [226, 235, 236]. The $parS$ sites are the DNA stretches located near the origin of replication, to which protein complex ParB binds forming the nucleoprotein complex ParB:$parS$. ATP-bound proteins ParA form dynamic filaments along the cell, whose ATPase activity is stimulated by ParB [237, 238]. A related study suggested, upon binding to ParB:$parS$, dissociation of ParA molecules happens at the contact site leading to shrinkage of filament away from the complex [239]. Such a filament has been proposed to actively determine the movement of the newly replicated ParB:$parS$ complex towards the opposite cell pole via various mechanisms like Brownian motion, "diffusion-rachet" and "DNA relay" model [240–242]. We did not consider the impact of such active processes in our model. (8) We did not consider the active cytoskeltal filaments (Ftsz, Mreb, ParM etc.) explicitly. In bacteria, cytoskeltal proteins are known to play active role in shaping cell wall, mediate cell division, and partitioning genetic material [42, 43, 243]. In E.coli, tubulin homologe Ftsz is involved in cell division in co-ordination with MinDE system. MinD and MinE oscillate along the long axis of the cell, inhibiting the polymerization of Ftsz ring.

Their localization near the cell poles leads to cytokinetic ring formation at the middle of the cell. The ring attaches to cell membrane and undergo constriction, allowing cell wall synthesis and cell division. Actin homologue Mreb guides the cell wall synthesis that is vital for maintainance of rod shaped morphology of the cell [244]. Similarly, ParM filaments grows in bi-directional fashion causing the segregation of plasmids [44].

Many of these aspects can be added in our basic chromosome model, to study their impact.

References

[1] B. Alberts, D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Essential Cell Biology*, 4th ed. (Garland Science, New York, USA, 2015).

[2] A. Wolffe, *Chromatin: Structure and function*, 3rd ed. (Academic press, London, 2012).

[3] M. Thanbichler, P. H. Viollier, and L. Shapiro, Curr. Opin. Genet. Dev. **15**, 153 (2005).

[4] A. B. Kolomeisky, Biophys. J. **109**, 459 (2015).

[5] G. Fudenberg and L. Mirny, Curr. Opin. Genet. Dev. **22**, 115 (2012).

[6] G. Fudenberg, M. Imakaev, C. Lu, A. Goloborodko, N. Abdennur, and L. A. Mirny, Cell Reports **15**, 2038 (2016).

[7] A. L. Sanborn, S. S. Rao, S. C. Huang, N. C. Durand, M. H. Huntley, A. I. Jewett, I. D. Bochkov, D. Chinnappan, A. Cutkosky, J. Li, K. P. Geeting, A. Gnirke, A. Melnikov, D. McKenna, E. K. Stamenova, E. S. Lander, and E. L. Aiden, Proceedings of the National Academy of Sciences of the United States of America **112**, E6456 (2015).

[8] R. T. Dame, in *Bact. Chromatin*, edited by R. T. Dame (Springer Netherlands, Dordrecht, 2018).

[9] F. Wu, P. Swain, L. Kuijpers, X. Zheng, K. Felter, M. Guurink, J. Solari, S. Jun, T. S. Shimizu, D. Chaudhuri, B. Mulder, and C. Dekker, Curr. Biol. **29**, 2131 (2019).

[10] P. Swain, B. M. Mulder, and D. Chaudhuri, Soft Matter **15**, 2677 (2019).

[11] W. T. Gray, S. K. Govers, Y. Xiang, B. R. Parry, M. Campos, S. Kim, and C. Jacobs-Wagner, Cell **177**, 1632 (2019).

[12] J. K. Fisher, A. Bourniquel, G. Witz, B. Weiner, M. Prentiss, and N. Kleckner, Cell **153**, 882 (2013).

[13] I. A. Berlatzky, A. Rouvinski, and S. Ben-Yehuda, Proc. Natl. Acad. Sci. U. S. A. **105**, 14136 (2008).

[14] W. Wang, G.-W. Li, C. Chen, X. S. Xie, and X. Zhuang, Science **333**, 1445 (2011).

[15] D. Bates and N. Kleckner, Cell **121**, 899 (2005).

[16] 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, Science (80-.). **326**, 289 (2009).

[17] L. A. Mirny, Chromosom. Res. **19**, 37 (2011).

[18] D. Chaudhuri and B. M. Mulder, Physical Review Letters **108**, 268305 (2012).

[19] D. Chaudhuri and B. M. Mulder, in *Bacterial Chromatin*, edited by R. T. Dame (Springer Protocols, Humana Press, New York, 2018), Chap. Molecular Dynamics Simulation of a Feather-Boa Model of a Bacterial Chromosome, p. 403.

[20] J. E. Dawson, *Dynamics of endosomal trafficking*, PhD Thesis, Institut für Theoretische Physik, Fakultät Mathematik und Naturwissenschaften, Technische Universität, Dresden, Germany, 2012.

[21] N. Parker, M. Schneegurt, A.-H. T. Tu, and B. M. Forster, *Microbiology*, 1st ed. (OpenStax, Houston, Texas, USA, 2016).

[22] S. Sazer and H. Schiessel, Traffic **19**, 87 (2018).

[23] D. E. Olins and A. L. Olins, Nature Reviews Molecular Cell Biology **4**, 809 (2003).

[24] J. M. Varley and G. T. Morgan, Chromosoma **67**, 233 (1978).

[25] A. Badrinarayanan, T. B. K. Le, and M. T. Laub, Annual Review of Cell and Developmental Biology **31**, 171 (2015).

[26] L. Postow, C. D. Hardy, J. Arsuaga, and N. R. Cozzarelli, Genes Dev. **18**, 1766 (2004).

[27] T. Vellai and G. Vida, Proceedings of the Royal Society B: Biological Sciences **266**, 1571 (1999).

[28] B. Pierce, *Genetics: A conceptual approach*, 2nd ed. (W. H. Freeman and Company, USA, 2005).

[29] D. N. Wilson and J. H. Doudna Cate, Cold Spring Harbor perspectives in biology **4**, a011536 (2012).

[30] K. Luby-Phelps, in *Microcompartmentation and Phase Separation in Cytoplasm*, Vol. 192 of *International Review of Cytology*, edited by H. Walter, D. E. Brooks, and P. A. Srere (Academic Press, USA, 1999), pp. 189–221.

[31] A. C. Miklos, C. Li, and G. J. Pielak, Methods Enzymol. **466**, 1 (2009).

[32] J. Howard, *Mechanics of Motor Proteins and the Cytoskeleton* (Sinauer Associates, Publishers, NC 27513 USA, 2001).

[33] A. R. Bausch and K. Kroy, Nature Physics **2**, 231 (2006).

[34] M. J. Shelley, Annual Review of Fluid Mechanics **48**, 487 (2016).

[35] K. E. Kasza, A. C. Rowat, J. Liu, T. E. Angelini, C. P. Brangwynne, G. H. Koenderink, and D. A. Weitz, The cell as a material, 2007.

[36] J. Sapp, Microbiology and molecular biology reviews : MMBR **69**, 292 (2005).

[37] R. Ellis, Trends in Biochemical Sciences **26**, 597 (2001).

[38] J. Spitzer and B. Poolman, FEBS Letters **587**, 2094 (2013).

[39] J. T. Trevors, J. D. van Elsas, and A. K. Bej, Current Issues in Molecular Biology **15**, 1 (2013).

[40] B. Parry, I. Surovtsev, M. Cabeen, C. O'Hern, E. Dufresne, and C. Jacobs-Wagner, Cell **156**, 183 (2014).

[41] C. M. Thomas and D. Summers, Bacterial Plasmids, 2008.

[42] E. Bi and J. Lutkenhaus, Nature **354**, 161 (1991).

[43] Z. Gitai, Cell **120**, 577 (2005).

[44] G. Ebersbach and K. Gerdes, Annu. Rev. Genet. **39**, 453 (2005).

[45] W. S. Klug, M. Cummings, C. A. Spencer, M. A. Palladino, and D. Killian, *Essentials of Genetics*, 10th ed. (Pearson, ADDRESS, 2020).

[46] R. T. Dame, M. C. Noom, and G. J. Wuite, Nature **444**, 387 (2006).

[47] T. Cremer and C. Cremer, European journal of histochemistry : EJH **50**, 161 (2006).

[48] T. Cremer and C. Cremer, European journal of histochemistry : EJH **50**, 223 (2006).

[49] J. H. Tjio and A. Levan, Hereditas **42**, 1 (1956).

[50] N. McNeil and T. Ried, Expert reviews in molecular medicine **2000**, 114 (2000).

[51] N. Ganai, S. Sengupta, and G. I. Menon, Nucleic Acids Research **42**, 4145 (2014).

[52] F. Mueller, P. Wach, and J. G. McNally, Biophysical Journal **94**, 3323 (2008).

[53] P.-G. de Gennes, *Scaling concepts in polymer physics* (Cornell University Press, Ithaca, 1979).

[54] G. van den Engh, R. Sachs, and B. Trask, Science **257**, 1410 (1992).

[55] Y. Xiang, I. V. Surovtsev, Y. Chang, S. K. Govers, B. R. Parry, J. Liu, and C. Jacobs-Wagner, bioRxiv (2020).

[56] C. Robinow and E. Kellenberger, Microbiology and Molecular Biology Reviews **58**, 211 (1994).

[57] N. Hadizadeh Yazdi, C. C. Guet, R. C. Johnson, and J. F. Marko, Molecular Microbiology **86**, 1318 (2012).

[58] C. Butan, L. M. Hartnell, A. K. Fenton, D. Bliss, R. E. Sockett, S. Subramaniam, and J. L. S. Milne, Journal of Bacteriology **193**, 1341 (2011).

[59] J. A. Hobot, Journal of Structural Biology **104**, 169 (1990).

[60] E. Kellenberger, A. Ryter, and J. Sechaud, The Journal of biophysical and biochemical cytology **4**, 671 (1958).

[61] D. J. Mason and D. M. Powelson, Journal of bacteriology **71**, 474 (1956).

[62] S. Bakshi, A. Siryaporn, M. Goulian, and J. C. Weisshaar, Molecular microbiology **85**, 21 (2012).

[63] D. P. Bremer H, EcoSal Plus **3**, (2008).

[64] J. O. Ortiz, F. Frster, J. Krner, A. A. Linaroudis, and W. Baumeister, Journal of Structural Biology **156**, 334 (2006).

[65] N. Turgeon, C. Laflamme, J. Ho, and C. Duchaine, Plasmid **60**, 118 (2008).

[66] R. Milo and R. Phillips, *Cell Biology by the Numbers*, 1st ed. (Garland Science, New York, USA, 2015).

[67] R. D. Kornberg, Science **184**, 868 (1974).

[68] K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond, Nature **389**, 251 (1997).

[69] R. Schleif, Annual Review of Biochemistry **61**, 199 (1992), pMID: 1497310.

[70] T. M. Yusufzai, H. Tagami, Y. Nakatani, and G. Felsenfeld, Molecular Cell **13**, 291 (2004).

[71] C. Hou, H. Zhao, K. Tanimoto, and A. Dean, Proceedings of the National Academy of Sciences of the United States of America **105**, 20398 (2008).

[72] Y. Kim, Z. Shi, H. Zhang, I. J. Finkelstein, and H. Yu, Science **366**, 1345 (2019).

[73] E. Kim, J. Kerssemakers, I. A. Shaltiel, C. H. Haering, and C. Dekker, Nature **579**, 438 (2020).

[74] D. Rieder, Z. Trajanoski, and J. G. McNally, Frontiers in Genetics **3**, 1 (2012).

[75] D. A. Jackson, P. R. Cook, and S. B. Patel, Nucleic acids research **12**, 6709 (1984).

[76] B. Tolhuis, R.-J. Palstra, E. Splinter, F. Grosveld, and W. de Laat, Molecular Cell **10**, 1453 (2002).

[77] M. Simonis and W. de Laat, Biochimica et Biophysica Acta (BBA) - Molecular Cell Research **1783**, 2052 (2008), functional Architecture of the Cell Nucleus.

[78] S. Schoenfelder, I. Clay, and P. Fraser, Current Opinion in Genetics and Development **20**, 127 (2010), chromosomes and expression mechanisms.

[79] C. G. Spilianakis and R. A. Flavell, Nature Immunology **5**, 1017 (2004).

[80] S. S. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, and E. L. Aiden, Cell **159**, 1665 (2014).

[81] J. Dekker, K. Rippe, M. Dekker, and N. Kleckner, Science **295**, 1306 (2002).

[82] C. S. Osborne, P. A. Ewels, and A. N. C. Young, Briefings in functional genomics **10**, 11 (2011).

[83] E. de Wit and W. de Laat, Genes & development **26**, 11 (2012).

[84] D. J. Brocken, M. Tark-Dame, and R. T. Dame, Curr. Opin. Syst. Biol. **8**, 137 (2018).

[85] S. C. Dillon and C. J. Dorman, Nat. Rev. Microbiol. **8**, 185 (2010).

[86] H. B. Brandão, P. Paul, A. A. van den Berg, D. Z. Rudner, X. Wang, and L. A. Mirny, Proceedings of the National Academy of Sciences of the United States of America **116**, 20489 (2019).

[87] E. J. Banigan, A. A. van den Berg, H. B. Brandão, J. F. Marko, and L. A. Mirny, Elife **9**, (2020).

[88] T. Ali Azam, A. Iwata, A. Nishimura, S. Ueda, and A. Ishihama, Journal of bacteriology **181**, 6361 (1999).

[89] K. K. Swinger, K. M. Lemberg, Y. Zhang, and P. A. Rice, The EMBO journal **22**, 3749 (2003).

[90] A. I. Prieto, C. Kahramanoglou, R. M. Ali, G. M. Fraser, A. S. N. Seshasayee, and N. M. Luscombe, Nucleic acids research **40**, 3524 (2012).

[91] O. Huisman, M. Faelen, D. Girard, A. Jaff, A. Toussaint, and J. Rouvire-Yaniv, Journal of bacteriology **171**, 37043712 (1989).

[92] S. Kar, R. Edgar, and S. Adhya, Proceedings of the National Academy of Sciences **102**, 16397 (2005).

[93] P. A. Rice, S. wei Yang, K. Mizuuchi, and H. A. Nash, Cell **87**, 1295 (1996).

[94] S. Stella, D. Cascio, and R. C. Johnson, Genes & development **24**, 814 (2010).

[95] H. Delius and A. Worcel, Cold Spring Harb. Symp. Quant. Biol. **38**, 53 (1974).

[96] R. Kavenoff and B. Bowen, Chromosoma **59**, 89 (1976).

[97] N. Trun and J. Marko, ASM News **64**, 276 (1998).

[98] S. Kadauke and G. A. Blobel, Biochimica et biophysica acta **1789**, 17 (2009).

[99] D. S. Johnson, A. Mortazavi, R. M. Myers, and B. Wold, Science **316**, 1497 (2007).

[100] B. Sites and B. Site, 5 (2010).

[101] P. Collas, Molecular Biotechnology **45**, 87 (2010).

[102] M. Doi and S. F. Edwards, *The theory of polymer dynamics* (Clarendon Press, Oxford, 1988).

[103] S. F. Edwards and P. Singh, J. Chem. Soc. Faraday Trans. **75**, 1001 (1979).

[104] M. Barbieri, M. Chotalia, J. Fraser, L.-M. Lavitas, J. Dostie, A. Pombo, and M. Nicodemi, Proc. Natl. Acad. Sci. U. S. A. **109**, 16173 (2012).

[105] G. Le Treut, F. Képès, and H. Orland, Biophys. J. **110**, 51 (2016).

[106] C. A. Brackley, B. Liebchen, D. Michieletto, F. Mouvet, P. R. Cook, and D. Marenduzzo, Biophys. J. **112**, 1085 (2017).

[107] J. J. Hermans and R. Ullman, Physica **18**, 951 (1952).

[108] A. Dhar and D. Chaudhuri, Phys. Rev. Lett. **89**, 65502 (2002).

[109] P. J. Hagerman, Annual Review of Biophysics and Biophysical Chemistry **17**, 265 (1988).

[110] H. G. Garcia, P. Grayson, L. Han, M. Inamdar, J. Kondev, P. C. Nelson, R. Phillips, J. Widom, and P. A. Wiggins, Biopolymers **85**, 115 (2007).

[111] C. Frontali, E. Dore, A. Ferrauto, E. Gratton, A. Bettini, M. R. Pozzan, and E. Valdevit, Biopolymers **18**, 1353 (1979).

[112] Z. Kam, N. Borochov, and H. Eisenberg, Biopolymers **20**, 2671 (1981).

[113] J. Widom, Quarterly Reviews of Biophysics **34**, 269 (2001).

[114] C. Bustamante, J. Marko, E. Siggia, and S. Smith, Science **265**, 1599 (1994).

[115] J. F. Marko and E. D. Siggia, Macromolecules **28**, 8759 (1995).

[116] P. A. Wiggins and P. C. Nelson, Phys. Rev. E **73**, 031906 (2006).

[117] T. Odijk, Macromolecules **12**, 688 (1979).

[118] G. S. Manning, Biopolymers **20**, 1751 (1981).

[119] C. B. Post, Biopolymers **22**, 1087 (1983).

[120] K. Bystricky, P. Heun, L. Gehlen, J. Langowski, and S. M. Gasser, Proceedings of the National Academy of Sciences **101**, 16495 (2004).

[121] C. A. Brackley, S. Taylor, A. Papantonis, P. R. Cook, and D. Marenduzzo, Proc. Natl. Acad. Sci. U. S. A. **110**, E3605 (2013).

[122] M. Rubinstein and R. H. Colby, Polymer physics, 2003.

[123] P. A. Wiggins, R. T. Dame, M. C. Noom, and G. J. Wuite, Biophys. J. **97**, 1997 (2009).

[124] A. Worcel and E. Burgi, J. Mol. Biol. 71: **71**, 127 (1972).

[125] V. Holmes and N. Cozzarelli, Proc. Natl. Acad. Sci. **97**, 1322 (2000).

[126] S. B. Zimmerman, Journal of Structural Biology **156**, 255 (2006).

[127] R. A. Van Der Valk, J. Vreede, F. Crémazy, and R. T. Dame, J. Mol. Microbiol. Biotechnol. **24**, 344 (2014).

[128] R. T. Dame and M. Tark-Dame, Curr. Opin. Cell Biol. **40**, 60 (2016).

[129] M. S. Luijsterburg, M. C. Noom, G. J. Wuite, and R. T. Dame, J. Struct. Biol. **156**, 262 (2006).

[130] R. T. Dame, Mol. Microbiol. **56**, 858 (2005).

[131] E. Peeters, R. P. C. Driessen, F. Werner, and R. T. Dame, Nat. Rev. Microbiol. **13**, 333 (2015).

[132] W. A. Bickmore and B. van Steensel, Cell **152**, 1270 (2013).

[133] K. Nasmyth and C. H. Haering, Annu. Rev. Genet. **43**, 525 (2009).

[134] J. E. Phillips and V. G. Corces, Cell **137**, 1194 (2009).

[135] M. Nicodemi and A. Prisco, Biophys. J. **96**, 2168 (2009).

[136] M. Nicodemi and A. Pombo, Curr. Opin. Cell Biol. **28**, 90 (2014).

[137] J. Johnson, C. A. Brackley, P. R. Cook, and D. Marenduzzo, J. Phys. Condens. Matter **27**, 064119 (2015).

[138] C. A. Brackley, J. Johnson, S. Kelly, P. R. Cook, and D. Marenduzzo, Nucleic Acids Res. **44**, 3503 (2016).

[139] M. Tark-Dame, M. S. Luijsterburg, D. W. Heermann, and R. van Driel, in *Genome Organization and Function in the Nucleus*, edited by K. Rippe (Wiley, Winheim, Germany, 2012), Chap. Understanding genome function: Quantitative modelling of chromatin folding and chromatin-associated processes, pp. 535– 555.

[140] V. F. Scolari and M. Cosentino Lagomarsino, Soft Matter **11**, 1677 (2015).

[141] V. F. Scolari, G. Mercy, R. Koszul, A. Lesne, and J. Mozziconacci, Phys. Rev. Lett. **121**, 057801 (2018).

[142] E. Alipour and J. F. Marko, Nucleic Acids Research **40**, 11202 (2012).

[143] C. A. Brackley, J. Johnson, D. Michieletto, A. N. Morozov, M. Nicodemi, P. R. Cook, and D. Marenduzzo, Physical Review Letters **119**, 1 (2017).

[144] T. B. K. Le, M. V. Imakaev, L. A. Mirny, and M. T. Laub, Science (80-.). **342**, 731 (2013).

[145] M. Marbouty, A. Cournac, J. F. Flot, H. Marie-Nelly, J. Mozziconacci, and R. Koszul, Elife **3**, e03318 (2014).

[146] T. Cremer and C. Cremer, Nat. Rev. Genet. **2**, 292 (2001).

[147] J. Dekker, Science (80-.). **319**, 1793 (2008).

[148] S. Meyer, S. Reverchon, W. Nasser, and G. Muskhelishvili, Curr. Genet. **64**, 555 (2018).

[149] M. Fritsche, S. Li, D. W. Heermann, and P. A. Wiggins, Nucleic Acids Res. **40**, 972 (2012).

[150] P. M. Llopis, A. F. Jackson, O. Sliusarenko, I. Surovtsev, J. Heinritz, T. Emonet, and C. Jacobs-Wagner, Nature **466**, 77 (2010).

[151] O. B. Ptitsyn and Y. E. Eisner, Biofizika **10**, 3 (1965).

[152] Y. E. Eisner, Polym. Sci. U.S.S.R. **11**, 409 (1969).

[153] P. G. de Gennes, J. Phys. Lettres **36**, 55 (1975).

[154] I. M. Lifshitz, A. Y. Grosberg, and A. R. Khokhlov, Rev. Mod. Phys. **50**, 683 (1978).

[155] I. M. Lifshitz, A. Y. Grosberg, and A. R. Khokhlov, Zh. Eksp. Teor. Fiz. **71**, 1634 (1976).

[156] Y. A. Budkov, A. L. Kolesnikov, N. Georgi, and M. G. Kiselev, Journal of Chemical Physics **141**, (2014).

[157] J. D. Weeks, D. Chandler, and H. C. Andersen, J. Chem. Phys. **54**, 5237 (1971).

[158] G. Grest and K. Kremer, Phys. Rev. A **33**, 3628 (1986).

[159] J. D. Watson, T. A. Baker, S. P. Bell, A. Gann, M. Levine, and R. Losick, *Molecular Biology of the Gene*, 7th ed. (Benjamin-Cummings, Menlo Park, CA., 2003).

[160] D. Frenkel and B. Smit, *Understanding molecular simulation: from algorithms to applications* (Academic press, NY, 2002).

[161] H. Limbach, A. Arnold, B. Mann, and C. Holm, Comput. Phys. Commun. **174**, 704 (2006).

[162] B. Duenweg, M. Stevens, and K. Kremer, *Monte Carlo and Molecular Dynamics Simulations in Polymer Science* (Oxford University Press, New York, U.S.A, 1995).

[163] W. Humphrey, A. Dalke, and K. Schulten, Journal of Molecular Graphics **14**, 33 (1996).

[164] M. P. Allen and D. J. Tildesley, *Computer simulation of liquids* (Oxford University Press, Oxford, 1989).

[165] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, 2012).

[166] A. Mashaghi, R. J. Van Wijk, and S. J. Tans, Structure **22**, 1227 (2014).

[167] A. Y. Grosberg, S. Nechaev, and E. Shakhnovich, J. Phys. **49**, 2095 (1988).

[168] J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J. S. Liu, and B. Ren, Nature **485**, 376 (2012).

[169] M. Bohn, D. W. Heermann, and R. van Driel, Phys. Rev. E **76**, 051805 (2007).

[170] J. Mateos-Langerak, M. 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, Proc. Natl. Acad. Sci. **106**, 3812 (2009).

[171] T. Mora and W. Bialek, Journal of Statistical Physics **144**, 268 (2011).

[172] A. Routh, S. Sandin, and D. Rhodes, Proc. Natl. Acad. Sci. **105**, 8872 (2008).

[173] G. L. Lukacs, P. Haggie, O. Seksek, D. Lechardeur, N. Freedman, and A. S. Verkman, J. Biol. Chem. **275**, 1625 (2000).

[174] Y. Tseng, J. S. H. Lee, T. P. Kole, I. Jiang, and D. Wirtz, J. Cell Sci. **117**, 2159 (2004).

[175] L. A. Mirny, M. Imakaev, and N. Abdennur, Current Opinion in Cell Biology **58**, 142 (2019).

[176] S. S. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, and E. L. Aiden, Cell **159**, 1665 (2014).

[177] A. Goloborodko, M. V. Imakaev, J. F. Marko, and L. Mirny, Elife **5**, 1 (2016).

[178] M. Renda, I. Baglivo, B. Burgess-Beusse, S. Esposito, R. Fattorusso, G. Felsenfeld, and P. V. Pedone, Journal of Biological Chemistry **282**, 33336 (2007).

[179] A. Pombo, D. A. Jackson, M. Hollinshead, Z. Wang, R. G. Roeder, and P. R. Cook, The EMBO journal **18**, 2241 (1999).

[180] A. G. Larson, D. Elnatan, M. M. Keenen, M. J. Trnka, J. B. Johnston, A. L. Burlingame, D. A. Agard, S. Redding, and G. J. Narlikar, Nature **547**, 236 (2017).

[181] P. de Gennes, J. Phys. Lettres **39**, 299 (1978).

[182] T. M. Birshtein and V. A. Pryamitsyn, Macromolecules **24**, 1554 (1991).

[183] B. Chu, Q. Ying, and A. Y. Grosberg, Macromolecules **28**, 180 (1995).

[184] C. Wu and X. Wang, Physical Review Letters **80**, 4092 (1998).

[185] J. M. Polson and M. J. Zuckermann, Journal of Chemical Physics **113**, 1283 (2000).

[186] J. M. Polson and N. E. Moore, Journal of Chemical Physics **122**, (2005).

[187] A. Grosberg and A. Khokhlov, *Statistical physics of macromolecules SE - AIP series in polymers and complex materials*, 1st ed. (AIP-Press, New York, USA, 1994).

[188] P. Therizolsa, T. Duong, B. Dujon, C. Zimmer, and E. Fabre, Proceedings of the National Academy of Sciences of the United States of America **107**, 2025 (2010).

[189] M. Emanuel, N. H. Radja, A. Henriksson, and H. Schiessel, Physical Biology **6**, 025008 (2009).

[190] C. Münkel, R. Eils, S. Dietzel, D. Zink, C. Mehring, G. Wedemann, T. Cremer, and J. Langowski, Journal of Molecular Biology **285**, 1053 (1999).

[191] H. Yokota, G. van den Engh, J. E. Hearst, R. K. Sachs, and B. J. Trask, The Journal of cell biology **130**, 1239 (1995).

[192] A. Goloborodko, J. Marko, and L. Mirny, Biophys. J. **110**, 2162 (2016).

[193] V. S. Lioy, A. Cournac, M. Marbouty, S. Duigou, J. Mozziconacci, O. Espéli, F. Boccard, and R. Koszul, Cell **172**, 771 (2018).

[194] X. Wang, P. M. Llopis, and D. Z. Rudner, Nature Reviews Genetics **14**, 191 (2013).

[195] M. S. Luijsterburg, M. F. White, R. Van Driel, and R. Th. Dame, Critical Reviews in Biochemistry and Molecular Biology **43**, 393 (2008).

[196] M. C. Pereira, C. A. Brackley, J. S. Lintuvuori, D. Marenduzzo, and E. Orlandini, Journal of Chemical Physics **147**, (2017).

[197] A. Kumar and D. Chaudhuri, Journal of Physics: Condensed Matter **31**, 354001 (2019).

[198] D. Song and J. J. Loparo, Trends Genet. **31**, 164 (2015).

[199] Y. Kim, Z. Shi, H. Zhang, I. J. Finkelstein, and H. Yu, Science **366**, 1345 (2019).

[200] R. De Vries, Polymer Science - Series C **54**, 30 (2012).

[201] T. N. Shendruk, M. Bertrand, H. W. De Haan, J. L. Harden, and G. W. Slater, Biophys. J. **108**, 810 (2015).

[202] S. Jun, Biophysical Journal **108**, 785 (2015).

[203] T. Odijk, Biophys. Chem. **73**, 23 (1998).

[204] S. Cunha, C. L. Woldringh, and T. Odijk, Journal of structural biology **136**, 53 (2001).

[205] D. Marenduzzo, C. Micheletti, and P. R. Cook, Biophys. J. **90**, 3712 (2006).

[206] C. Jeon, Y. Jung, and B. Y. Ha, Soft Matter **12**, 9436 (2016).

[207] C. Jeon, Y. Jung, and B. Y. Ha, Scientific Reports **7**, 1 (2017).

[208] H. Kang, P. A. Pincus, C. Hyeon, and D. Thirumalai, Phys. Rev. Lett. **114**, 68303 (2015).

[209] J. Mondal, B. P. Bratton, Y. Li, A. Yethiraj, and J. C. Weisshaar, Biophys. J. **100**, 2605 (2011).

[210] A. Maritan, C. Micheletti, and A. Trovato, Nature **406**, 287 (2000).

[211] D. Chaudhuri and B. M. Mulder, Phys. Rev. E **83**, 031803 (2011).

[212] R. Wang, P. Virnau, and K. Binder, Macromolecular Theory and Simulations **19**, 258 (2010).

[213] H.-P. Hsu, W. Paul, and K. Binder, Macromolecules **43**, 3094 (2010).

[214] P. E. Theodorakis, H. P. Hsu, W. Paul, and K. Binder, Journal of Chemical Physics **135**, 164903 (2011).

[215] P. G. Bolhuis, A. A. Louis, J. P. Hansen, E. J. Meijer, and I. Introduction, Chemical Physics **114**, 4296 (2001).

[216] K. Kremer and G. S. Grest, J. Phys. Condens. Matter **2**, SA295 (1990).

[217] S. Asakura and F. Oosawa, Journal of Polymer Science **33**, 183 (1958).

[218] A. Narros, A. J. Moreno, and C. N. Likos, Macromolecules **46**, 9437 (2013).

[219] S. G. Krantz, *Handbook of Complex Variables* (Birkhäuser, Boston,MA, 1999).

[220] H. Niki, Y. Yamaichi, and S. Hiraga, Genes & development **14**, 212 (2000).

[221] M. Valens, S. Penaud, M. Rossignol, F. Cornet, and F. Boccard, The EMBO journal **23**, 4330 (2004).

[222] O. Espeli, R. Mercier, and F. Boccard, Molecular Microbiology **68**, 1418 (2008).

[223] R. Sinden and D. E. Pettijohn, Proc. Natl. Acad. Sci. 78: **78**, 224 (1981).

[224] R. Mercier, M.-A. Petit, S. Schbath, S. Robin, M. El Karoui, F. Boccard, and O. Espéli, Cell **135**, 475 (2008).

[225] P. Dupaigne, N. Tonthat, O. Espéli, T. Whitfill, F. Boccard, and M. Schumacher, Molecular Cell **48**, 560 (2012).

[226] D. A. Mohl, J. Easter Jr, and J. W. Gober, Molecular Microbiology **42**, 741 (2001).

[227] E. Toro, S.-H. Hong, H. H. McAdams, and L. Shapiro, Proceedings of the National Academy of Sciences of the United States of America **105**, 15435 (2008).

[228] G. R. Bowman, L. R. Comolli, J. Zhu, M. Eckart, M. Koenig, K. H. Downing, W. E. Moerner, T. Earnest, and L. Shapiro, Cell **134**, 945 (2008).

[229] G. Ebersbach, A. Briegel, G. J. Jensen, and C. Jacobs-Wagner, Cell **134**, 956 (2008).

[230] O. Danilova, R. Reyes-Lamothe, M. Pinskaya, D. Sherratt, and C. Possoz, Molecular microbiology **65**, 1485 (2007).

[231] F. R. Blattner, G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao, Science **277**, 1453 (1997).

[232] S. B. Zimmerman and S. O. Trach, Journal of Molecular Biology **222**, 599 (1991).

[233] A. Minton, Biophysical Journal **63**, 1090 (1992).

[234] E. A. Libby, M. Roggiani, and M. Goulian, Proc. Natl. Acad. Sci. **109**, 7445 (2012).

[235] M. A. Fogel and M. K. Waldor, Genes and Development **20**, 3269 (2006).

[236] D. C. H. Lin and A. D. Grossman, Cell **92**, 675 (1998).

[237] J. Easter and J. W. Gober, Molecular Cell **10**, 427 (2002).

[238] A. C. Leonard and J. E. Grimwade, Molecular microbiology **55**, 978 (2005).

[239] J. L. Ptacin, S. F. Lee, E. C. Garner, E. Toro, M. Eckart, L. R. Comolli, W. E. Moerner, and L. Shapiro, Nature Cell Biology **12**, 791 (2010).

[240] L. C. Hwang, A. G. Vecchiarelli, Y.-W. Han, M. Mizuuchi, Y. Harada, B. E. Funnell, and K. Mizuuchi, The EMBO journal **32**, 1238 (2013).

[241] A. G. Vecchiarelli, K. C. Neuman, and K. Mizuuchi, Proceedings of the National Academy of Sciences of the United States of America **111**, 4880 (2014).

[242] H. C. Lim, I. V. Surovtsev, B. G. Beltran, F. Huang, J. Bewersdorf, and C. Jacobs-Wagner, eLife **3**, e02758 (2014).

[243] Y.-L. Shih and L. Rothfield, Microbiology and molecular biology reviews : MMBR **70**, 729 (2006).

[244] S. van Teeffelen, S. Wang, L. Furchtgott, K. C. Huang, N. S. Wingreen, J. W. Shaevitz, and Z. Gitai, Proceedings of the National Academy of Sciences **108**, 15822 (2011).

[245] J.-K. Ryu, C. Bouchoux, H. W. Liu, E. Kim, M. Minamino, R. de Groot, A. Katan, A. Bonato, D. Marenduzzo, D. Michieletto, F. Uhlmann, and C. Dekker, bioRxiv (2020).

[246] D. Amiad-Pavlov, D. Lorber, G. Bajpai, S. Safran, and T. Volk, bioRxiv (2020).

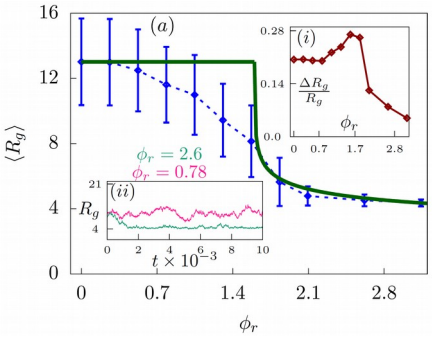
Thesis Highlight

Name of the Student: Amit Kumar
Name of the CI/OCC: Institute of Physics, Bhubaneswar
Enrolment No.: PHYS07201404012
Thesis Title: Polymeric models for chromosome organization: Impact of cross-linkers, crowders and confinement
Discipline: Physical Sciences **Sub-Area of Discipline: Biophysics**
Date of viva voce: 27/11/2020

The highly folded structure of chromosome consists of DNA and associated proteins. At the small length scales, DNA associates with Histone proteins to form chromatin. In the first part of the thesis, using molecular dynamics simulation, we investigated the impact of changing density of multi- valent binder proteins on chromatin morphology. We model the chromatin as a self- avoiding chain, and assume an attractive potential between the binders (cross-linkers) and monomers. The increasing binder density leads to chromatin- folding, a continuous coil- globule transition mediated by clustering of binders. The prediction from a mean field model agrees reasonably well with the simulation results for the folding transition. Along the transition, the chromatin undergoes local morphological changes which we characterize in terms of contact probabilities, loops of various orders, and zippering.

In the second part of this thesis, we focus on the impact of confinement and cytosolic crowding on the morphology of the chromosome in rod- shaped bacteria as a whole. Motivated by the observation of large loops in the chromosome, we consider the feather-boa model of chromosome, consisting of polymeric loops attached to a backbone chain. Molecular dynamics simulations of the feather-boa polymer subjected to cylindrical confinement describe the emergent helical organization of the chromosome observed in experiments. A further coarse-grained model is used to probe the impact of cytosolic crowding more systematically.

Small crowders remain homogeneously distributed while the chromosome remains open. In the range of intermediate crowder sizes, a transverse segregation of the chromosome and crowders leads to a complimentary helical organization of the two. For even larger size, crowders are completely expelled from the central region, leading to a longitudinal segregation along the long axis of cylinder.



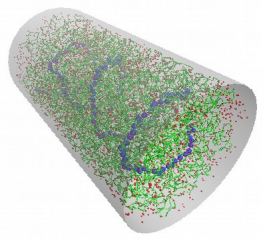


Figure: Top shows decrease in radius of gyration with binder density. Bottom shows helical organization of feather boa chain in cylindrical confinement.