SISSA
Scuola
Internazionale
Superiore di
Studi Avanzati

Scuola Internazionale Superiore di Studi Avanzati

Doctoral Thesis

Statics and dynamics of physical linking: from confined polymers to chromatin filaments

Author:
Giulia Amici

Supervisor:
Prof. Cristian Micheletti
Co-supervisor:
Prof. Angelo Rosa

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

PhD course in Physics and Chemistry of Biological Systems
Molecular and Statistical Biophysics Group

Academic year 2019-2020

Via Bonomea, 265
34136 Trieste – Italy
T +39 0403787111
E info@sissa.it
sissa.it

Alla mia famiglia.

Contents

Part I

Linked rings under confinement

Introduction

During last decades, the improvement of nanotechnologies led to the possibility of performing experiments at a single molecule level in a variety of conditions. This allowed us to investigate biomolecules on a scale that was not previously accessible to experiments, while studied in physical theoretical models for years.

Indeed, traditional experimental techniques were based on bulk measurements, where results were averaged over molecules and time [1]. Single-molecule experiments (SME) instead combine manipulation techniques (such as AFM, LOTs, MTs and BFP [2–8]) and visualization techniques (fluorescence techniques, as SMF [9–11]) in order to detect and follow single biomolecules in real time. In addition, the development in microscale manufacturing methods enabled also the study of single biomolecules in narrow confinement conditions [12], a typical scenario in biology that has a central interest in present studies. Indeed, using optical tweezers it is possible to trap DNA-filaments into channels and slits with nanometric dimensions [13], probing then the physical properties with high accuracy.

These progresses gave the possibility to procede in parallel with experiments and physical modeling, validating existing theories and also elaborating new models. A good example of this fruitful combination are the single-molecule pulling experiments, that measure force as a function of the molecular extension [14]. These experiments demonstrated that the elastic response of double-stranded DNA was described by the worm-like chain model, introduced in polymer theory in 1949 [15].

This is just one of the possible examples of the boost that SME gave to the understanding of the physics of single chains. Over the years, the static, dynamical and mechanical properties of single linear chains have been well characterized, in various conditions such as in response to stimuli like electrical fields, convergent fluxes, and in both bulk and confinement [16–19].

Understanding the single chain case is obviously the first step to achieve in order to

describe more complicated systems, made of more than one chain constrained in narrow volumes and mutually entangled. This condition is typical of many biological systems, such as catenanes inside bacteria or chromosomes into eukaryotic nuclei [20–22]. Here the interplay between confinement and entanglement is a key feature in determining structural and functional properties [23–25].

Given the biological relevance and the physical interest in such systems, polymer physicists were motivated in studying entangled chains under confinement, starting from the simplest case of single knotted chains, a step further the single linear chain. Intra-chain entanglement has been explored both theoretically and experimentally, firstly revealing the physical properties of knotted chains in bulk and secondly showing that the different properties of knotted and unknotted polymers can be significantly enhanced by spatial confinement [26–37]. These findings have explained the role of knots in many biological processes, such as transcription and replication [24, 38–42] or protein folding [43–46].

Even more representative of biological systems is the case of inter-chain entanglement (linking) in confinement, with more chains entangled together. This is the natural step further after the characterization of single knotted chains. Nowadays using SME also systems made of two or more entangled chains can be experimentally studied [47], and this is an unprecedented chance to explore the physical laws underling very complicated biological systems. Think for example to the case of chromosomes inside eukaryotic nuclei and to the importance of having a good knowledge of their properties. But, despite the biological relevance and the experimental feasibility, and in contrast with the intra-chain entanglement case, the case of linked chains under confinement remained unexplored.

We then decided to address this study, starting from the simplest case of two linked rings under channel confinement, asking ourselves if our findings could be framed in existing theories about single molecules, or in a more general picture. We provided a description of the metric and topological observables of the system, investigating both the equilibrium and the kinetic aspects, using molecular dynamics simulations.

More in details, this first part of the thesis is organized as follows.

In chapter 2 I provide a primer on links, which sets a reference for concepts and methods used in subsequent chapters. In particular, I give the definition of links and describe their topological properties. Afterwards, I introduce the computational technique used to

detect and localize the physically linked portion of a link, presenting also the first application of the method to the case of a pair of linked rings in bulk. This is indeed a reference for the interpretation of my results for linked rings in confinement.

In Chapter 3 I first overview a set of key theoretical results for single linear chains under channel confinement, preliminary for the interpretation of the results for confined linked rings. Then, I illustrate the study of the equilibrium metric and topological properties of two linked rings under channel confinement. I describe the model and the methods and present the results. The most surprising results regard the properties of the length of the physically linked portion of the system, studied upon entering the weak confinement regime. The linked portion length is independent on chain length, while, at fixed chain length, it increases monotonically with channel diameter. Moreover, even when confinement is strong enough to pull apart and segregate unlinked rings, the linked portion stays much larger than in the highly stretched limit.

These findings were totally unexpected and are in contrast with the intra-chain entanglement case. They shed light on the physics of confined linked chains for the first time, revealing the interesting consequences of the balance between the topological linking constraint and the confinement constraint. I finally frame them together in a theoretical interpretation and give an idea of possible biological applications.

In Chapter 4 I extend the study of the previous chapter to the case of slit confinement and to the characterization of the kinetic evolution of the system. I start presenting some concepts of polymer kinetics used in the chapter. Then, I illustrate the methods and the obtained results. At the state of the art, little is known on the effect that mutual entanglement has on the kinetics of chains under confinement and we discovered that the kinetic properties of the linked portion are really sensitive to confinement dimensionality. The most impressive results regards the diffusion of the linked portion along the rings' contour. At the maximum considered confinement, the renewal time of the topological constraint along the chain countour is nearly an order of magnitude larger for channels than for slits, indicating a stronger effect of the channel geometry on topological kinetic properties. I compare also the renewal time with the system unlinking time, measured once a cut of the chains is made. I show that the mechanism of unlinking is qualitatively different from the one governing the

stochastic contour displacement of the linked portion.

The presented results give a first significative contribution to the characterization of the kinetics of linked polymers in confinement and can have a significative biological relevance, whose implications are here suggested.

The above results are collected in the following articles, on which this thesis is based:

- G. Amici, M. Caraglio, E. Orlandini, C. Micheletti, *Topologically Linked Chains in Confinement*, ACS Macro Lett. 8, 442 (2019).
- G. Amici, M. Caraglio, E. Orlandini, C. Micheletti, *Effect of Dimensionality and degree of Confinement on the Dynamics of Linked rings*, in preparation.

Links

2.1 Definition of link

In everyday life we experience links: for example when we use the chain of a bicycle, or we wear a bracelet, that are made of small rings linked together (figure 2.1). This familiar images help us in the definition of links, some examples of which are given in figure 2.2. In figure we see quasi-planar geometrical representations of links, that actually are objects in 3D represented through their 2D *projections*. Looking at the figure, we can thus intuitively define a link as a set of closed curves entangled together. Each closed curve is a *component* of the link and the number of components is called *multiplicity*, m [48]. For example in figure 2.2 the so-called *unlink* or *trivial* link, the *Hopf* link and the *Star of David* have 2 components.

Hopf, Star of David are traditional names used for links. A rigorous notation is the one introduced by Rolfsen [49], in which links are grouped according to their multiplicity and their *crossing number*, C, the minimal number of crossings over all possible projections and geometrical representations of the link. The complete notation is C_s^m, where m is the multiplicity and the subscript s is an enumerative index distinguishing between links with

FIGURE 2.1: Example of links in everyday life.

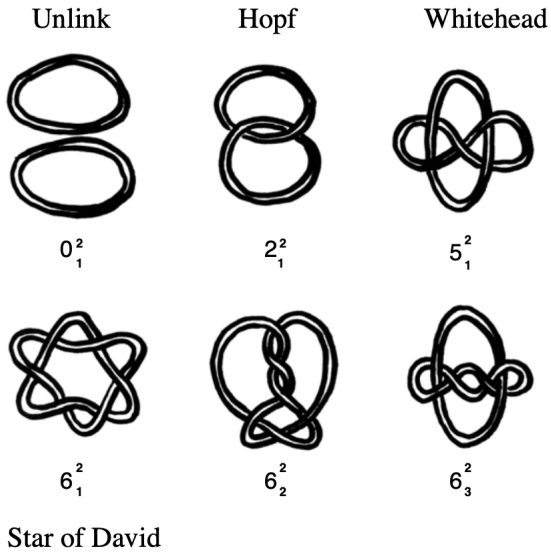


FIGURE 2.2: Examples of links and associated notation.

same C and m but different topology. According to the Rolfsen notation, for example the *unlink* is denoted as 0_1^2, while the *Hopf link* is denoted as 2_1^2.

Note that the links represented in figure 2.2 are intuitively perceived as topologically distinct. But in general, to distinguish between link topologies could be not an easy task: for example looking at snapshots of linked molecules from an experiment, or looking at the conformations of linked chains obtained in a simulation. A practical example is provided in figure 2.3c: could you say, at a first sight, that the links in figure have the same topology?

We are facing the issue to associate a geometrical representation of a link to its topological state. In order to do this, in principle we could rearrange the link geometry until we reach its simplest quasi-planar representation (the one with the minimum number of crossings) and then classify it looking at the tabulated links. But this method is hardly feasible in practice. And in any case, we need to find a tool that is not associated with the visual comparison with tabulated links, especially in cases in which one has to classified a lot of conformations (both in experiments and in numerical studies).

To tackle the problem, mathematicians have identified the so-called *topological invariants* of a link. They are quantities that are independent on the link geometric realization, i.e. for which there is no succession of singular geometrical transformations that turn the link into another with different topology. This set of transformations is denoted as *Reidemeister moves*: the only possible succession of planar moves that change the geometric realization

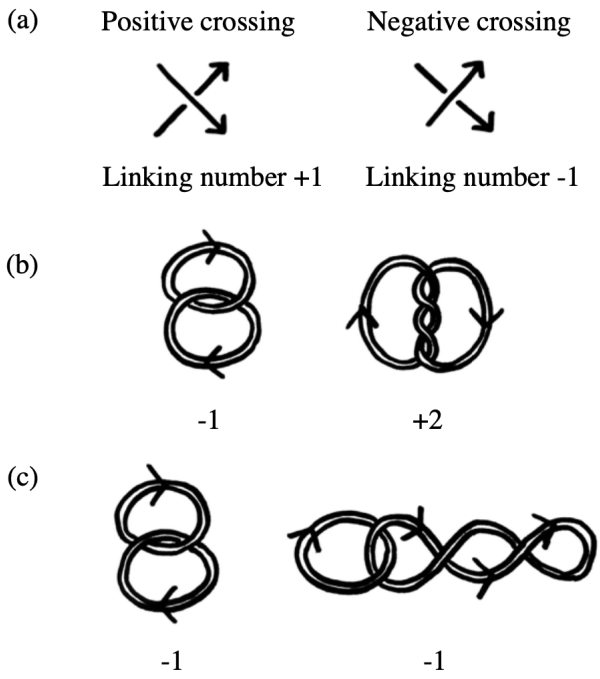


FIGURE 2.3: Computing the linking number.

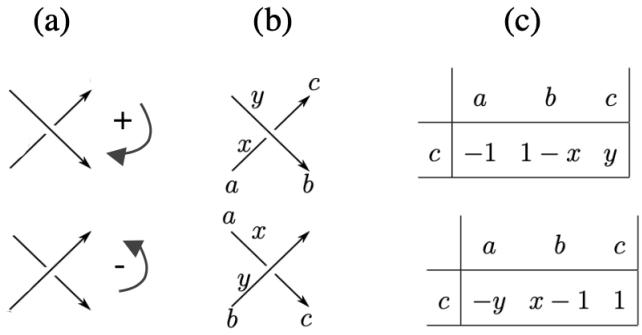
without altering the topology of the link [48].

2.2 Topological invariants: linking number and polynomials

An intuitive topological invariant is the *linking number* [48]. The linking number is an invariant of the oriented link, that is, it is defined once the *orientation* on the components is chosen. This means that there is a preferred direction to travel along each path. Orientation is indicated with an arrow and oriented diagrams have two types of crossings, positive and negative, as shown in figure 2.3a.

In order to calculate the linking number one proceeds as follows: assign an orientation to the projection of the link and then count $+1$ or -1 for the two types of possible crossings (figure 2.3a). The linking number is given by the sum of the +1s and -1s over all crossings divided by two. In figure 2.3 there are some examples: in panel b there are two links with different topology and different linking number; in panel c there are instead different geometrical realizations of the same link, whose linking number remains invariant.

But there are some problematic issues: some link topologies share the same linking number. For example the so-called *Whitehead link* has linking number 0 as the unlink. Thus, the



	a	b	c
c	-1	$1-x$	y

	a	b	c
c	$-y$	$x-1$	1

FIGURE 2.4: Alexander rule. The matrix has one row for each crossing, and one column for each arc.

linking number is useful and intuitive, but its degeneracy for the simplest link types poses the necessity for more powerful invariants, such as the *multivariable Alexander polynomial* [50].

2.2.1 Multivariable Alexander polynomial

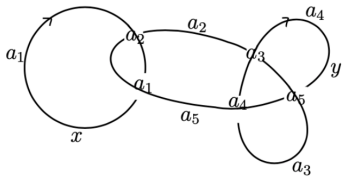
The multivariable Alexander polynomial (MVA) is associated to a link, with the polynomial order equal to the number of components of the link. In this thesis I use an algorithm that calculates the MVA in section 2.3.2.

The MVA is calculated from the *Alexander matrix* of a link projection, built according to the following points [51] (illustrated in figure 2.4):

- Assign an orientation to the link projection and establish the sign of each crossing, following for instance the right-hand rule (panel a);
- Label all the arcs in the projection with a progressive numbering index n, in such a way that the n-arc starts at the n-crossing. For example in panel b there are two components x and y, one crossing and three arcs: $n = a, b, c$;
- The Alexander matrix has dimension given by the *number of crossings x number of arcs* and its entries are calculated with the rule in figure 2.4c.

In figure 2.5 there is an example of matrix calculation.

If the Alexander matrix is denoted as M, the MVA is calculating then considering the determinant of a minor of the matrix M_i^j, with the $i-th$ row and $j-th$ column of M removed:



$$M = \begin{array}{c|ccccc} & a_1 & a_2 & a_3 & a_4 & a_5 \\ \hline a_1 & y-1 & 0 & 0 & 0 & 1-x \\ a_2 & 1-y & x & 0 & 0 & -1 \\ a_3 & -y & 0 & 1 & y-1 & 0 \\ a_4 & 0 & 0 & -y & 1 & y-1 \\ a_5 & 0 & 0 & y-1 & -y & 1 \end{array}$$

FIGURE 2.5: Figure from [51]. Example of Alexander matrix.

$$MVA = (-1)^{i+j} \frac{\det(M_j^i)}{w_i(x_j - 1)} \prod_k x_k^{\frac{rot(k)-\mu(k)}{2}} \tag{2.1}$$

x_j is the variable associated to the component of the link at which the arc of the $j-th$ column belongs;

$w_j = \sum_{l=1}^{m} \pm x_l$;

$\mu(k)$ is the numbers of times the $k-th$ component is the over strand in a crossing;

$rot(k)$ is the rotation number of the $k-th$ component (i.e how many times the tangent vector to the $k-th$ component rotates in the counter clockwise direction as we travel around it once).

2.3 Physical links

Until now I have described the conventional topological features of links, in this section I focus instead on the physical properties of linked chains.

The physical properties of a molecule are significantly influenced by the presence of entanglement. For example it is important to classify the topology of a linked molecule, but also to quantify the localization of its entangled region, because this influences the physical response of the molecule to external conditions and stimuli.

In this section I refer in particular to the Hopf link, since in chapters 3 and 4 I illustrate precisely the case of two chains linked in the Hopf topology.

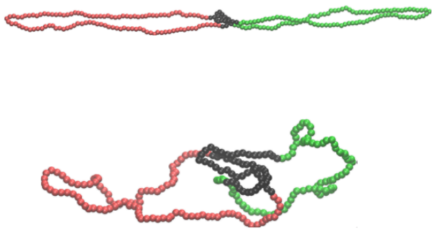


FIGURE 2.6: Links with same topology but different degree of localization of the entangled region.

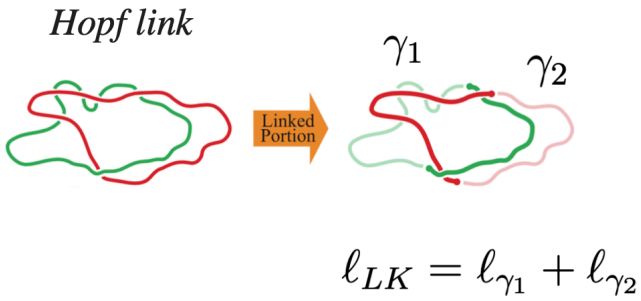


FIGURE 2.7: Linked portion for the Hopf Link. The linked portion is the shortest portion that, upon closure, has the same topology of the original link; its length is given by the sum of the constitutive sub-arcs, γ_1 and γ_2.

2.3.1 Linked portion

In figure 2.6 there is an example of two links of the same topology but with different degree of entanglement. We would like an unsupervised method that algorithmically can establish the degree of entanglement of a link.

The degree of entanglement can be quantified identifying the so-called *linked portion* of the system, introduced in ref. [52]. The linked portion is a topological feature peculiar of linked chains, and can be defined as the portion of the component chains where the physical entanglement resides. An example is given in figure 2.7: here from the two component rings, the linked region is identified by two smaller sub-arcs, labeled as γ_1 and γ_2.

According to definition given in [52], the requirement for the choice of the sub-arc pair γ_1 and γ_2 is that they have to form the pair with the smallest total arc-length that has the same topology of the entire link. And here a problem arises: γ_1 and γ_2 are a system of open curves, but mathematically the topological entanglement of two curves is defined only if they are both closed. By definition they are actually unlinked and their topological state cannot

be established using topological invariants. Thus, there are no mathematical instruments for quantifying the physical entanglement among open chains.

The problem was solved in [52] introducing the concept of *physical link*: a physical link is made of open entangled curves and its topological state is the one of the mathematical link obtained closing its component chains in a proper way. The concept of physical link can be used then as a tool for the definition of the linked portion of a topological link (see next section).

2.3.2 Identification of the linked portion

In this section I illustrate the algorithm for the location of the linked portion of a link, developed in [52].

Algorithm

Given a link with components Γ_1 and Γ_2 we want to identify its linked portion. To do that, in principle one should start to consider the whole ensemble of sub-arc pair γ_1 and γ_2 (see figure 2.7). Since this is very computationally demanding, one can adopt a stochastic top-down search scheme based on a bisection method (illustrated in figure 2.8):

- Start from the curves (Γ_1, Γ_2) with topology τ and length $L(\Gamma_1, \Gamma_2)$;
- Generate the set of sub-arcs $\{\gamma_{12}\} = \{(\gamma_1, \gamma_2)\}$ with length $\lambda'(\gamma_1, \gamma_2)$. At the beginning $\lambda' = 2^k$ with $k = max_j \left\{j | 2^j < L(\Gamma_1, \Gamma_2)\right\}$, while in next iterations $k \to k-1$ and $\lambda' \to \lambda' \pm 2^k$ (bisection rule);
- At each step, for every pair (γ_1, γ_2) of the set $\{\gamma_{12}\}$ one should verify if it has topology $\tau' = \tau$, closing properly the two curves (see below in *Closure procedure*), calculating the MVA and then keeping a pair only if its length is smaller than the previous ones. This is very expensive: to save computational cost not all the pairs are analyzed and a self-learning procedure is considered. Given a pair (γ_1, γ_2): if $\tau' = \tau$, the pair is chosen as temporary linked portion and one moves to the next step (without analyzing the rest of the set) setting $k \to k-1$ and $\lambda' \to \lambda' - 2^k$ (i.e. the sub-arc length is reduced); if not, the pair is assigned to a new set $\{\overline{\gamma_{12}}\}$ that contains all the pair of a certain length

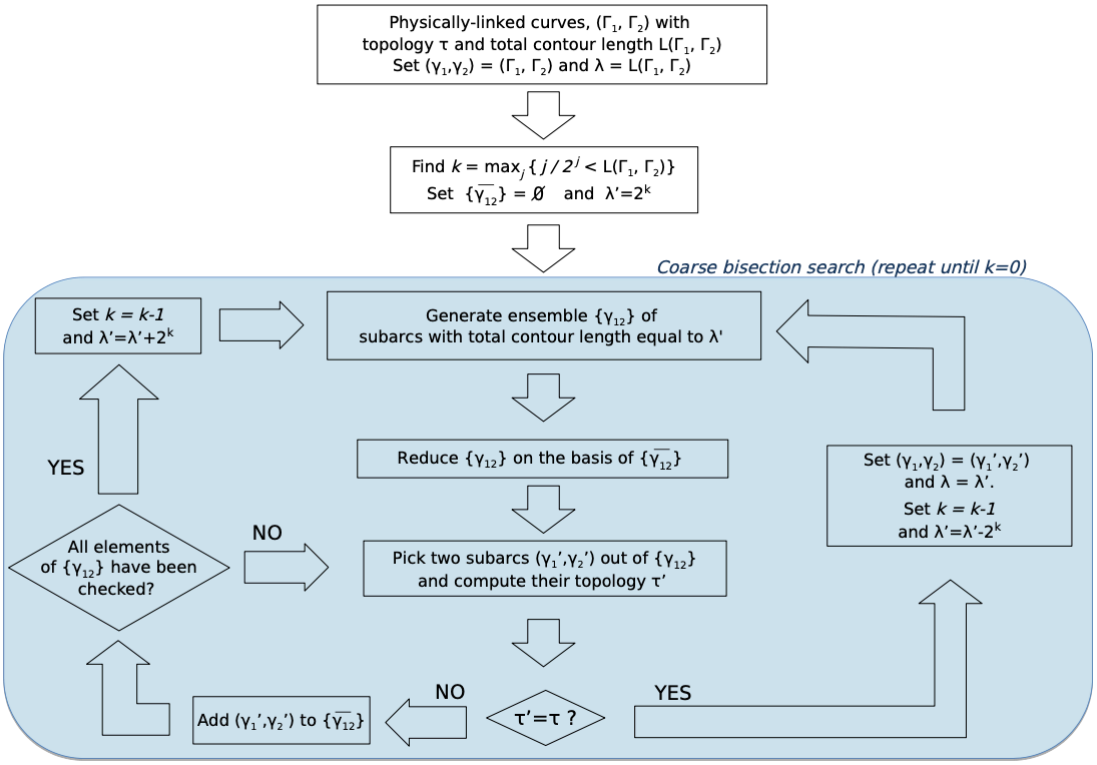


FIGURE 2.8: Linked portion search algorithm.

with $\tau' \neq \tau$; if no pairs is compatible with τ, one sets $k \to k-1$ and $\lambda' \to \lambda' + 2^k$ (the sub-arc length is increased) and moves to next step;

- At the next step the previous $\{\overline{\gamma_{12}}\}$ is used in a self-iterative manner to discard the pairs that are not compatibles with τ (without calculating their MVA);
- The iteration procedure stops when $k = 0$.

Closure procedure

The detection of the linked portion requires the transformation of a pair of open chains (γ_1, γ_2) into closed curves. The closure procedure is sketched in figure 2.9. Given two open chains (γ_1, γ_2) one first computes their centre of mass. Then draws two segments starting from the extremities of γ_1 and pointing away from the centre of mass of γ_2. The transformation of γ_1 into a loop is then performed by joining the ends of these segments along the surface of a sphere with radius very large compared to the extension of the pair (red dashed arc). This is repeated for γ_2. In order to minimize the additional entanglement that may

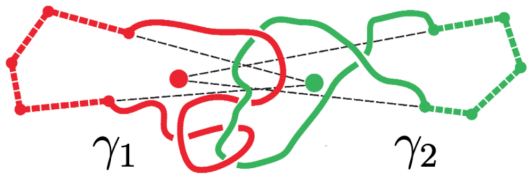


FIGURE 2.9: Closure procedure.

arise during closure, the segments pointing away from the centre of mass of the other chain are chosen to be very long compared to the radius of gyration of the chains.

This scheme is named *closure at infinity*. After the closure the two chains (γ_1, γ_2) are turned into the components of a proper link, whose topology is established with the two-variable Alexander polynomial.

2.3.3 Properties of the linked portion in bulk

The study of the linked portion of an entangled polymer system is a very recent problem. The concept of linked portion has been introduced in ref. [52], where also the first systematic method to find location and length of the linked portion has been presented (section 2.3.2). Contextually to its introduction, the first application of the method is given in the article, for linked chains in bulk.

In particular, Caraglio et al in ref. [52] considered a pair of linked rings of equal length, tied in the Hopf and Solomon topologies. They investigated the physical properties of the length of linked portion of the considered systems, varying the ring length N. The linked portion length is defined as the sum of the length of its constitutive arcs γ_1 and γ_2 (see figure 2.7).

In figure 2.10 there are the curves of the average linked portion length as a function of N.

The data points are interpolated by the dotted lines that correspond to power scaling laws:

$$\ell_{LK} \sim N^{\alpha} \text{ [52]} \tag{2.2}$$

with $\alpha = 0.36 \pm 0.05$ and $\alpha = 0.54 \pm 0.05$ respectively for the Hopf and Solomon link.

The fact that in both cases α is well below 1, is a strong indication that, even discounting finite size corrections, the average length of the linked region grows sublinearly with N. This

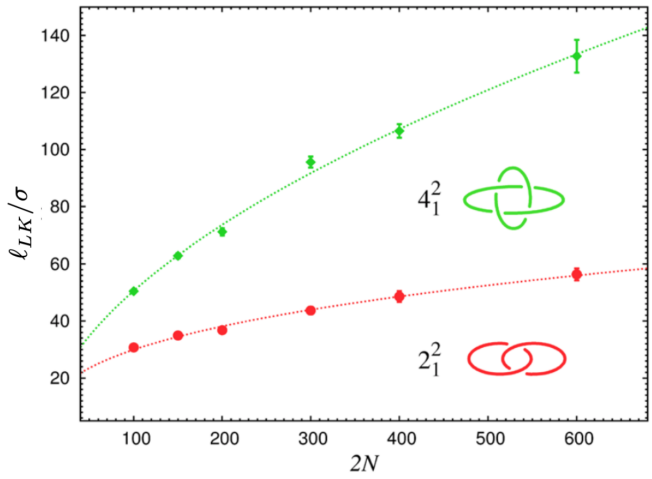


FIGURE 2.10: Figure adapted from [52]. Average contour length of the linked portions as a function of the total contour length $2N$, for Hopf (red circles) and Solomon (green diamonds) links. For each set of data the dotted curve corresponds to power law fits.

implies that the average inter-chain entanglement of linked rings in bulk is weakly localised in asymptotical long chains.

The study of unconstrained linked chains is an important reference for the interpretation of the confined case, of interest in this thesis. The results presented in this section have been indeed fundamental for the theoretical interpretation of the problem studied in chapter 3, about a pair of confined linked rings tied in the Hopf topology.

Topologically Linked Chains in Confinement

This chapter is based on the article *Topologically Linked Chains in Confinement* [53], published on April 2019 in ACS Macro Letters.

3.1 Background concepts for linear chains under channel confinement

In this section I will overview a set of key theoretical results for single linear chains under channel confinement. These results for open and hence unlinked chains are preliminary for the interpretation of our results for linked rings.

Linear semiflexible chains in channel confinement exhibit various scaling properties depending on the interplay of the relevant scales in the problem, namely, the channel width, D, the persistence length, l_p, and the thickness of the chain, σ [54–58]. In particular, for $D > l_p$ (the case of interest in this thesis) a polymer will only feel moderately the effect of confinement and its properties are describable in terms of the so-called *De Gennes* and *Extended De Gennes* scaling regimes (figure 3.1). A detailed description of these regimes is provided in the next sections.

3.1.1 De Gennes regime

De Gennes developed the classic model of a linear semiflexible polymer under confinement in 1977 [54,55], considering the case of a single chain of length $L = \sigma N$ (N number of beads, σ bead size) confined in a channel of diameter D, where $l_p << D << L$.

According to De Gennes model, confinement causes the chain to organize itself in a sequence of self-excluding *isometric blobs*, each of size equal to the channel diameter D (as

sketched in figure 3.1a). This means that the characteristic size, the gyration radius, of the portion of chain contained in every blob is equal to D:

$$R^0_{blob} \sim D \tag{3.1}$$

The contour length of the portion of chain in a blob is defined as $L_b = \sigma N_b$. The key ansatz introduced by De Gennes is that chain portions within blobs are not affected by confinement, so that each blob behaves as a self-avoiding walk in bulk. One can then derive the following scaling law for R^0_{blob}:

$$R^0_{blob} \sim L_b^{3/5} \tag{3.2}$$

where the (approximate) scaling exponent for the self-avoiding walk in bulk is used [54], $\nu = 3/5$.

The self-excluding blobs do not overlap and distribute along the channel in a sequential array, so the longitudinal span of the whole chain, $R^0_{\parallel}$, is given by the size of the blob, R^0_{blob}, times the number of blobs, L/L_b :

$$R^0_{\parallel} \sim D\frac{L}{L_b} \sim \frac{L}{D^{2/3}} \tag{3.3}$$

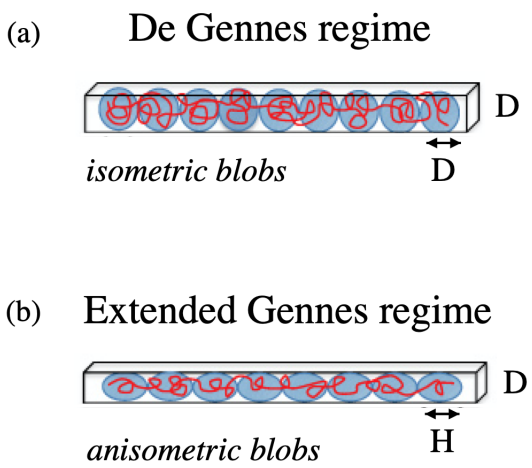


FIGURE 3.1: Sketches of a linear chain under channel confinement, in the De Gennes (isometric blobs) and Extended De Gennes regime (anisometrc blobs). Image adapted from [19].

3.1.2 Extended De Gennes regime

In 2008 Odijk established the threshold value for D above which the De Gennes regime holds: he argued that the De Gennes regime is defined for D larger than $D^{**} = \frac{l_p^2}{\sigma}$ [19, 56, 57]. Moreover, Odijk predicted the existence of a further scaling regime for $D < D^{**}$, holding in the range $2l_p < D < \frac{l_p^2}{\sigma}$.

Indeed, the threshold value $D^{**} = \frac{l_p^2}{\sigma}$ corresponds to the critical size above which the self-excluded volume effects become significant. This is established heuristically by noting that at D^{**}, ideal and self-excluding chains have comparable size: $l_p^{1/2} L^{**1/2} = (l_p \sigma)^{1/5} L^{**3/5}$, with $D^{**} = l_p^{1/2} L^{**1/2}$ being the scaling law for a portion of chain in a blob that behaves as an ideal polymer (with exponent $\nu = 1/2$) [19, 56, 57, 59].

Thus, below $D^{**} = \frac{l_p^2}{\sigma}$ excluded-volume interactions maintain the linear ordering of the blobs, but are not strong enough to drive the blob statistics to that of a self-avoiding polymer. The system is not describable anymore in terms of the De Gennes regime. According to Odijk, below D^{**} the blobs become *anisometric*, with transverse size D and longitudinal size H, figure 3.1b, in contrast with the isometric ones of the De Gennes regime.

In the Odijk picture, excluded volume interactions weakly affect the blob statistics in this new regime, consequently the blob extension H scales with the blob contour length as an ideal chain [57]:

$$H \sim l_p^{1/2} L_b^{1/2} \tag{3.4}$$

In ref. [56,57] the latter equation is used in conjunction with the fact that, at the crossover between the ideal and the self-avoiding chain limits, one has that $z = \sigma L_b^2 / D^2 H \sim 1$, where the ratio z is the so-called *chain interaction parameter* [15]. Thus, one finds that:

$$L_b \sim \frac{l_p^{1/3} D^{4/3}}{\sigma^{2/3}} \tag{3.5}$$

And, substituting equation 3.5 in 3.4:

$$H \sim D^{2/3} \tag{3.6}$$

Finally, despite the profound differences, the chain longitudinal size $R_{\parallel}^0$ is given by the same scaling law of De Gennes regime:

$$R_{\parallel}^{0} \sim H\frac{L}{L_b} \sim \frac{L}{D^{2/3}} \tag{3.7}$$

This regime has been called *extended De Gennes regime*.

The theoretical results illustrated in this section and in section 3.1.1 have been derived for single linear chains and there is no extension for circular chains, neither isolated nor linked.

3.2 Background concepts for ring chains under channel confinement

The case of circular chains under channel confinement has been investigated by numerical and experimental studies [60–69]. These studies have highlighted an unexpected effect for the case of two unlinked polymers under channel confinement: the so-called *entropic segregation* effect.

Entropic segregation occurs when two chains are confined together in a channel of diameter $D < R_{bulk}$, with R_{bulk} chain size in bulk, both for circular and linear chains [64–69]. In this condition, when a polymer is overlapped with a second one it loses conformational entropy and feels an effective repulsive force. This leads to the spatial separation of the two, with the repulsive force that increases with chain length and reducing the confining dimension: $f_{repulsion} \propto N/D$ [66, 69].

An example of entropic segregation is given in figure 3.2, for the case of two rings in a channel of diameter D.

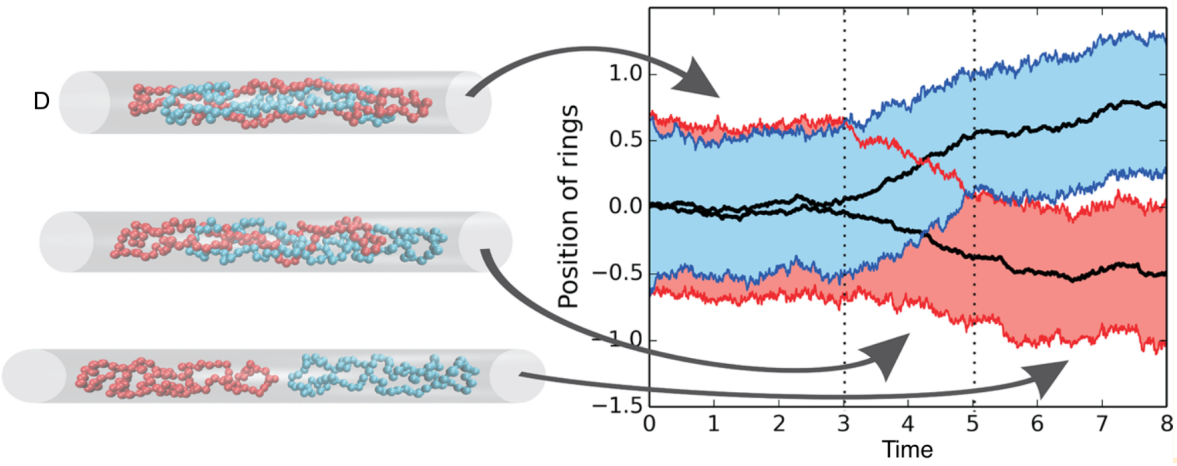


FIGURE 3.2: Figure from [66]. Two overlapping rings under channel confinement experience entropic segregation.

Chain topology determines a difference in terms of the strength of the entropic repulsion. Indeed, ring polymers experience a stronger entropic repulsion than linear ones with same length and in the same confinement conditions. This occurs due to the additional self-constraining effect of the rings. Indeed, considering a ring topology instead of a linear chain has been argued to be equivalent to reducing the confinement size to $D/\sqrt{2}$, and so the repulsive entropic force becomes larger [60, 66] (see figure 3.3). This is an interesting effect of the interplay between topology and segregation.

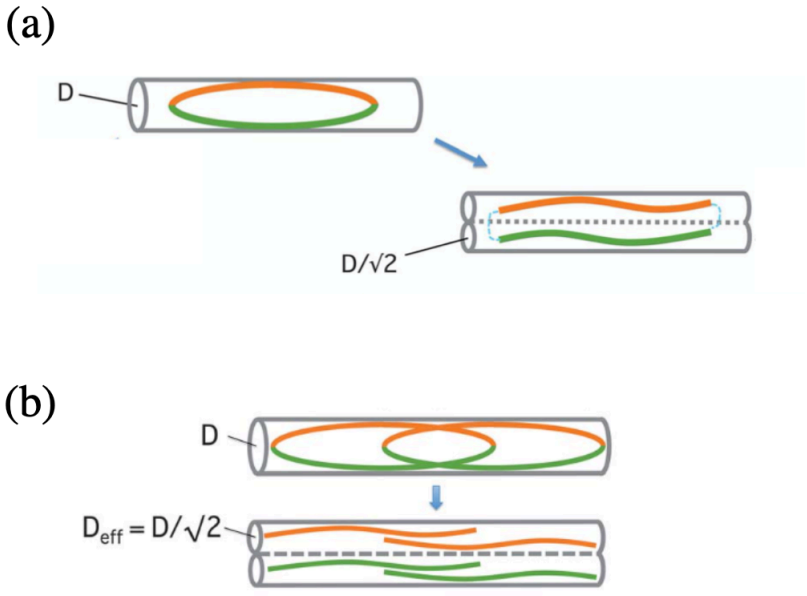


FIGURE 3.3: Figure adapted from [60]. (a) A ring chain is mapped into a "parallel connection" of two linear subchains, trapped in an "imaginary" tube with effective diameter $D/\sqrt{2}$. (b) The same mapping for two unlinked rings.

Entropic segregation has been indicated as the mechanism that drives the spatial segregation of bacterial catenanes after cellular replication [64,69] (see figure 3.4). Bacteria have only one circular chromosome that is duplicated in two identical molecules during the replication phase of the cell cycle. The daughter chromosomes are topologically linked and are referred to as catenases. Only at the end of the replication they become unlinked, thank to the action of the topoisomerase enzyme. After they are unlinked, they segregate in opposite sides of the cell, allowing the cell division into two new-born cells. Numerical studies modeled this biological process and showed that entropic segregation is responsible for the spatial separation of the unlinked daughter chromosomes.

In this chapter I study a pair of topologically linked rings confined in a channel, that can not be separated for construction. We ask if the segregation forces have a role also in this case and how the entropic repulsion deals with inter-chain topological entanglement.

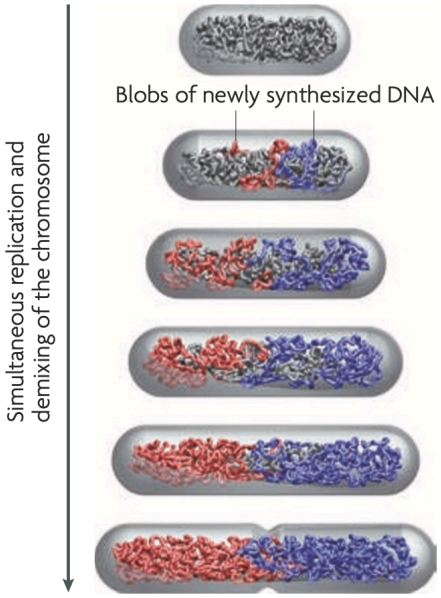


FIGURE 3.4: Figure adapted from [64]. Entropic segregation drives the spatial separation of bacterial chromosomes.

3.3 Model and methods

We considered a pair of equally long semiflexible rings of N beads tied in the Hopf link topology and confined inside a cylindrical channel of diameter D and axis corresponding to

the x Cartesian axis (see figure 3.5). The length of the rings and channel diameters are, respectively, varied in the $120 \leq N \leq 720$ and $11\sigma \leq D \leq 72\sigma$ range, where σ is the nominal diameter of the beads.

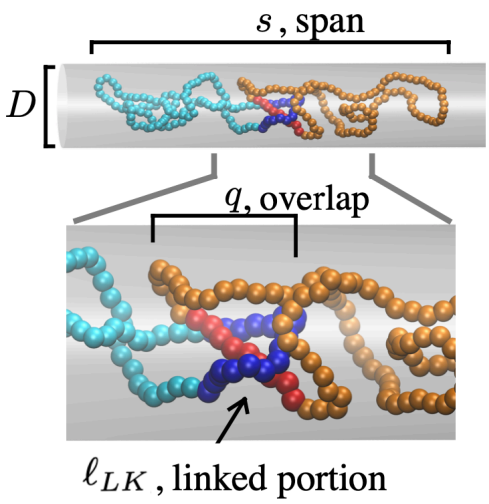


FIGURE 3.5: Typical conformation of two Hopf-linked semiflexible rings of $N = 120$ beads inside a channel of diameter D. The longitudinal span, s, overlap, q, and the linked portion, ℓ_{LK} are highlighted.

The system is studied with Langevin molecular dynamics simulations, with the chain potential energy given by:

$$U = U_{LJ} + U_{FENE} + U_{bend}, \tag{3.8}$$

where

$$U_{LJ} = \sum_{i,j>i}^{N} 4\epsilon \left[\left(\frac{\sigma}{d_{i,j}} \right)^{12} - \left(\frac{\sigma}{d_{i,j}} \right)^{6} + \frac{1}{4} \right] \theta(2^{1/6}\sigma - d_{i,j}) \tag{3.9}$$

$$U_{FENE} = -\sum_{i}^{N-1} 15\epsilon \left(\frac{R_0}{\sigma} \right)^2 \ln \left[1 - \left(\frac{d_{i,i+1}}{R_0} \right)^2 \right] \tag{3.10}$$

$$U_{bend} = \sum_{i=2}^{N-1} \epsilon \left(\frac{l_p}{\sigma} \right) \left(1 - \frac{\vec{b}_{i-1} \cdot \vec{b}_i}{|\vec{b}_{i-1}| \cdot |\vec{b}_i|} \right) \tag{3.11}$$

$$\tag{3.12}$$

In the above expression, $d_{i,j} = |\vec{r}_i - \vec{r}_j|$ represents the distance between two beads i and

j , $\vec{b}_i = \vec{r}_{i+1} - \vec{r}_i$ is the bond vector connecting two consecutive beads of the chain and θ is the Heaviside function: $\theta(x) = 0$ if $x < 0$ and $\theta(x) = 1$ if $x \geq 0$. ϵ is the characteristic unit of energy of the system, set equal to the thermal energy k_bT, and l_p is the persistence length of the chains, chosen equal to 5σ.

U_{LJ} is a truncated and shifted Lennard-Jones potential describing the excluded volume interaction between any pair of beads, including consecutive ones: two beads repel if their distance is less than $2^{1/6}\sigma$, which corresponds to the minimum of the potential. The U_{FENE} potential enforces the connectivity of the chain, so that two consecutive particles cannot distance more than $R_0 = 1.5$. U_{bend} is the bending energy, which penalizing consecutive bond vectors that are not parallel.

In addition, the chain is subject to the excluded volume interaction with the channel walls. This interaction is described with a truncated and shifted Lennard-Jones potential too:

$$Uwall = \sum_{i,j>i}^{N} 4\epsilon \left[\left(\frac{\sigma}{\delta_{i,j}} \right)^{12} - \left(\frac{\sigma}{\delta_{i,j}} \right)^{6} + \frac{1}{4} \right] \theta(2^{1/6}\sigma - \delta_{i,j}) \tag{3.13}$$

where $\delta_i = \frac{D}{2} - \Delta_i$ and Δ_i is the distance of the ith bead from the channel axis.

The dynamical evolution of the chain (initially prepared in a stretched conformation lying along the channel axis), follows the equation:

$$m\ddot{\vec{r}}_i = -\gamma\dot{\vec{r}}_i - \nabla_i U_i + \vec{\eta}_i \tag{3.14}$$

where m is the bead mass, γ is the friction coefficient and η is a Gaussian noise acting on each bead. η has zero mean and is delta-correlated, satisfying the fluctuation-dissipation relation:

$$\langle \eta_{i,a}(t) \rangle = 0 \tag{3.15}$$

$$\left\langle \eta_{i,a}(t)\eta_{j,b}(t') \right\rangle = 2K_bT\gamma\delta_{i,j}\delta_{\alpha,\beta}\delta(t - t') \tag{3.16}$$

where α, β and i, j respectively represent the Cartesian coordinates and the particle indices, $\delta_{i,j}$ is the Kronecher delta and $\delta(t - t')$ is the Dirac delta.

We use $m/\gamma = 2\tau_{LJ}$ [70], being $\tau_{LJ} = \sigma\sqrt{m}/\epsilon = \sqrt{m}/K_bT$ the characteristic simulation time. The dynamical equation is integrated numerically using the LAMMPS simulation package, setting as customary $K_bT = m = \sigma = 1$ [70, 71]. The integration timestep was set to $dt = 0.005\tau_{LJ}$ and the length of the simulated trajectories is $10^6 - 10^7\tau_{LJ}$.

Averages values of metric and topological observables were computed over hundreds of uncorrelated conformers for each (N, D) values.

3.4 Results

3.4.1 Equilibrium metric properties

The first set of results is shown in figure 3.6 and regards the D-dependence of the longitudinal *overlap* of the linked rings, q, for various chain lengths. The overlap is defined as the projection on the channel axis of the region in which the two rings are superposed (see the sketch in figure 3.5). Its measure is given by the difference between the longitudinal coordinate of the rightmost bead of the ring on the left and the longitudinal coordinate of the leftmost bead of the ring on the right.

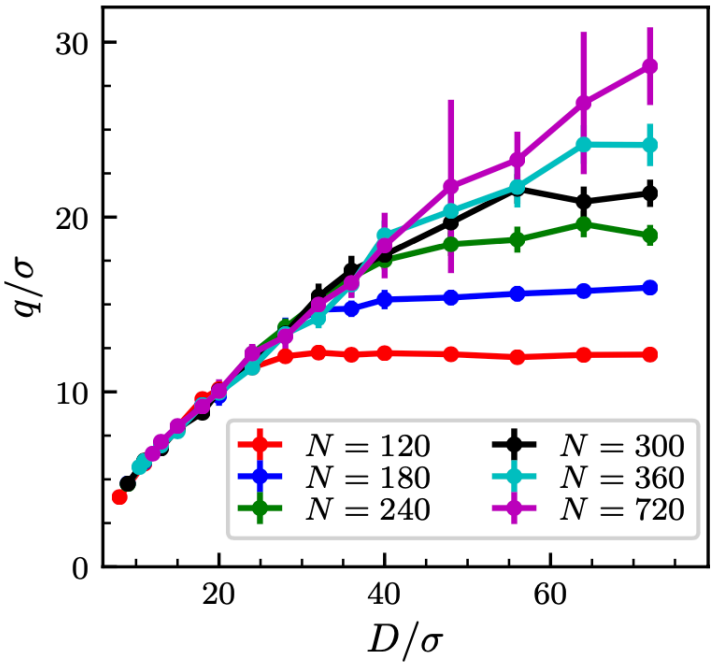


FIGURE 3.6: Average overlap of Hopf-linked rings as a function of D and for various chain lengths.

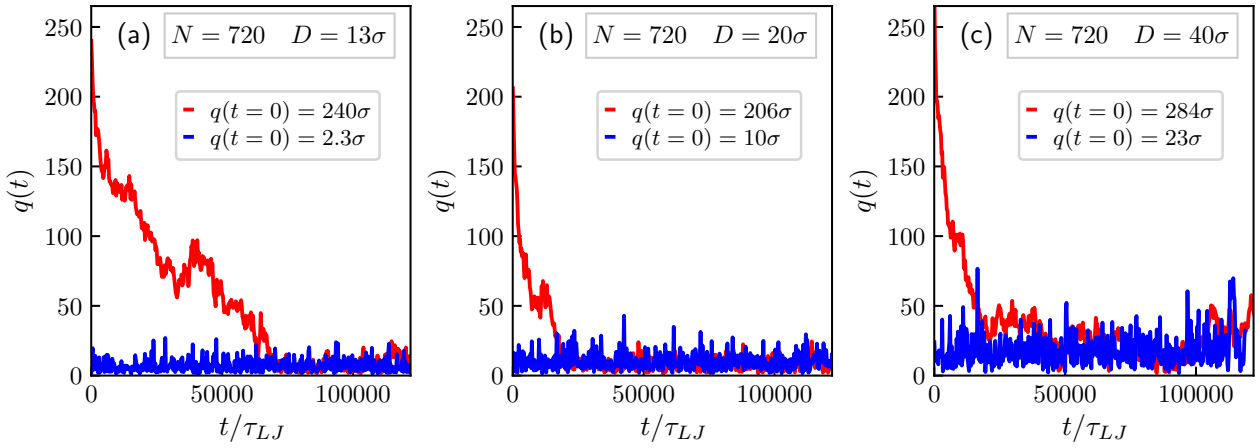


FIGURE 3.7: Time evolution of the overlap, q, in trajectories started from configurations where the rings in the link are compenetrated (red) or pulled tout (blue). The equilibrium values of the overlap are reached independently the initial degree of compenetration within a timespan comparable to the system autocorrelation time. Data are for $N = 720$ and $D = 13\sigma, 20\sigma, 40\sigma$.

In figure 3.6 we can see that for $D \gtrsim 50\sigma$ the curves for the overlap are already levelled to their asymptotic bulk values (which depend on N) at the smaller ring lengths, and that in general the curves for all N are well separated. In this weak confinement regime, the N-dependent overlap is thus well spread out at fixed D.

For $D < 25\sigma$, instead, this N-dependence is entirely lost. All curves are practically superposed despite the sixfold length variation, from $N = 120$ to $N = 720$. This is not due to the system being trapped in metastable states preventing equilibration. We checked this by performing different simulations starting from configurations where the linked rings are compenetrated or pulled tout (see figure 3.7). We did this for a representative value of N, 720, and for three values of D $(13, 20, 40)$. We verified then that the equilibrium values for the overlap are reached independently the initial degree of compenetration.

The significance and unexpectedness of this N-independent regime is aptly conveyed by figure 3.8a, where we contrast the overlap of linked and unlinked rings placed inside channels that are not infinitely extended but bounded by transverse walls. For each N and D, the wall spacing is set equal to the average span, s, of linked rings in infinitely extended channels (figure 3.8c). This setup creates the simplest conditions for an equal footing comparison of linked and unlinked rings: it weakly perturbs the former while constraining the latter to having approximately the same longitudinal footprint.

The comparison of the curves in figure 3.8a shows that the overlap at fixed $D < 25\sigma$ is practically N-independent for linked rings but decreases appreciably with N for unlinked

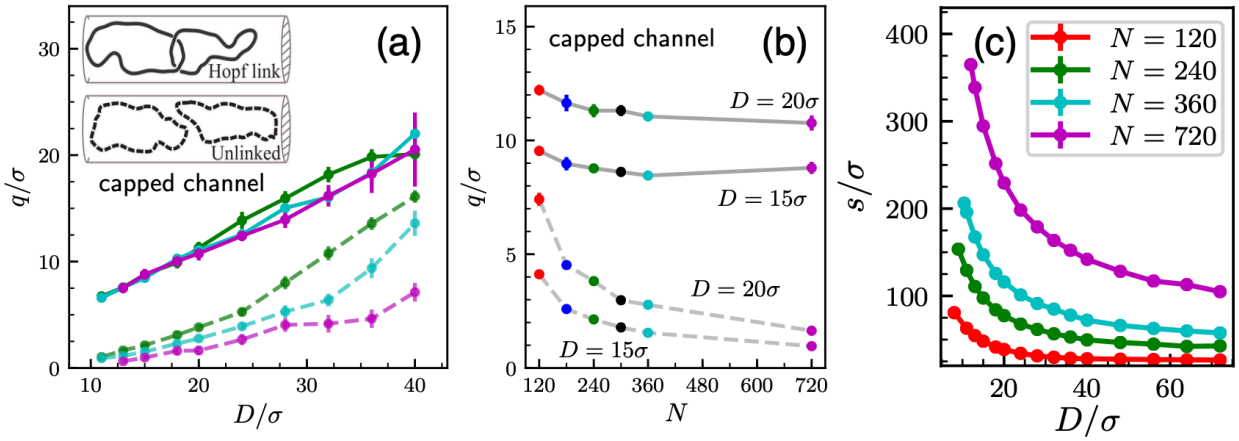


FIGURE 3.8: (a) Average overlap of linked and unlinked ones in finite capped cylinders of diameter D, as a function of D. (b) Average overlap as a function on N. (c) Average span of N beads Hopf-linked rings as a function of D, in infinitely extended channels. Error bars, often smaller than data point symbols, represent 95% confidence intervals.

ones, see also figure 3.9b. For Hopf links the typical overlap variation across $120 \leq N \leq 720$ is only about 10% and, in addition, q is never close to the smallest possible value found in highly stretched links, $q_{min} \approx \sigma$, see figure 3.7a. For unlinked rings, instead, as N is increases from 120 to 720 at fixed D, the overlap has a six-fold reduction so that, for longest rings, $N = 720$, it is typically smaller than one bead size. This is a characteristic signature of spatial segregation (illustrated in section 3.2), see also figure 3.9, and is consistent with the length-dependent increase of entropic segregation previously reported for confined separable rings [60, 66].

Altogether, the results of figure 3.8 make the following two key points. First, even when confinement is strong enough that longer and longer unlinked rings are more easily pulled apart by entropic forces, linked rings maintain a length-independent degree of intermingling. This is equivalent to saying that the topological linking constraint balances the length-dependent segregation forces in a manner that yields a surprising N-independent value of the overlap. Second, even when $D \sim l_p$, the overlap is never as small as the one of fully-tightened links.

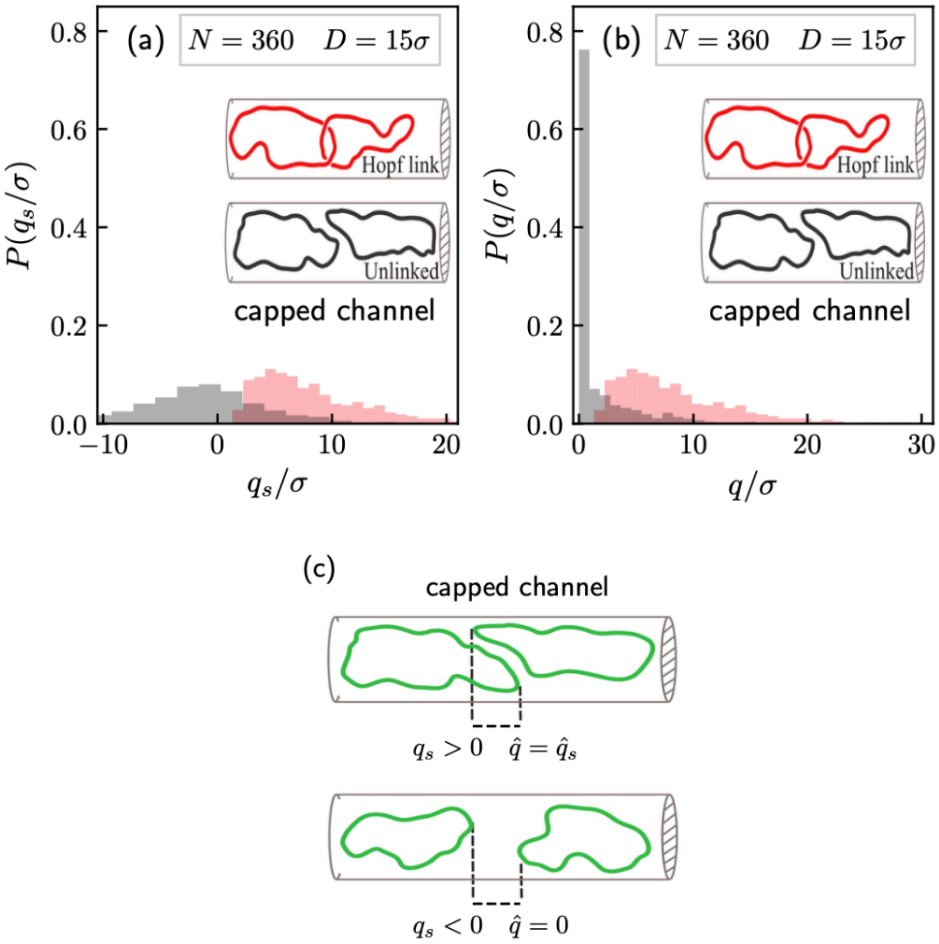


FIGURE 3.9: Normalised distributions of the (a) *signed* overlap, q_s, and (b) proper overlap, q, for linked and unlinked rings of $N = 360$ beads and inside capped channels with diameter $D = 15\sigma$. q_s is defined as the signed difference between the longitudinal coordinate of the rightmost bead of the ring on the left and the longitudinal coordinate of the leftmost bead of the ring on the right. Negative values of q_s therefore correspond to configurations that do not overlap but leave a gap, of size $|q_s|$, between them. The standard, proper overlap, q, is taken equal to q_s if the latter is positive and equal to zero otherwise. The peak at $q = 0$ in panel (b) shows that in most of the sampled configurations the two unlinked rings are fully segregated.

3.4.2 Equilibrium topological properties

To clarify the mechanisms underpinning these results, we examined the portion of the link where the inter-chain entanglement resides. We thus identify the linked portion (see chapter 2).

Examining the length of the linked portion, ℓ_{LK}, i.e. the summed length of its constitutive arcs, opens a totally new way to probe inter-molecular entanglement in spatial confinement. Indeed, ℓ_{LK} is an explicit measure of the total amount of contour length that is actually sequestered by inter-chain entanglement and, although nowadays is not yet accessible in

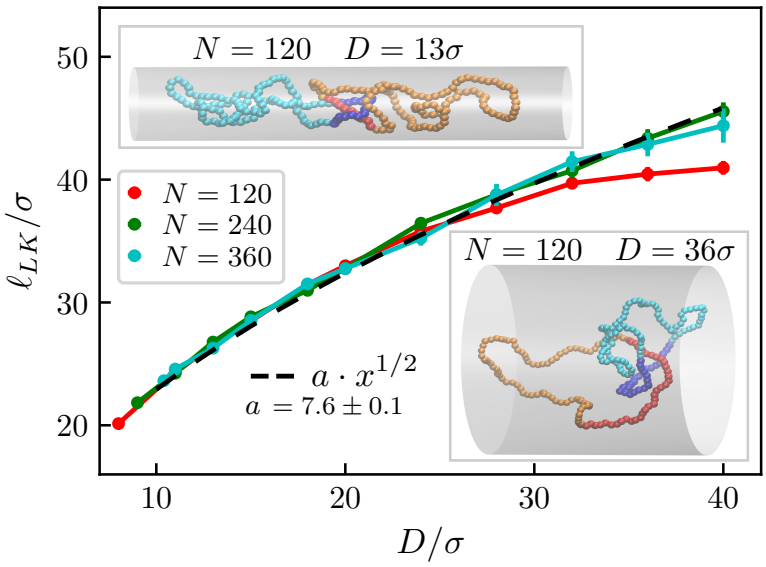


FIGURE 3.10: The average length of the linked portion is practically independent of N for $D < 25\sigma$ and follows closely a $D^{0.5}$ power law (dashed line). Error bars, often smaller than data point symbols, represent 95% confidence intervals. Two representative snapshots at different levels of confinement are shown in the insets.

microfluidic experiments, it directly affects experimentally measurable quantities such as the longitudinal span and overlap of the chains.

The ℓ_{LK} dependence on N and D is shown in figure 3.10. Although ℓ_{LK} is related to the overlap, as we can see in figure 3.11, it provides a more direct and stringent indication of the asymptotic independence of inter-chain entanglement on N for sufficiently high confinement. It is important to stress that in the considered range of D, the rings are still only weakly elongated by channel confinement and, in fact, the size of the linked portion, like the overlap, is far from the minimal value of the highly stretched case [52]. The collapse of the curves in figure 3.10 is thus unrelated to the expected insensitivity of ℓ_{LK} on N in tightly pulled links, which would be independent on D too. As a matter of fact, a fit of the data (dashed line in figure 3.10) shows that the limiting $\ell_{LK}(D)$ curve has a simple power law dependence on D:

$$\ell_{LK} \propto D^{0.5} \ . \tag{3.17}$$

3.4.3 Comparison with confined knotted rings

At this point it is interesting to do the comparison with the topological properties at equilibrium of intra-chain entanglement (knotting). The case of knots under channel confinement has been largely studied. We found that the inter-chain entanglement properties, in figure

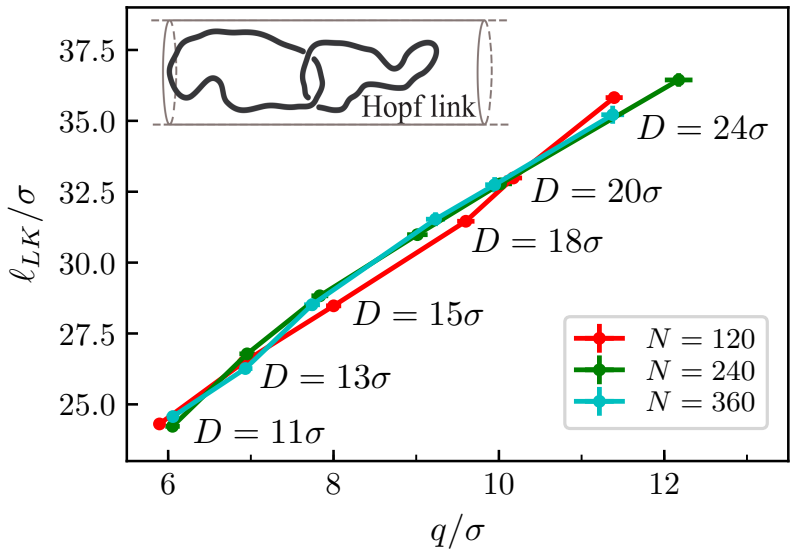


FIGURE 3.11: Relationship between the average length of the linked portion, ℓ_{LK}, and the overlap, q, for linked rings of $N = 120$, 240 and 360 beads in channels of widths D ranging from 11σ to 24σ. Error bars, often smaller than data point symbols, represent 95% confidence intervals.

3.10, contrast with those previously established for intra-chain entanglement in at least two ways. First, at all levels of channel confinement, the average length of the knotted portion (defined as the shortest portion of the ring that, once circularised with a suitably-chosen auxiliary arc has the same topology of the entire ring [72]), varies significantly with N due to the slowly decaying tails of the probability distribution of knot lengths (see figure 3.12) [30,73–75]. Secondly the different confinement compliance of the unknotted and knotted portions create a non-monotonic dependence of knot length on D [30, 31, 34], while the length of the linked portion varies monotonically with D (figure 3.12) .

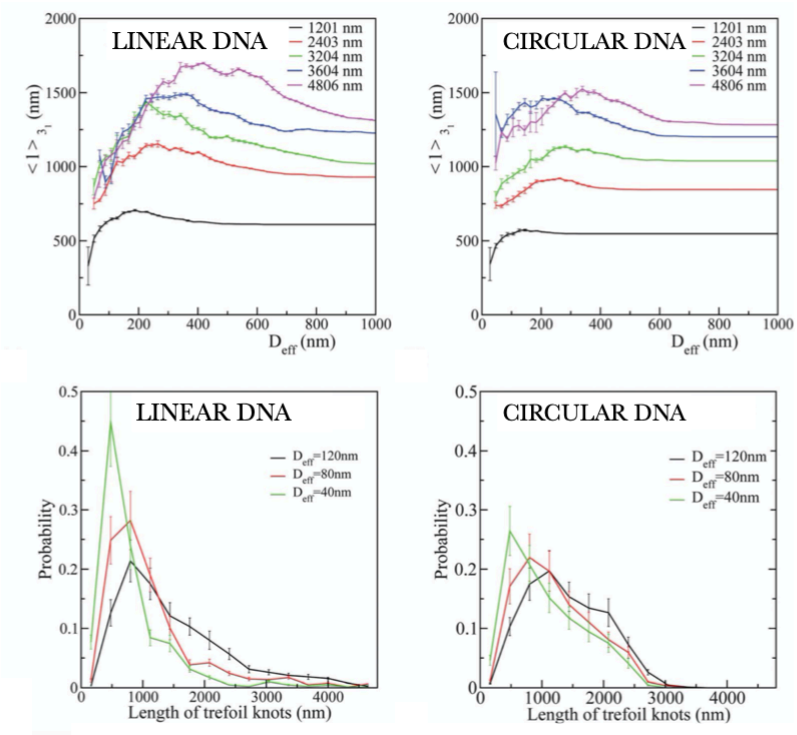


FIGURE 3.12: Figure adapted from [30]. Top panel: average lengths of trefoil knots occurring in confined linear and circular chains. Bottom panel: distribution of knot lengths for open and closed chains for three different values of D.

3.5 Theoretical interpretation

In summary, we observed two notable properties for the pair of linked rings under channel confinement. First, upon entering the weak confinement regime, the length of the linked portion, ℓ_{LK}, becomes independent of chain length. Second, at fixed N, ℓ_{LK} varies approximately as $D^{0.5}$.

An interpretation of these properties is given in this section.

In section 3.1 I have reviewed some important results about existing theories for linear polymers under confinement. There is no extension of them for the linked case, but they are useful in order to build an interpretative framework for our case of study.

In particular, we surmise that the collapse of the curves of the linked portion length in figure 3.10 and the observed scaling $\ell_{LK} \sim D^{0.5}$ can be rationalised in terms of the anisometric blob regime. Indeed, the range of channel widths of interest here is $11\sigma \lesssim D \lesssim 25\sigma$ where, for

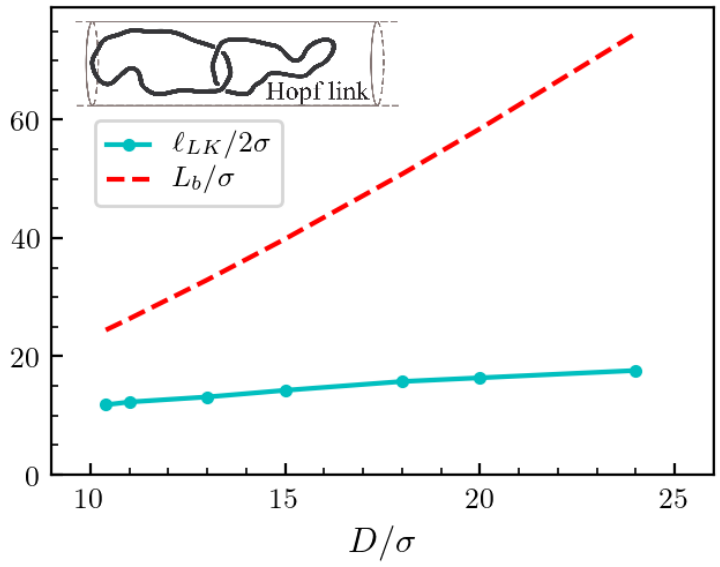


FIGURE 3.13: The observed average length of the linked portion, ℓ_{LK}, and the theoretical estimate of the blob size, L_b, are shown as a function of D for Hopf-linked links of $N = 360$. Error bars are smaller than data point symbols.

increasing N, one observes well-defined limiting curves for $\ell_{LK}(D)$ and $q(D)$, see figures 3.10 and 3.6. This range overlaps well with the interval $2\sqrt{2}l_p < D < \sqrt{2}l_p^2/\sigma$ where, according to the earlier works presented in section 3.1.2, the mesoscopic metric properties are describable in terms of anisometric blobs. Respect to the bound definition given in section 3.1.2, we actually added a factor $\sqrt{2}$, following the results reviewed in section 3.2. This factor accounts indeed for the effective reduction of channel section experienced by ring chains compared to linear ones [60] .

The estimated contour length of the anisometric blobs, L_b, is given by equation 3.5, which is shown as a function of D in figure 3.13 along with the semi-length of the linked portion. At all D's the semi-length of the linked portion is appreciably smaller than L_b, implying that the entangled region of each ring is entirely contained within one anisometric blob. Though this result is plausible *a posteriori*, it could not have been established *a priori* given the heretofore lack of systematic results for channel-confined linked rings.

It also explains, again *a posteriori*, why increasing N has no impact on the length of the overlapping or physically-linked region seen in figures 3.6 and 3.10. Since the linked region is entirely contained within only two juxtaposed anisometric blobs, one per each component of the Hopf link, it is practically screened from the other blobs. Consequently, as the latter grow in number with increasing N, there is no substantial effect on ℓ_{LK}.

The above argument can be used to predict the observed D-dependence of ℓ_{LK} when used

in conjuction with the fact [52] that ℓ_{LK} of unconstrained Hopf-linked rings with $l_p = 5\sigma$ scales as

$$\ell_{LK}(N, D = \infty) \sim N^{\alpha} \tag{3.18}$$

with $\alpha = 0.36 \pm 0.05$ (more details in section 2.3.3).

Indeed, the previous observations, imply that the entanglement properties of the link ought to be equivalent to those of two *unconstrained* Hopf-linked rings with the same persistence length $l_p = 5\sigma$ but with contour length equal to the blob one, $N = L_b$. Using equation 3.18 for $N = L_b$, one has $\ell_{LK} \sim (L_b)^{\alpha}$, that, after substituting equation 3.5, yields

$$\ell_{LK} \sim D^{0.48 \pm 0.02} \ , \tag{3.19}$$

which matches the phenomenological relationship of eq. 3.17 and figure 3.10.

3.6 Conclusions

In conclusion, the observed insensitivity of ℓ_{LK} on N and its specific power law dependence on D can be transparently accounted for by the localization of the linked portion inside a single anisometric blob for each ring.

This latter property could be relevant for the action of topisomerases that unlink newly-replicated chromosomes inside bacterial cells. New daughter chromosomes are inevitably created in a catenated state and their topological constraint must be removed by topoisomerases via DNA-strand passages catalysed at specific points, likely corresponding to hooked DNA juxtapositions [20, 23, 24]. Based on our results it appears plausible that DNA confinement inside the prolate bacterial cells helps to localize the linked region in a small portion of the entire genome. This would favour hooked geometries in correspondence of essential juxtapositions and would also cluster the latter together, thus assisting topoisomerase action. For a crude estimate of this effect one can consider double-stranded DNA filaments, for which $l_p \sim 50$nm and $\sigma \sim 2.5$nm, inside a channel with diameter equal to the typical *E. coli* transverse size, $D \sim 500$nm. From equation 3.5, the estimated length of the blob size is $L_b \sim 5\mu$m, which is much smaller than the mm-long bacterial DNA, and so would also be the physically-linked region contained inside it. We believe that models of these systems

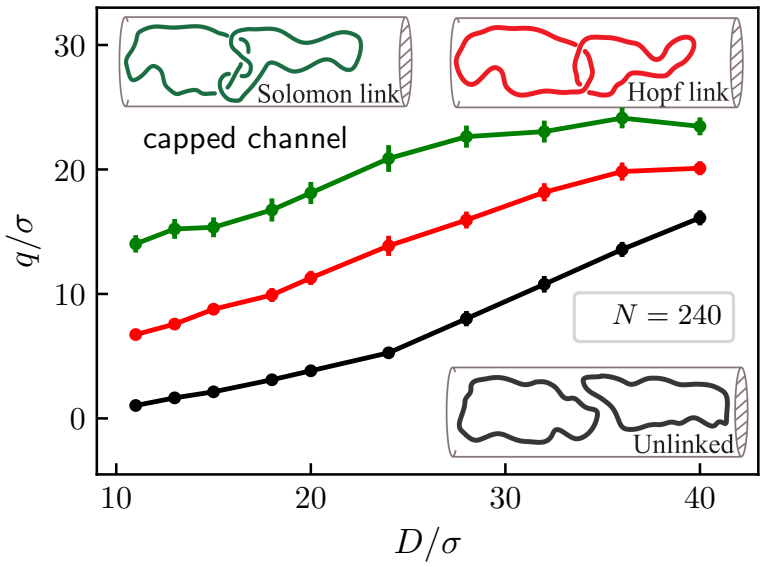


FIGURE 3.14: Comparison of the average overlap of unlinked rings and of Hopf- and Solomon-linked ones confined in equally-long capped cylinders of diameter D. Error bars, often smaller than data point symbols, represent 95% confidence intervals.

would be worth addressing in future studies particularly for the additional effects that the confinement along the longitudinal direction of the channel could introduce on top of the radial one.

More in general, it would be important to study how the geometry and dimensionality of the confining compartments affects the physical properties of linked chains. This aspect could be relevant in order to model different kinds of biological systems.

Another fundamental aspect to consider is also the study of the kinetics of the confined linked chains. In fact, biological processes cannot be comprehensively studied regardless of their time evolution.

Further promising extensions are suggested by the results of figure 3.14 for topological links more complex than the Hopf link. The graph compares the D-dependent overlap profiles of Solomon links with the unlinked and Hopf-linked pairs. The results show that the difference respect to the equivalent unlinked case is much stronger than for the Hopf link. It would therefore be interesting to study systematically the effect of increased topological complexity on the size of the linked portion at same channel width.

Finally, note that the properties predicted here for topologically linked rings under confinement ought to be detectable with current experimental setups. A possible model system would be DNA rings confined and linked inside narrow chambers [13, 76]. Similarly to recent advanced circular-to-linear unfolding experiments [12], flourescence kymographs or ultrathin

nanopore membranes [77] could then be used to probe the expected length-independent size of the catenated region and its undoing after the chains are made separable by an irradiation-induced double-strand cut [12] or, again, by topoisomerase action as in ref [76].

Effect of Dimensionality and degree of Confinement on the Equilibrium and Kinetics of Linked rings

This chapter is based on the draft of the article *Effect of Dimensionality and degree of Confinement on the Dynamics of Linked rings*, authors G. Amici, M. Caraglio, E. Orlandini, C. Micheletti.

In the previous chapter I gave a description at equilibrium of the metric and topological properties of a pair of linked rings under channel confinement. In this chapter I further investigate the physical properties of that system.

In particular, two very important issues remained open and totally unexplored. First, the effect of the dimensionality of the confining geometry on the equilibrium properties. Second, the characterization of the kinetics of the confined system.

To address these issues we considered two linked rings confined into channels (1-D confinement) or slits (2-D confinement), investigating then the equilibrium and kinetic properties.

4.1 Background concepts of polymer kinetics

In this section I present some concepts of polymer kinetics that I use in this chapter.

Relaxation time

An important observable for polymer kinetics is the polymer *relaxation time*. Its definition is given in the Rouse model [15].

The Rouse model describes the dynamics of polymers pictured as chains of beads connected by harmonic springs and subject to the stochastic motion arising from random collisions with the surrounding medium. The model neglects otherwise important contributions arising from the subtle interplay between monomer motion and the surrounding medium (hydrodynamic effects).

The description of the polymer kinetics is provided by the spectrum of Rouse times, τ_p, with $p = 0, 1...N-1$, that characterize the normal collective modes of the polymer (being N the number of monomers). According to this model, the first mode, with $p = 1$, is the *relaxation time* of the polymer, τ_r, i.e. the longest relaxation time of the chain.

Autocorrelation function

The Rouse model provides the theoretical definition of the relaxation time for the kinetic evolution of a polymer. The study of the kinetics of a simulated polymer, for which one has access to conformations collected according to a discrete time step, poses the necessity of a practical estimate for the relaxation time. In this context, the estimate of the relaxation time sets the time threshold above which polymer conformations can be considered uncorrelated.

The relaxation time for a simulated system is usually calculated from the autocorrelation function of the gyration radius [78]:

$$C(t) = \frac{\langle (R_g(t) - \langle R_g \rangle)(R_g(0) - \langle R_g \rangle) \rangle}{\sigma^2}, \tag{4.1}$$

where σ is the standard deviation.

The autocorrelation function $C(t)$ typically decays exponentially and, in general, with more than one characteristic time τ.

An unique estimate for the relaxation time of the polymer can be obtained integrating $C(t)$ and bypassing the multiplicity of characteristic decay times.

Using the integral criteria, the estimated relaxation time is defined as:

$$\tau_r = \int_0^\infty dt C(t) \tag{4.2}$$

Practically, $C(t)$ is summed in the limited range above which it is larger than 0.

Mean square displacement, diffusion coefficient and self-diffusion time

Another relevant observable for the kinetics of a polymer is the diffusion coefficient D_N. It can be determined from the *mean square displacement* (MSD) of the center of mass of the chain:

$$MSD(\tau) = \left\langle (\vec{r}_{CM}(t+\tau) - \vec{r}_{CM}(t))^2 \right\rangle_t \tag{4.3}$$

where $\vec{r}_{CM}(t)$ is the coordinate of the center of mass at time t and the MSD is calculated referring to a time lag τ.

The MSD for a polymer of length N satisfies the following relation:

$$MSD(\tau) = 2dD_N\tau \tag{4.4}$$

where d is the space dimensionality.

The diffusion coefficient of the chain, D_N, is proportional to the diffusion coefficient of a single monomer, D_0: $D_N \propto D_0/N$.

$D_0 = K_bT/\gamma$, where γ is the friction coefficient that appears in the Langevin equation (section 3.3). In our study the Langevin equation is integrated using LAMMPS [71]. Here the friction coefficient is defined as $\gamma = m/damp$, where $damp$ is a coefficient that has the dimension of a time. Adopting the choice proposed by Kremer and Grest in ref. [70], $damp$ is set equal to:

$$damp = 2\sigma\sqrt{m/k_bT} \tag{4.5}$$

In reduced units where $m = k_bT = \sigma = 1$ one has $damp = 2$, $D_0 = 2$ and $D_N = 2/N$.

Finally:

$$MSD(\tau) = \frac{4d\tau}{N} \tag{4.6}$$

For example, for channel confinement $d = 1$:

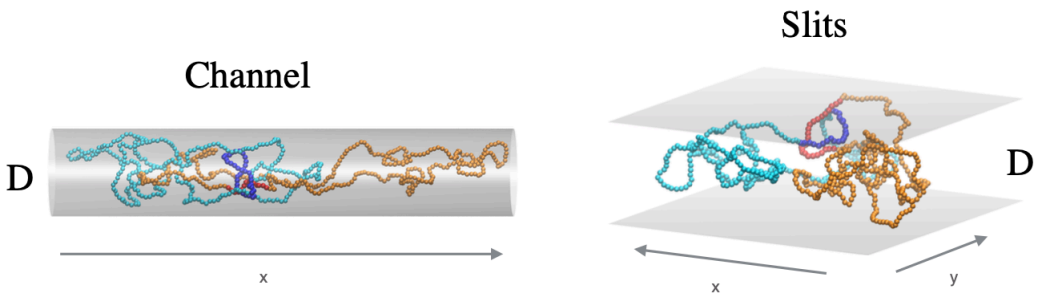


FIGURE 4.1: Typical conformations of two Hopf-linked semiflexible rings inside a channel of diameter D and slits of width D.

$$MSD(\tau) = \frac{4\tau}{N} \tag{4.7}$$

while for slits $d = 2$:

$$MSD(\tau) = \frac{8\tau}{N} \tag{4.8}$$

Finally, the *self-diffusion time*, τ_{sd}, is defined as the time at which the MSD of the polymer is equal to its square size, provided by the gyration radius:

$$MSD(\tau_{sd}) = R_g^2 \tag{4.9}$$

And so:

$$\tau_{sd} = \frac{NR_g^2}{4d} \tag{4.10}$$

4.2 Model and methods

We considered the same system of chapter 3, a pair of equally long semiflexible rings of N beads tied in the Hopf link topology. This time we confined them in channels of diameter D or in slits of width D (figure 4.1). The ring length is fixed equal to $N = 360$, while D is varied in the $10.4\sigma \leq D \leq 40\sigma$ range.

Again, the system is studied with Langevin molecular dynamics simulations, using the LAMMPS simulation package (for details see section 3.3) with integration timestep $dt = 0.005\tau_{LJ}$. The typical duration of the simulated trajectories at each value of D was $\sim 10^7\tau_{LJ}$.

4.3 Results

4.3.1 Equilibrium metric properties

We started by measuring the equilibrium metric properties of the system, comparing them between channels and slits.

In the previous chapter we considered the overlap and the span of the linked rings, measured along the privileged direction corresponding to the channel axis, which has no analog for slits. A more useful quantity to measure in the present case, for an equal comparison of channels and slits, is the gyration tensor. In particular, we looked at both the overall gyration radius and its *longitudinal* component (Rg and $Rg_{||}$). The latter is defined along the axis for channels (corresponding to the x cartesian axis) and lying on the plane for slits (corresponding to the xy cartesian plane).

The results for the gyration radius and its longitudinal component are shown in figure 4.2 (respectively in panel a and b). First of all, in both panels for small D the gyration radius is larger in the channel case. This is because in slits the system can spread freely in two directions, while in channels the linked rings are elongated along the axis. For larger D instead, the confinement effect is weaker in both cases and the distance between the curves is reduced.

Note that the intersection of the curves for $Rg_{||}$ in panel b follows from the different dimensionality of the longitudinal gyration radius in channels and slits, which implies that in bulk $Rg_{||}^{slit}/Rg_{||}^{channel} = \sqrt{2}$. Consistent with this, we observe $Rg_{||}^{slit} > Rg_{||}^{channel}$ in the large D limit of figure 4.2b.

However, the ratio of the longitudinal gyration radius at the maximum value of D, 40σ, is smaller than $\sqrt{2}$ (it is about ~ 1.1). This means that at the maximum D we are far from the bulk case, despite $40\sigma >> l_p$.

Finally, another interesting feature of figure 4.2 regards the relative variation of both Rg and $Rg_{||}$ from the minimum to the maximum D. This is about 30% for channels, and only 3% in slits: an interesting effect of confinement dimensionality.

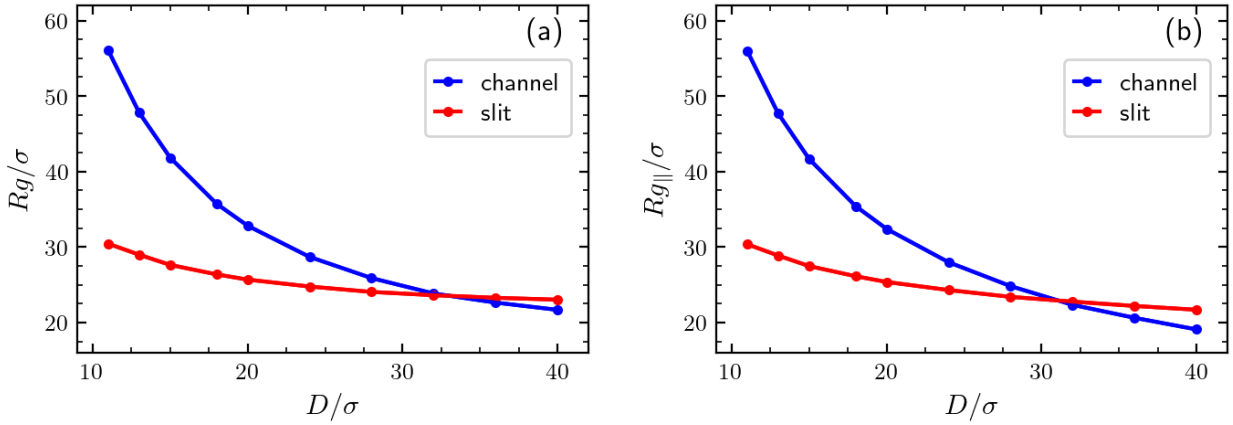


FIGURE 4.2: (a) Gyration radius, Rg, of the system confined into channels (blue) and slits (red) as a function of D. (b) Longitudinal component of the gyration radius, $Rg_{||}$. Error bars are smaller than point data symbols.

4.3.2 Equilibrium topological properties

Next, we studied the linked portion of the system at equilibrium, investigating the effect of the different confinement dimensionality on the region in which the entanglement resides. We measured the linked portion length, l_{LK}, and we compared the trends for channels and slits varying D. The result is shown in figure 4.3. Again, we observed the effect of a stronger constraint in channels. Indeed the linked portion length, in the considered range of D, is always smaller for channels than for slits. In particular, at small D, l_{LK} is about half for channels. This means that the linked region is more localised in channels than in slits.

4.3.3 Kinetic evolution of metric properties

After having explored the equilibrium properties, we moved to the kinetic ones, starting from the metric *relaxation time* of the whole system (see section 4.1). We measured the relaxation time from the autocorrelation function of the longitudinal component of the gyration radius, $\tau_{Rg_{||}}$.

In figure 4.4 we can see the relaxation time as a function of D, for channels and slits. There are opposite trends for the two confining geometries: increasing for channels and decreasing for slits. Moreover, for small D the relaxation time is larger for slits than channels, while at large D it is the opposite.

We asked if this opposite trend was a general kinetic consequence of the different confining dimensionality, in which case it ought to hold for confined unlinked rings too, or if it was

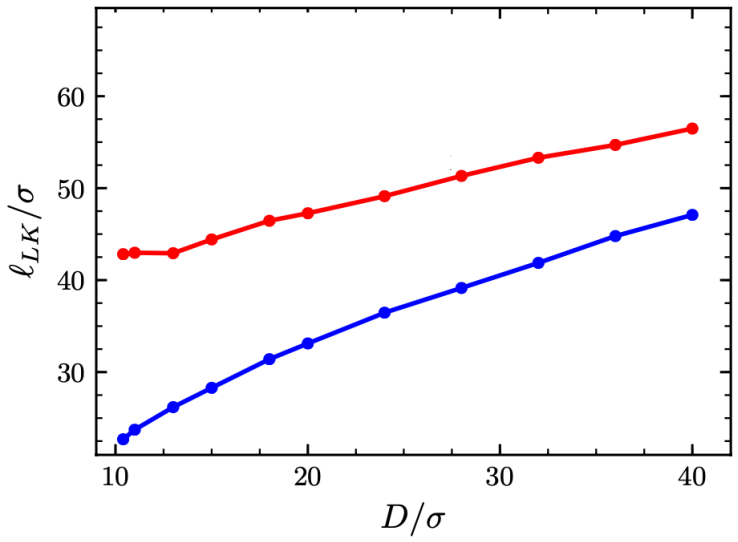


FIGURE 4.3: Linked portion length of the system confined into channels (blue) and slits (red) as a function of D. Error bars are smaller than point data symbols.

due to the presence of the topological constraint. We compared then the relaxation time of the pair of linked rings, each of length N=360, with the relaxation time for a circular chain of length N=720, in the same confinement conditions. The curves for the circular chain are provided in figure 4.4 in dotted lines.

The relaxation time trends are the same for the pair of linked rings and the circular chain. We conclude that the increasing or decreasing trend of the relaxation time is a general consequence of 1-D and 2-D confinement. The presence of the topological constraint does not change that behaviour. Its effect regards instead the magnitude of $\tau_{Rg_{||}}$: it determines a systematic reduction of $\tau_{Rg_{||}}$ for all types of confinement.

4.3.4 Kinetic evolution of topological properties

Motion in space of the linked portion

The measure of $\tau_{Rg_{||}}$ give us information about the overall kinetics of the system, but we were also interested in the internal kinetics of the linked portion, that has never been studied until now. In particular, we aimed to understand how the linked portion moves in space and along the rings' contour.

We studied the motion of the linked portion in space by measuring the mean square displacement of its central beads, MSD_{LK}. In figure 4.5 there are the cases of minimum and maximum D, for channels (panel a) and slits (panel b). First, we note that MSD_{LK} is larger

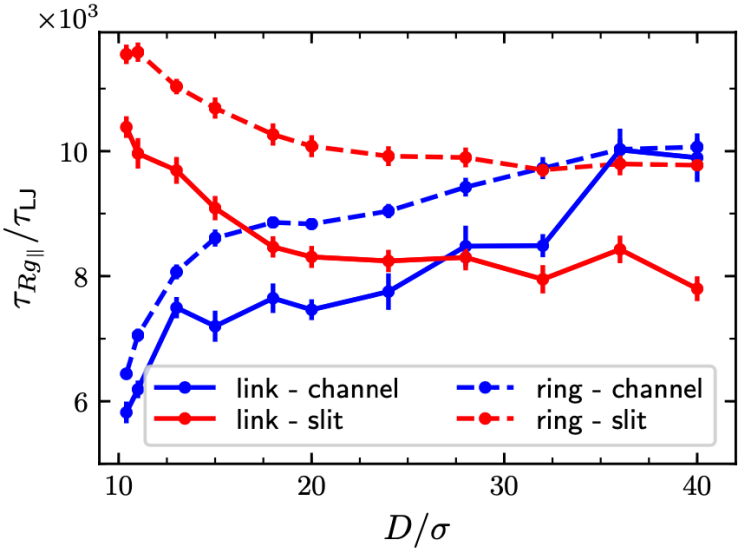


FIGURE 4.4: Relaxation times for channel (magenta) and slits (orange) for the Hopf link with rings of length N=360, compared with the relaxation times of a circular chain of N=720 in dotted lines.

in slits, consistently with the larger dimensionality. More interestingly, in particular at the minimum value of D, the curve in channels has two different slopes, at small and large times. This is not observed in slits instead.

To understand this feature, we compare the curves for channels at small times with the expected trend for the MSD of the center of mass of an isolated ring with N=360 (whose definition is given in 4.1):

$$MSD(t) = \frac{4}{N}t = \frac{t}{90} \tag{4.11}$$

The relative curves (for minimum and maximum D) are provided in figure 4.5a in black dotted lines.

The good agreement with the expected trend for an isolated ring suggests that there are two regimes for MSD_{LK} in channels: the linked portion beads initially diffuse like they belonged to an isolated ring, and only after a certain time, because of the topological constraint, follow the center of mass of the entire system. In slits, instead, we surmise that the first regime is so fast that it is not resolved, and MSD_{LK} soon follows the expected trend for center of mass of the whole system.

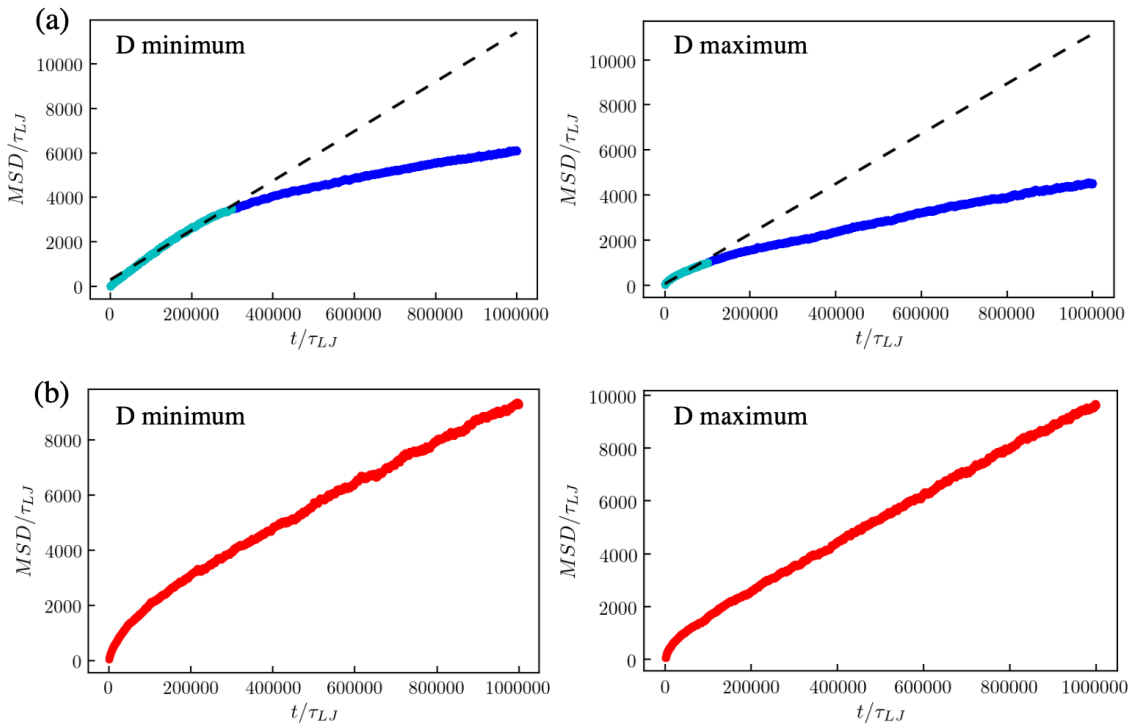


FIGURE 4.5: (a) MSD_{LK} as function of time in channels (blue) at minimum and maximum channel diameter. (b) MSD_{LK} as function of time in slits (red) at minimum and maximum slit width.

Motion along the chain of the linked portion

We next investigated the motion of the linked portion along the rings' contour.

In order to do this, we studied the diffusion of the linked portion central beads along the chains. In figure 4.6 there is the diffusion coefficient as a function of D for channels and slits. We see that it is always larger in slits. In particular, at small D it is almost an order of magnitude larger. This indicates a slower motion of the linked portion beads along the contour in channels compared to slits.

This observation suggests that the two confining geometries determine different responses for the diffusion of the linked portion along the contour. This is clear looking at the self-diffusion time, τ_{sd}, in figure 4.7, defined as the time required by the linked portion to diffuse over a contour distance equal to its average size $\langle \ell_{LK} \rangle$ (at fixed D). Indeed, first of all the self-diffusion time shows pronounced differences at small D between channels and slits, with τ_{sd} larger in channels. Second, and more interesting, changing the confining dimensionality has a strong effect on the relative variation of the self-diffusion time from the minimum to the maximum D. Indeed, at the minimum D in channels τ_{sd} is twice the value at the maximum

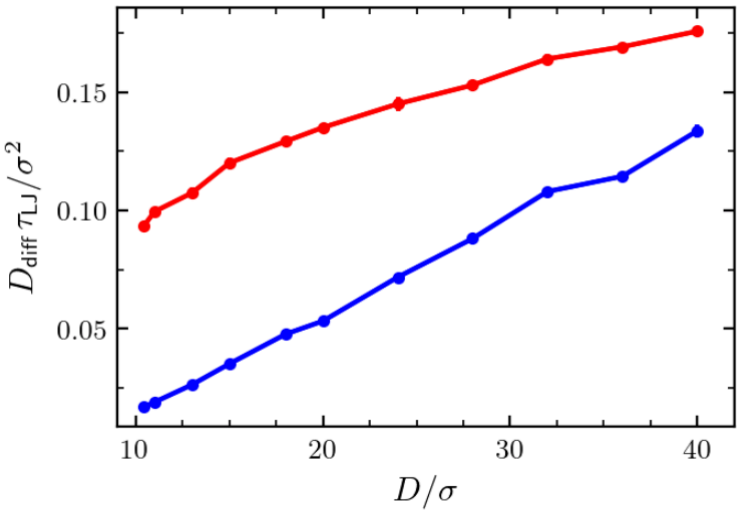


FIGURE 4.6: Diffusion coefficient of the linked portion length beads along the ring contour, for channel (blue) and slits (red). Error bars are smaller than point data symbols.

D (from more than 15000 at 10.4σ to $8000\tau_{LJ}$ at 40σ). In slits instead τ_{sd} has a smaller excursion, and overall the curve explores a limited time range: from 10000 to $8000\tau_{LJ}$. These results indicate that in channels the variation of the degree of confinement has stronger impact on the contour diffusivity.

To further investigate the effect of the confining dimensionality on the diffusion of the linked portion along the rings' contour, we measured the linked portion *renewal time*: the time required by the linked portion to reach a diametrically opposite point along the ring. $\tau_{renewal}$ is then defined as the time at which the mean squared displacement of the linked portion central bead along the contour is equal to the semi-length of the rings, $N_{semi} = 180$. Results in figure 4.8 show significative differences in the relative variation from the minimum to the maximum D between channels and slits. There is a variation of about an order of magnitude in channels, while in slits $\tau_{renewal}$ is practically constant upon varying D. This indicates that the entangled region is very persistent in small channels compared to larger ones and that this is a special property of channel confinement, that has no analog in slits.

At this point is interesting to do a comparison of the timescales involved in the internal kinetics of the linked portion, captured by the measure of $\tau_{renewal}$, and in the kinetics of the entire system, captured by $\tau_{Rg_{||}}$ (figure 4.4). For both confinement dimensions and at all explored values of D, $\tau_{renewal}$ is at least one order of magnitude larger than $\tau_{Rg_{||}}$. The relative

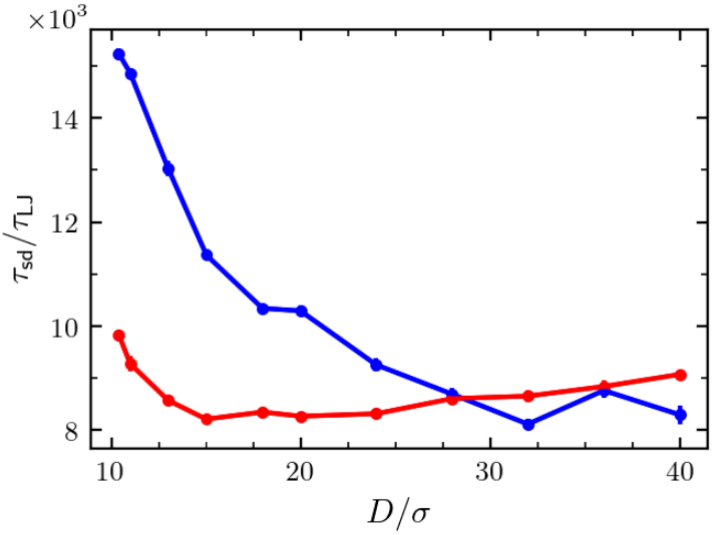


FIGURE 4.7: Self-diffusion time of the linked portion length beads along the ring contour, for channel (blue) and slits (red). Error bars are smaller than point data symbols.

difference is largest for the narrowest channel, where it is more than two orders of magnitude. This clarifies that the timescales of fluctuations of overall metric properties, as captured by $Rg_{||}$, are largely decoupled from the internal (contour) evolution of the linked portion that evolves much more slowly. Besides the theoretical implications, this has practical ones too, as it indicates that the experimental profiling of equivalent overall metric features, such as the span, are not good of the linked portion dynamics, which is much more persistent and has a qualitatively different dependence on D.

Unlinking time

We wonder if the renewal time could be a proxy for the system *unlinking time*, i.e. the time needed for the untying of the two components of the system if we cut ones of the ring bonds at the diametrically opposite point away from the linked portion. Doing this means changing the nature of the entanglement from topological or permanent to geometrical or transient. The results for $\tau_{unlinking}$ is shown in dotted lines in figure 4.8. We found that the unlinking time is much smaller than the renewal time and is also much less sensitive both of the confining dimensionality and of the degree of confinement. Once the cut is made, the geometrical linked chains come untied very fast. From this, we establish two conclusions. First, the sensitivity of the renewal time on the dimensionality and degree of confinement is largely due to the closed, ring nature of the chains. Second, because this timescale largely

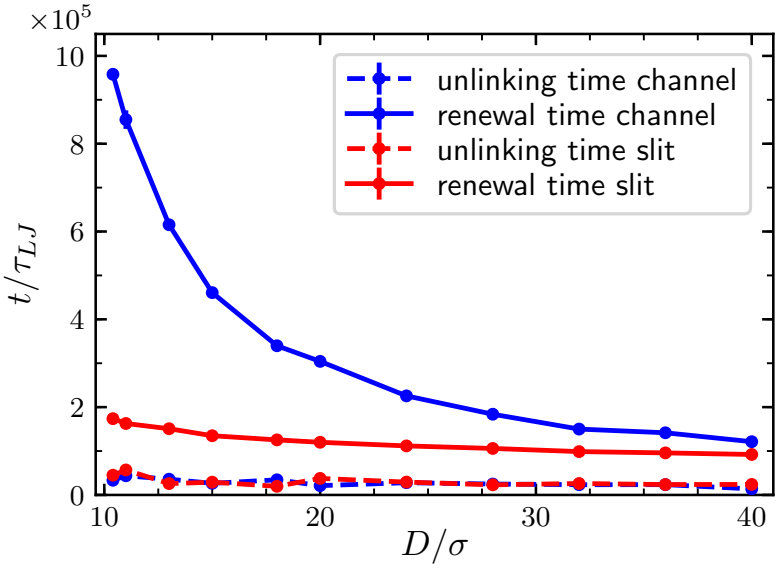


FIGURE 4.8: Renewal time (full lines) and unlinking time (dotted lines) of the linked portion, for channel (blue) and slits (red). Error bars are smaller than point data symbols.

exceeds the one of unlinking at all values of D, one has that the mechanism of unlinking is qualitatively different from the one governing the stochastic contour displacement of the linked portion.

4.4 Conclusions

In conclusion, we characterized the metric and topological properties of a system made of two linked rings under channel and slit confinement, both at equilibrium and out of equilibrium.

The study of the system at equilibrium has shown that the different confining dimensionality results in a more elongated system in the channel case, and in a more localised linked portion, whose length is twice as in slits at fixed D. For both the metric and topological properties these are the natural consequences of the passage from 1D to 2D confinement.

The study of the kinetic properties provided interesting and unexpected features, in particular regarding the diffusion along the rings' contour of the linked portion. Indeed we found that its diffusion properties are strongly dependent on the confining dimension and the degree of confinement. We showed that the diffusion along the contour is much slower in channels compared to slits, and in particular that the renewal time required to the linked portion to diffuse to the diametrically opposite portion of the ring is nearly an order of magnitude larger for channels than for slits. This suggests then that in channels the beads in the linked portion tend not to leave their position along the chain, and so that the linked portion constitutive

beads are persistent and does not change rapidly, in comparison with the timescales involved in the fluctuations of the overall system, captured by the whole system relaxation time. Indeed, we found that the internal dynamics of the linked region is largely decoupled from the fluctuations of overall metric properties.

We also compared the linked portion renewal time with the unlinking time of the system. This measure allowed us to investigate the change in the nature of the entanglement from permanent to transient, thus providing a first quantitative insight into how folded linear polymers can negotiate topological constraints. We discovered that the unlinking kinetics of the linear chains is qualitatively different from the renewal kinetics of topological constraints in channels. The disengagement mechanism of the linear system can be extremely faster than the linked portion renewal time, especially when the effective dimensionality of the entangled chains is small.

These observations taken together show that the interplay between confinement and linking results in different kinetic properties, depending on the geometry of confinement. In particular, channel geometry induces very peculiar properties, resulting in a persistence of the beads at the entangled region between the two rings.

This could have a significative biological relevance. In fact, the emergence and time evolution of linking between fluctuating filaments is a ubiquitous phenomenon in nature. It is observed in biological systems such as interlocked DNA rings in kinetoplasts [79], domain-swapped protein complexes [80], intermingling of neighbouring chromosomes [81] and catenanes in bacteria [20].

According to our findings, the confinement of entangled biomolecules can determine different physical properties according to the type of compartments in which they are contained. This could make the difference in the biological processes in which they are involved. Also the topology of the chains can be determinant in the kinetics of such mechanisms. Having a ring or a linear topology, i.e permanent or transient entanglement, may lead to different kinetic properties and consequences on the biological processes. For example, our study could be extended in order to model catenanes in bacteria such as *E. Coli*. In chapter 3 we saw that in bacteria the two daughter ring chromosomes are linked in the Hopf topology until the topoisomerasis unlinked them, doing cuts in specific genomic sequences. The topological constraint is thus transformed into a transient one, allowing the chromosome separation.

In conclusion, the results we obtained here could have the potential capability to shed light

on the kinetic mechanisms underlying the biological processes in which entangled biomolecules are involved. Our work give a first significative contribution to the characterization of the kinetics of linked polymers in confinement, and it needs to be extended and further explored, in order to fully understand the physical behaviour of inter-locked molecular systems.

Part II

Loop extrusion and epigenomic-based compartimentalization as mechanisms underlying chromosome conformations

Introduction

The eukaryotic genome is contained in the cell nucleus and consists of long filaments of chromatin fibers, the latter resulting from the association of DNA with specific organizing proteins. The compaction of chromatin fibers inside the nucleus is impressive: for example human DNA is 2 meter long, and it is contained into a $10\mu m$-sized nucleus [82]. This level of compaction is achieved thanks to the peculiarities which determines how the genome folds. This structure evolves during the cell cycle, going from the mitotic initial shape to the interphase shape, and facilitating the biological processes in which the genome is involved.

The development of experimental tecniques that probe chromatin structure (such as FISH [83], Hi-C [84], high-resolution microscopy [85], GAM [86]) revealed a hierarchical organization for chromatin during interphase, made of domains into domains of increasing size [83–95]. From TADS to A/B compartments to chromosome territories.

In the last years, a huge amount of experimental information about this hierarchical structure have been collected, and many of the observed structural features were explained by relatively simple physical models. These models propose different mechanisms of chromatin folding, starting from minimal physical assumptions and try to capture various aspects of chromatin phenomenology.

For example it has been clarified the importance of a *passive* mechanism of chromatin folding in the formation of chromosome territories [96, 97]. It was shown that the positioning and contacts of the initial mitotic state are crucial in determining the final interphase organization, without the intervention of any kind of active process. Chromosomes territories emerge indeed as a consequence of the *memory* of the initial mitotic shape, with chromosomes relaxing in distinct positions in the nucleus reminiscent of their mitotic spatial separation.

On the other hand, also *active* mechanisms (ATP-dependent) of chromatin folding have been proposed, such as the loop extrusion mechanism, accredited for the formation of TADs

in interphase [98–101] and also for the compaction of mitotic chromosomes [102, 103]. The loop extrusion process consists in the formation of loops of chromatin fibers mediated by proteins, that leads to the organization in larger structural units.

Moreover, there are also phase separation models (such as block copolymer models) able to reproduce instead chromatin organization in A/B compartments, in which chromatin fibers fold according to an attractive interaction among fibers of the same type (A or B) [21,99,104].

Thus, the proposed models of chromatin folding are ad-hoc models able to reproduce different structural features. So far, people have investigated the role that each folding mechanism plays in chromosome organization, but a systematic study of their combination and relative relevance is still missing. Indeed, it is a priori not clear what is the impact of the different folding mechanisms when they act simultaneously or at different temporal steps during the cell cycle.

In order to have a chromosome model capable to comprehensively reproduce experimental features, it is important to understand the consequences of this interplay. The study of the time ordering in which the different mechanisms act could be a new and fundamental key of interpretation for chromatin folding.

In this work, we study a model for human chromosome 19, starting from a mitotic conformation and simulating then the effect of loop extrusion and/or phase separation, introducing them at different moments during the simulation. We want to compare the final structures, with reference to the structural features from experiments, in order to see which of the explored time orderings of folding mechanisms better reproduce the actual chromatin organization. This would allow us to speculate on what happens during the folding process in vivo.

More in details, this second part of the thesis is organized as follows.

In chapter 6 I the describe the chromatin structure captured by experiments. In particular, I illustrate the organization in TADs, A/B compartments and chromosome territories during interphase, and the compact structure of mitotic chromosomes. Then, I illustrate the physical models capable to reproduce the presented structural features, describing passive and active folding processes.

In chapter 7 I describe our ongoing work. I present the model and the methods and then

move to the obtained preliminary results. Finally, I illustrate the planned future steps and the perspectives of the study.

The presented study is realized with M. Di Stefano, A. Rosa and C. Micheletti.

Chromatin organization and folding processes

In this chapter I describe the spatial organization of chromatin inside eukaryotic nuclei, illustrating the structural features that emerged from experiments. Then, I present also the physical models that better succeeded in reproduce these experimental findings, proposing ad-hoc mechanisms of chromatin folding. These models allow us to do hypotheses on what happens during the folding in vivo, simplifying the problem and thus shedding light on the physical mechanisms that underlie the biological process.

6.1 Chromatin interphase organization

Genome inside eukaryotic cells is arranged in a spatial organization on more levels. At the root of this organization is the DNA double helix. The double helix is then wrapped around proteins, called *histones*, that are involved in the folding of DNA in a more compact structure: the *chromatin fibre* (figure 6.1). Chromatin fibers make up discrete functional units, the chromosomes.

The shape of the chromatin fibers is dynamic and evolves during cell life: from cell birth to the time in which the cell divides itself into two daughter cells. The sequence of events from the birth to the cellular division is known as cell cycle [82] and it is divided into two main phases (for somatic cells): the phase of growth, called interphase, that covers about 90% of the entire cycle, and the phase of division, called mitotic phase. In the mitotic phase chromosomes come in compact structures (figure 6.2b), while during interphase chromatin fibers are in a diffuse state (figure 6.2a).

Experiments revealed different scales of spatial organization of chromatin during interphase [105]. At the nuclear scale ($1-10\mu m$) *fluorescence in situ hybridization* (FISH) showed that the chromosomes are well distinguishable, each of them occupying a compact region of the nucleus, termed *chromosome territory* [83,95]. Other experimental tecniques [85,86], and

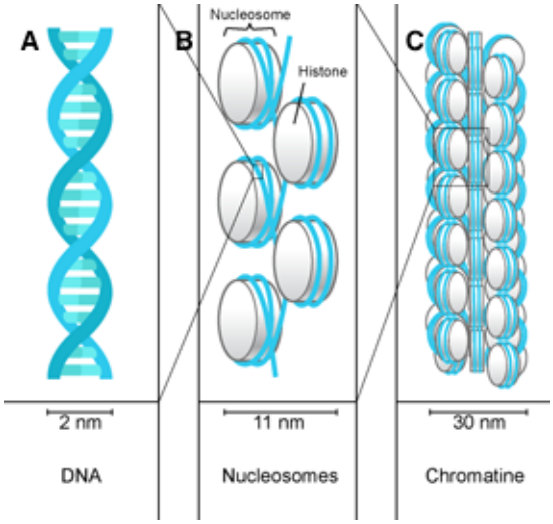


FIGURE 6.1: Different levels of DNA organization. (a) Two DNA strands wrap themselves around a common axis and form a double helix. (b) Interaction with histones packs DNA into a succession of bonded beads, the nucleosomes. (c) Nucleosome succession forms chromatin fibers.

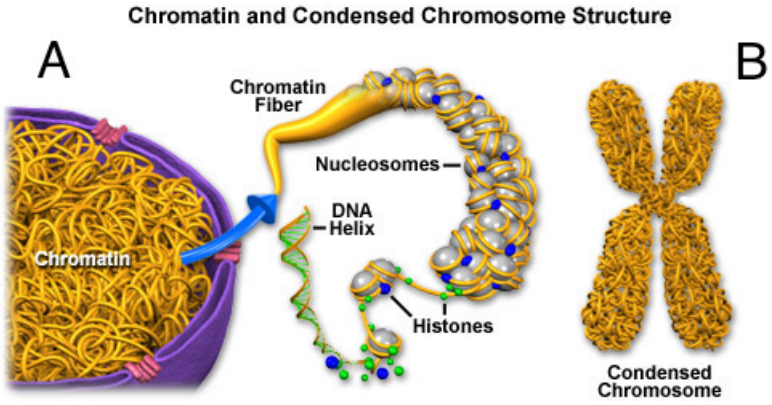


FIGURE 6.2: (a) Chromatin fibers inside the eukaryotic nucleus during interphase. (b) Condensed mitotic chromosome.

in particular the use of *chromosome conformation capture techniques* in their high-throughput version, the Hi-C tecnique, provided a multiscale knowledge of genome architecture between the small limit (chromatin fibers) and large limit (chromosome territories) [84, 87–91].

6.1.1 TADs

The study of Hi-C data reveals different levels of organization for chromatin. Indeed, with the Hi-C tecnique one can probe the structure of whole chromosomes and generate matrices of physical contacts among chromatin sequences of fixed size (see figure 6.3 for more details). This can be done for two different chromosomes (inter-chromosomal) or within a single chromosome (intra-chromosomal). The resulting contact matrices are represented as two-dimensional maps with genomic positions along the two axes (see figure 6.4). The size of the probed chromatin sequences, measured in kilo base pairs, *kbp* (the unit of measure for genomic distance), defines the resolution of the map. Using maps with different resolutions, one can study chromatin organization at different scales.

In particular, from Hi-C intra-chromosomal maps emerges that, on the scale of a few *kbp* up to about $1Mbp$, chromosomes are organized into self-interacting regions, called *topologically-associating domains*, TADs (figure 6.4) [93,94]. These domains show a high level of interaction among chromatin sequences within the domain, while inter-domain interactions are depleted.

TADs constitute regulatory regions where loci and their associated regulatory elements interact frequently with each other, and are isolated from the rest of the genome. Moreover, the organization of the genome in TADs seems to be a general property, well conserved across tissues within the same species [107] and even between different species [108]. They are also stable during cellular division: they are disrupted in the mitotic phase of the cell cycle and then quickly re-established [109].

All these findings about the genome organization into TADs suggest that these domains represent not only structural building blocks, but also functional units that promote biological processes. Indeed, inside TADs there are long-range interactions that ensure the physical proximity of genomic sequences involved in regulatory interactions (such as promoter-enhancer interactions) [106]. These interactions are mediated by the presence of two kinds of proteins, named cohesin and CTCF [93, 106, 110]. Cohesin and CTCF promote the contact among distant genomic sequences through the formation of chromatin loops (figure 6.5). In

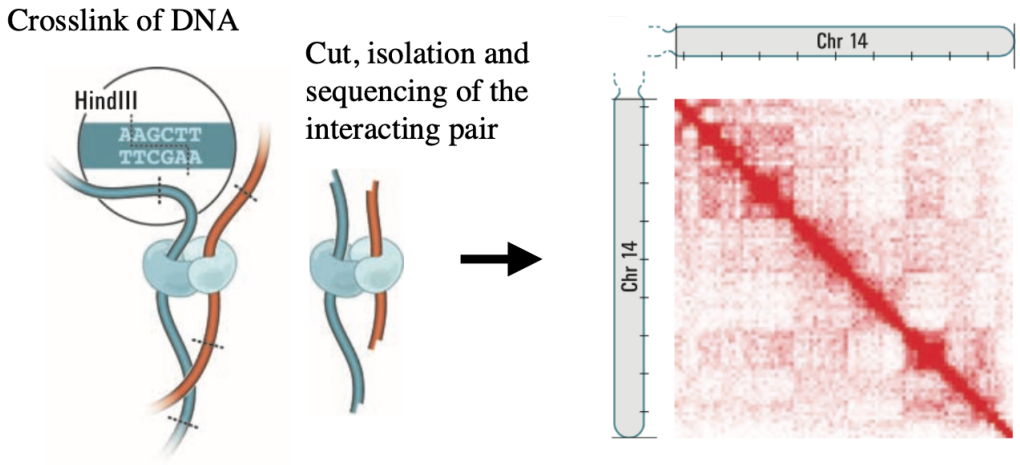


FIGURE 6.3: Figure adapted from [84]. Overview of Hi-C. With the Hi-C technique one can identify chromatin sequences that are in physical contact in a cell population. In order to detect them, the DNA of every cell is cross-linked with a molecule (the formaldehyde) that mediates the formation of chromatin links between any pair of genomic sequences in close enough spatial proximity to be cross-linked. The pair is then fragmented and isolated, in order to sequence it. A catalogue of interacting fragments with known base pair sequences is made, and the interacting pairs of chromatin fragments are assigned each at its position along the DNA chromatin sequence. So, from a Hi-C experiment one obtains a list of DNA-DNA contacts, which is called a "Hi-C library". The Hi-C library is used then to calculate the number of observed contacts between any pair of sequences with fixed size within a single chromosome or between two different chromosomes (averaging over the cell population). The inferred data are organized in a contact matrix, where the entries are the average number of contacts observed between pairs of loci. The contact maps is usually visualized as a heat map. Once obtained this map, and knowing the linear sequence of bases of a certain chromosome, one can track it back the 3D structure.

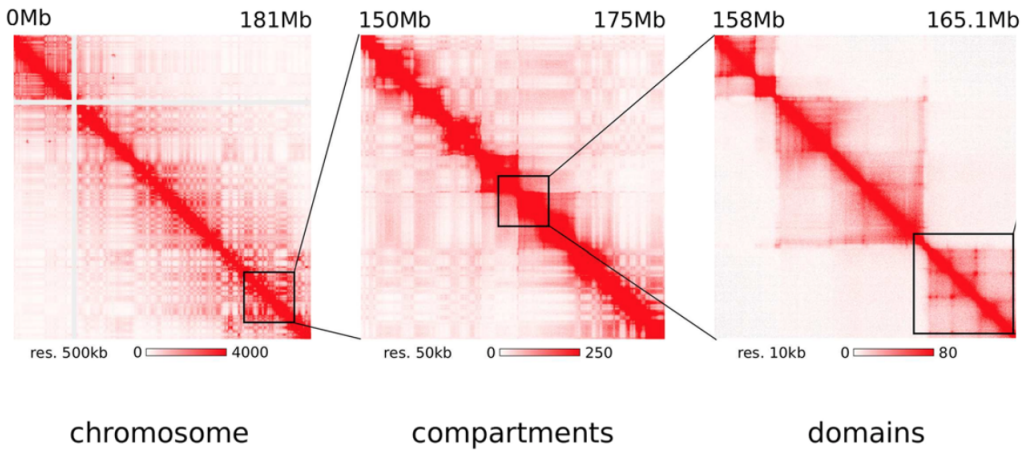


FIGURE 6.4: Hi-C intra-chromosomal maps show different levels of chromatin organization: from compartments to topologically associated domains.

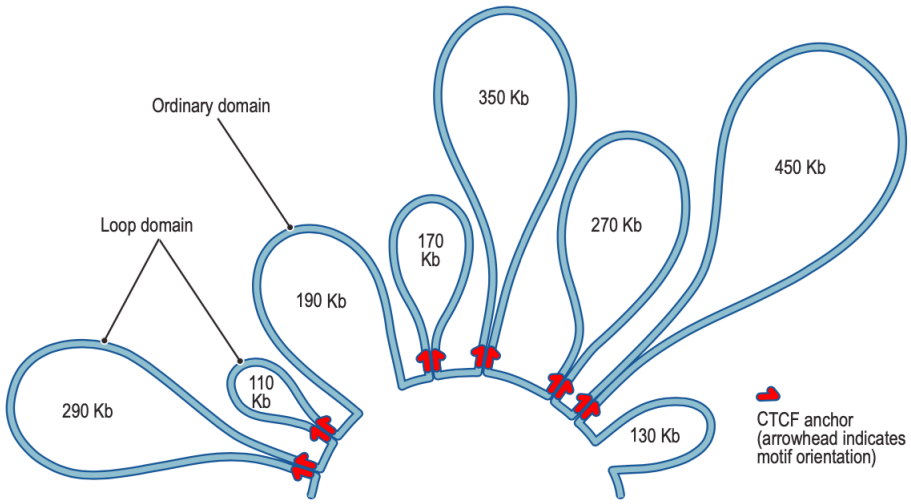


FIGURE 6.5: Figure from [106]. A schematic rendering of a 2.1 Mb region on human chromosome 20, in which there are loops anchored by convergent CTCF-binding sites.

particular, cohesin can topologically entrap DNA and slide along it and over small DNA-bound proteins and nucleosomes [111–113]. Experiments also showed a rich presence of cohesins at TAD boundaries [93]. CTCF is also enriched at TAD boundaries, and was found that the disruption of CTCF binding sites alters TAD structure [93, 94, 101, 114–116].

6.1.2 Compartments

At a larger scale, from Hi-C maps emerges that the next level of chromatin organization after TADs is into *compartments*, with size of few Mbp [84, 92] . These compartments appear as a "plaid pattern" on Hi-C maps (figure 6.4) and are characterized by the fact that elements of a compartment are rich of intra-compartment contacts, while inter-compartment interactions are depleted. They have been identified by using principal component analysis on Hi-C intra-chromosomal maps [84].

By comparing the identified compartments with known epigenetic features, it was found that the spatial organization of chromatin into compartments is strongly correlated to the chromatin state in the transcription process [117]. In particular, there are two types of compartments, identified as A and B. Compartment type A is associated to euchromatin (transcriptionally active) and B to eterochromatin (transcriptionally inactive).

The correlation of the genome compartment organization with the transcriptional state of chromatin is another signature feature of the fact that the chromatin 3D arrangement is closely linked to biological functionality. In this perspective, TADs and compartments are

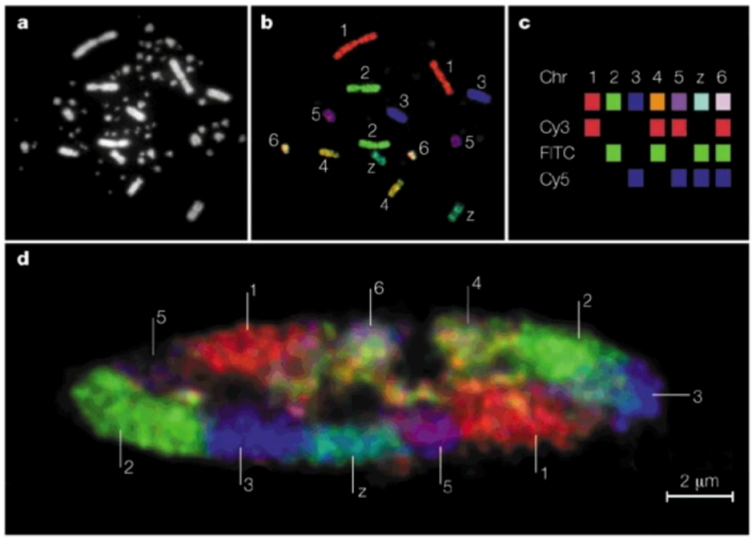


FIGURE 6.6: Figure from [83]. Chromosome territories during interphase capture with FISH.

not separated structural units, but are functional domains that in a concerted way promote the same biological processes, acting on different aspects and at different levels.

6.1.3 Chromosome territories

The last level of organization of chromatin during interphase, at the nucleus scale, is constituted by chromosome territories, revealed experimentally using the FISH tecnique, that probes chromosome position in the nucleus with fluorescence microscopy [83, 95] (see figure 6.6).

Each chromosome occupies a distinct region in the nucleus, with average diameter of $\sim 1\mu m$. Their position inside the nucleus is non-random, with gene rich chromosomes systematically closer to the nuclear center and gene poor chromosomes closer instead to the nucleus periphery [118]. This suggests a correlation between the location of loci in the nucleus and the gene expression.

Boundaries of the territories are not sharp but there is a chromosome intermingling which allows the formation of long-range interactions, depending on the physical proximity of the chromosomes [119]. Thus, sequences at regions of intermingling are more likely to interact than those within territories. Moreover, some of the long-range inter-chromosomal interactions are found to be maintained also after the cellular division, in new-born cells. This feature suggests the presence of some molecular mechanism that establishes and maintains the chromosomes' positions [120].

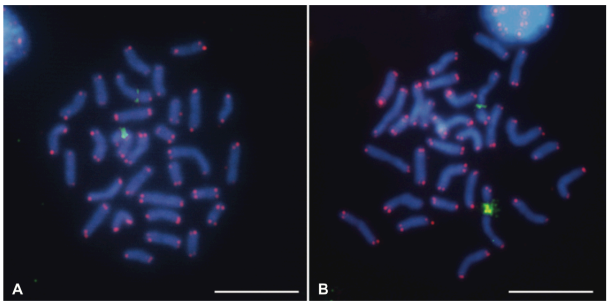


FIGURE 6.7: Figure from [121]. Mitotic chromosomes of *H.magnipapillata* captured with FISH.

6.2 Chromatin mitotic organization

During the mitotic phase chromosome are organized into highly condensed structures, well separated one from the other. At the nuclear scale, experimental tecniques such as light microscopy, electron microscopy, tomography, and mechanical measurements were able to capture images of these compact structures (see for example figure 6.7). The experimental observations suggest that this compaction is achieved via formation of loops along chromosomes [103, 121–125].

The use of the Hi-C tecnique allows us to have information on the mitotic structure on a smaller scale. Hi-C intra-chromosomal maps for mitotic chromosomes show that physical contacts between chromatin sequences are localised along the map diagonal. An evident sign of the linear ordered chromosome structure. See for example the interesting comparison between the same region of a human chromosome in the interphase and mitotic phase in figure 6.8. Finally, also the study of Hi-C data supported the picture according which mitotic chromosomes are made a series of closely packed chromatin loops [103].

6.3 Folding processes

6.3.1 Loop extrusion model

The loop extrusion folding process is the mechanism accredited for the formation of TADs in interphase chromosomes and for the compaction of mitotic chromosomes into loops (with different ingredients and molecular players).

According to the loop extrusion model, the folding of chromatin fibers proceeds like this (figure 6.9): there are active motors, the so-called loop extrusion factors, that attach to

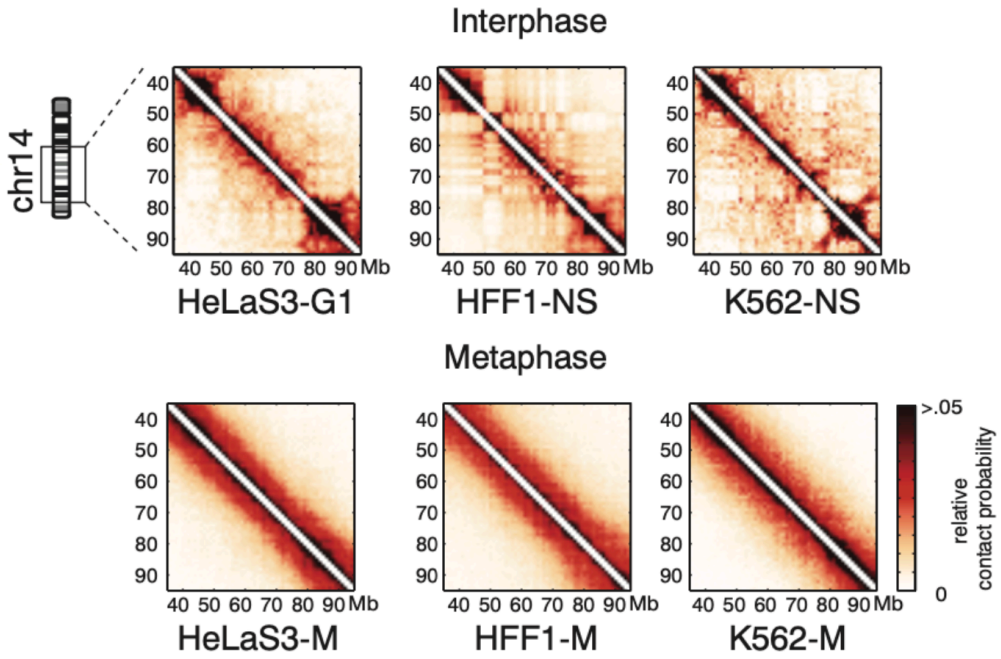


FIGURE 6.8: Figure from [103]. Hi-C maps for a region of human chromosome 14 during interphase and mitotic phase.

the chromatin fiber and start progressively enlarging a DNA loop bidirectionally (two-sided extrusion), until they either fall off, bump into each other, or bump into extrusion barriers.

Various groups proposed models based on loop extrusion, in which coarse-grained molecular dynamics simulations are used to study chromosome folding [98–102]. Among these, the Mirny group proposed a characterization of the loop estrusion process in terms of the ratio of two relevant length scales of the system [102]. These are the loop extrusion factor (LEF) processivity, λ, i.e the average size of a loop bounded by LEFs, and the average separation between two extruders, d (see figure 6.10). The ratio λ/d defines two regimes:

- For $\lambda/d << 1$ the LEFs are sparse, the loops are formed by single LEFs and separated by gaps (figure 6.10a). In this *sparse regime* the chromosome is in a poorly compacted state.

- For $\lambda/d >> 1$ there is instead a *dense regime* where the chromosome is compacted into an array of consecutive loops, each having multiple LEFs at its base (figure 6.10b).

Polymer simulations in the sparse regime, performed with LEFs and with the introduction of extrusion barriers that define the loop sizes, are able to reproduce the organization into TADs in interphase chromosomes (figure 6.9) [98]. The proposed molecular candidates for

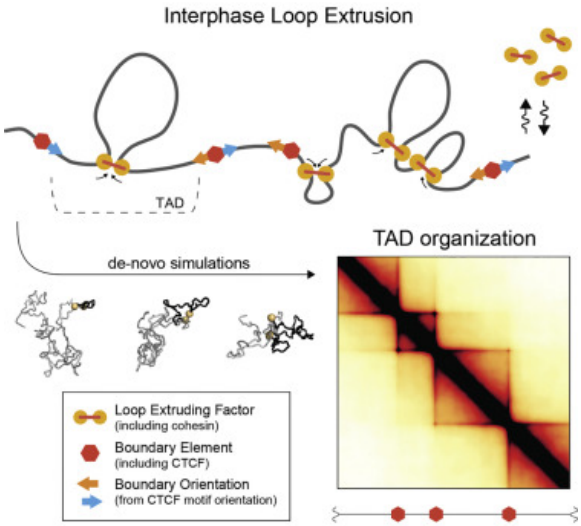


FIGURE 6.9: Figure from [98]. Sketch of the loop extrusion mechanism in interphase and simulated intra-chromosomal contact map obtained via loop extrusion. The simulated maps obtained through loop extrusion reproduce chromosome interphase organization into TADs.

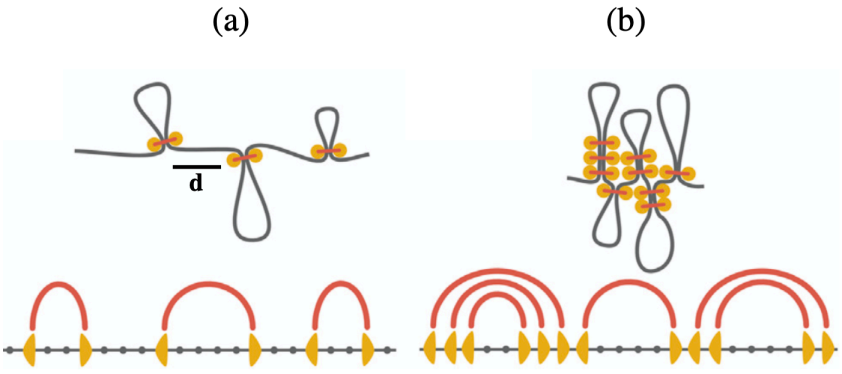


FIGURE 6.10: Figure adapted from [102]. Sketch of the loop extrusion sparse (a) and dense regime (b).

the extrusion factors are cohesins, and at the boundaries are supposed to be CTCF-binding sites linked to the relative proteine, and dispose in convergent orientation. The idea is that they oppose loop extrusion unidirectionally.

Polymer simulations in the dense regime reproduce instead a linear arrangement for chromosomes, with the formation of bottle-brush polymer conformations [102, 103]. These are the characteristic features of the chromosome mitotic shape. Experimental evidences suggest that the extruding motors that mediates the formation of mitotic chromosomes are condensin proteins [122, 126, 127]. Indeed, condensins can generate compaction without crosslinking of topologically distinct chromosomes, binding to two nearby points and then slide to generate a progressively larger loop [128]. This loop extrusion process creates an array of consecutive loops in individual chromosomes.

Finally, note that the initially proposed bidirectional mechanism of loop extrusion has been recently extended to include the combined action of two and one-sided extruders, to take into account experimental evidence that in vitro the yeast condensin can extrude loops asymmetrically, and the possibility that two extruding motors can surpass each other to form Z-loops [127].

6.3.2 Phase separation and block copolymer model

The process of folding accredited for compartment organization in interphase is the phase separation process, modeled with the so-called block copolymer model [21, 99, 104].

In block copolymer models chromatin, represented as a bead and spring chain, is divided into A and B blocks made of consecutive monomers that have the same local properties (monomer size and fiber flexibility) but interact differently. The different interaction between the blocks leads to the spatial segregation of A and B-type chromatin, reproducing then the compartment organization.

An example of block copolymer model is given in ref. [99]. Mirny et al positioned A and B blocks randomly and chose the sizes of blocks to yield the autocorrelation length of the compartment profile inferred from experimental Hi-C data. Then the spatial segregation of A and B-type chromatin was induced by a weak B-B attraction, that constitutes the "compartmental interaction".

They folded the system via molecular dynamics simulation and found that the model is able to reproduce compartments seen from Hi-C maps.

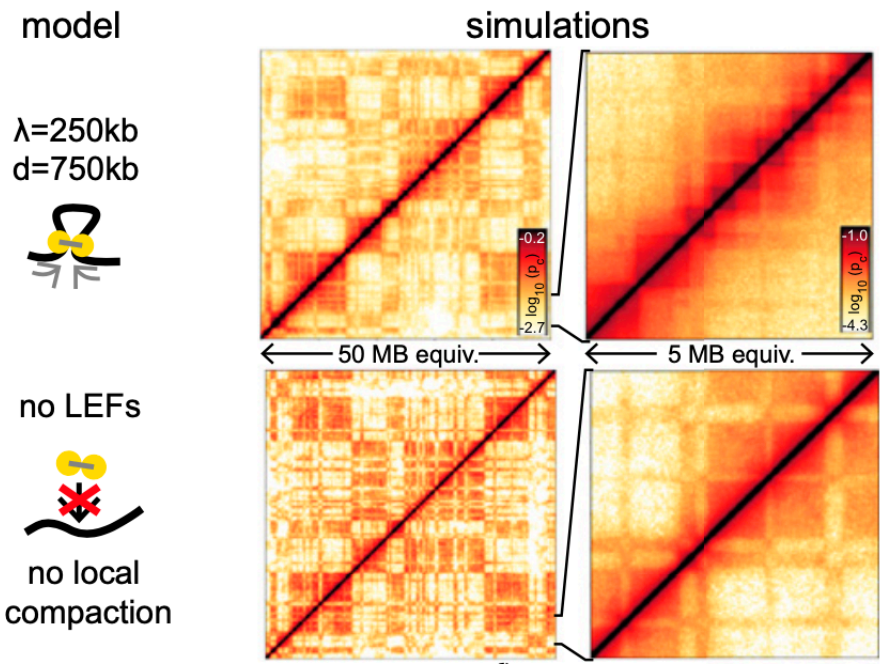


FIGURE 6.11: Figure from [99]. Loop extrusion factors (LEFs) removal leads to stronger and fragmented compartmentalization and loss of TADs.

They also used the phase separation model in conjunction with loop extrusion and found that loop extrusion interferes with compartmental segregation. Indeed, the simulations showed that depletion of loop extrusion factors reduced TADs and revealed finer compartments, while the increase of the number of loop extrusion factors strengthened large TADs and reduced compartmentalization (figure 6.11).

6.3.3 Chromosome memory and conservation of local topological states

Rosa and Everaers [96] proposed a model for chromosome territories based on the maintenance of the local topological state of chromosomes from the mitotic phase to the interphase [97]. Indeed, at the beginning of interphase each chromosome evolves from its initial mitotic conformation and starts swelling inside the nucleus. Rosa and Everaers argued that the relaxation time to reach the complete mixing of all chromosomes, starting from the fully unmixed state, exceeds 100 years for mammalians. Since the typical duration of a cell cycle is about 1 day, it is clear that chromosomes never fully relax. As a consequence, the spatial structures of chromosomes remain effectively stuck into territorial-like conformations, retaining the topological "memory" of the initial mitotic state. This mechanism explains then the chromatin organization into territories.

This was demonstrated doing a modeling of entire chromosomes of different species (and so of different lengths): human, Drosophila and yeast. In each case, they used a bead and spring model to simulate the equilibration of more copies of the same chromosome, that were inside simulation boxes that ensure a density equal to that of the chromatin in the nucleus. The studied systems consist of semi-diluite solutions of linear chains, subject to topological entanglement: the chains cannot cross each other and their movements are severely restricted by the topological constraints. The effect of these constraints is to slow the chain dynamics beyond a density dependent entanglement length, l_e, and a characteristic entanglement time, τ_e [96, 130], whose values depend on chain stiffness and on the contour length density of the polymer solution.

The initial mitotic state was reproduced with linear or ring-shaped helical structures and the system was then equilibrated via molecular dynamics simulations.

The length of the simulations covered the longest cell-cycle life-time among the considered species (i.e. the one for the longest chromosome, the human one). This time is actually sufficient to mix and equilibrate the shorter yeast chromosomes, and indeed in this case they found that apparently there was no territorial organization, while the much longer Drosophila and human chromosomes remained confined to distinct territories (figure 6.12).

Moreover, their model was able to reproduce the experimental scaling behavior of the contact probability as a function of the genomic distance, $P(\ell)$, in interphase chromosomes.

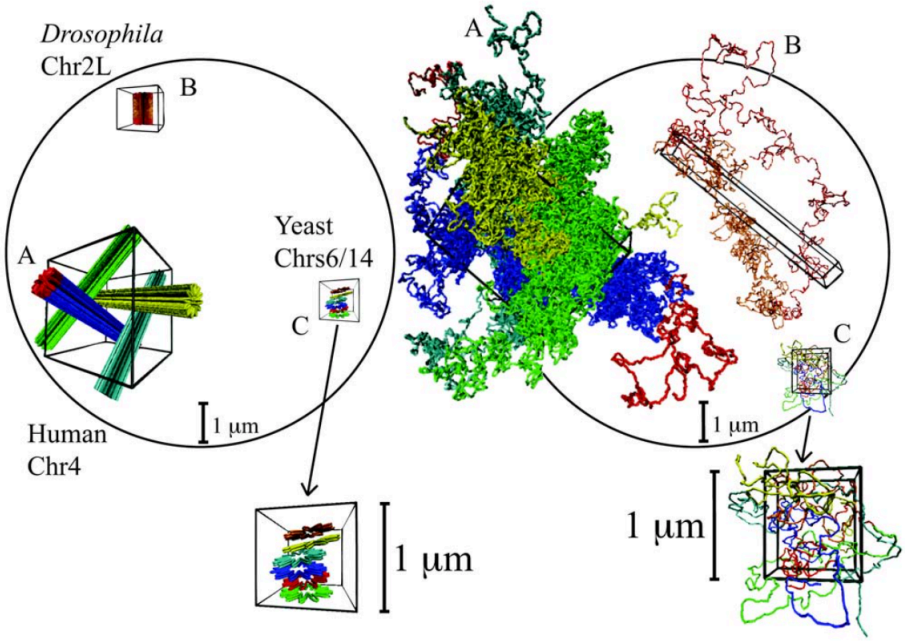


FIGURE 6.12: Figure from [96]. Initial and final configurations of human Chr4 (a), of Drosophila Chr2L (b) and of yeast Chr6 and Chr14 (c) shown together with the spherical nucleus (black circle) of 10 mm in diameter and the corresponding simulation boxes.

Ongoing work

In this chapter I study a semi-diluite solution of chromatin fibers, modeled as coarse-grained semiflexible chains. In particular I consider more copies of the same human chromosome, the chromosome 19, first studying the folding of this system in the mitotic conformation. Our aim is to simulate a model that reproduces the mitotic arrangement and contacts, using loop extrusion. Once obtained a mitotic conformation that fits the experimental evidences, we are interested in study on which degree the existent chromatin physical contacts are predictive of the interphase structure, letting the system folding using the mechanisms illustrated in chapter 6, with different time orderings. The comparison of the final simulated chromatin structures would allow us to understand which of the explored time orderings of folding mechanisms better reproduce the actual chromatin organization, in comparison to the experimental findings. We surmise that this would allow us to speculate on what happens during the folding process in vivo.

7.1 Model and methods

We simulated six copies of human chromosome 19 of length $L_c = 59.13Mpb$ as semi-flexible chains of N_{beads} with thickness $\sigma = 30nm$ and persistence length $l_p = 5\sigma$ [96, 129]. Accordingly with the mapping in ref. [96], each bead spans $3000bp$ and, hence, to account for the total contour length $L_c = 59.13Mpb$ of human chromosome 19, the number of beads N_{beads} was set equal to 19710.

The six copies are placed in a random, but non-overlapping manner inside a cubic simulation box such that the nuclear density is $\rho = 0.012bp/nm^3$, which matches the typical genomic density in human cells [96]. We prepared the six chain copies in a mitotic-like arrangement, as shown in figure 7.1. In particular, each chain was prepared in an elongated

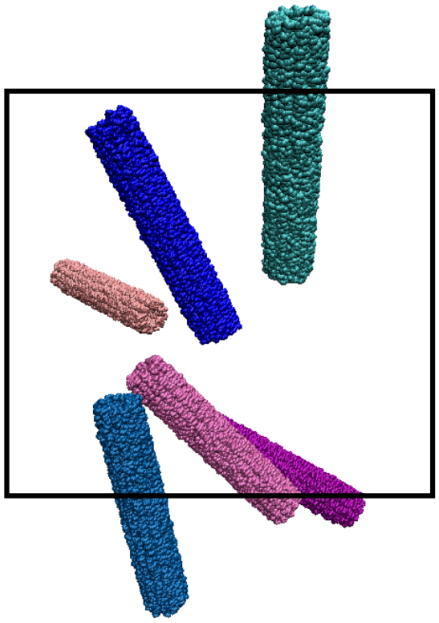


FIGURE 7.1: Initial mitotic-like arrangement, made of 6 copies of model human chromosome 19.

solenoidal-like configuration according to the procedure introduced by Rosa et al in [96], with loops of $\sim 50kpb$ each and departing radially from a central axis.

The system is studied with Langevin molecular dynamics simulations, with chain potential energy given by equation 3.8 and the dynamical evolution of each chain following equation 3.14. The equation of motion is integrated numerically using the LAMMPS simulation package [70, 71]. The integration timestep was set to $dt = 0.012\tau_{LJ}$.

Finally, the obtained semi-diluite solution of linear polymers is characterized by an entanglement length $l_e = 40\sigma$ and entanglement time $\tau_e = 1600\tau_{LJ}$, having adopted the choice proposed in refs [96, 130] (see also section 6.3.3).

7.1.1 Extrusion dynamics

Each initial conformation in figure 7.1 is subjected to a dynamics of loop extrusion with a number of extruding factors $N_{EF} = 100$, in order to create 100 intra-chromosomal loops in each of the 6 chromosome chains. This means that there are N_{EF} input target loops, disposed in random positions along the chains. The length of the loops is chosen equal to $l_{loop} = N_{beads}/N_{EF}$ for three of the copies (colored in pink tones in figure 7.1), while is randomly defined for the others (colored in blue tones in figure 7.1).

The loop extrusion dynamics is bidirectional: at each extrusion step the $loop_n$ (i, j) with $i < j$ gains one bead backward of the first bead $(i \to i-1)$ and one bead forward of the second bead $(j \to j+1)$ in the loop. The extrusion becomes unidirectional when one of the target loop ends is reached, either forward or backward. Finally, the extrusion is stopped when the target loop is formed at both ends. We chose to extrude a pair of beads every $0.01\tau_e$. The total duration of the extrusion process is given by the number of bead pairs of the longest loop multiplied for the loop extrusion simulation timestep, corresponding to $0.01\tau_e$.

Before being able to run the simulations, we largely tested the algorithm implemented for the extrusion dynamics.

7.2 Preliminary results

In this section I describe the first preliminary results of the study.

We started from the six copies of chromosome 19 in the mitotic conformation in figure 7.1. To remove any excessive intra-chain strain of the orderly designed mitotic arrangement, the model chromosomes of figure 7.1 were briefly evolved with an unbiased MD protocol of 10^5 MD timesteps. The resulting relaxed mitotic configuration is shown in figure 7.2a. This mitotic arrangement was further evolved for a much longer simulation time, corresponding to 10^8 MD timesteps, to obtain the decondensed arrangement shown in figure 7.2c. I will refer to the two different configurations as *condensed* or *decondensed*. We did this because we are interested in seeing the effect of the initial relaxation on the final structure, once simulated the effect of loop extrusion and/or phase separation, in different time ordering.

As a first thing we simulated the effect of loop extrusion on the two initial configurations. In both cases, we used 100 extruders and the final extruded configurations are shown in figure 7.2b and d, respectively for the condensed and decondensed initial configurations. The three chromosomes in blue tones correspond to the set with random loop length, while the three ones in pink tones to the set with constant loop length.

We checked the validity of our protocol for the loop extrusion by measuring the final cartesian distance between the pairs of target beads at the base of the loops. As we can see from figure 7.3 this distance is progressively reduced, until every pair of target beads reach

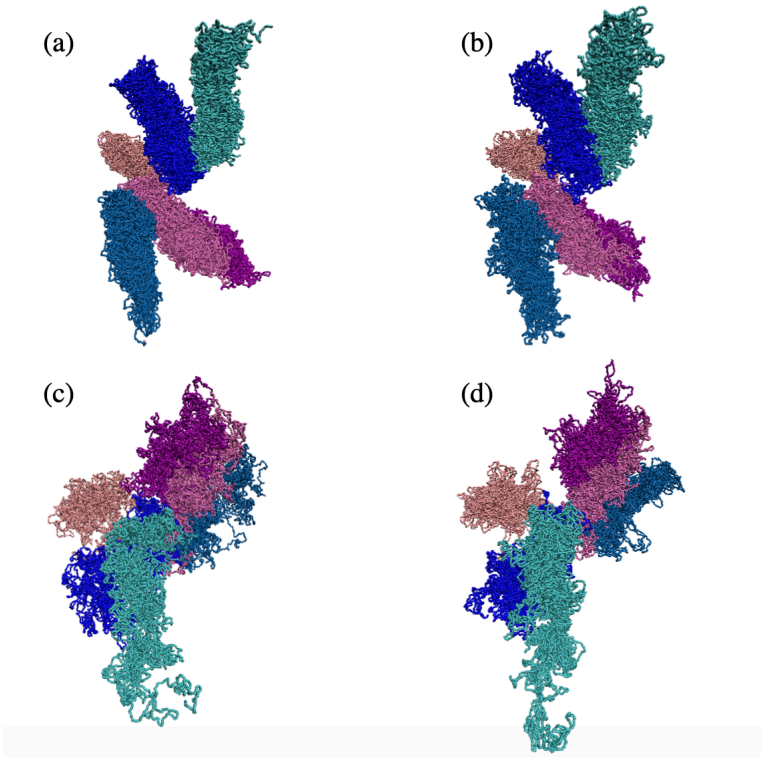


FIGURE 7.2: The system in the starting and final extruded configurations. (a) Condensed starting configuration, (b) condensed final configuration, (c) decondensed starting configuration, (d) decondensed final configuration.

a distance smaller than $2\sigma \sim 60nm$ (this is the chosen cutoff distance between target beads for a fully extruded loop).

7.2.1 Gyration radius

We analyzed then the physical properties of the obtained extruded chromosomes starting by measuring the average gyration radius for the two sets of chromosome copies, with random loop length and constant loop length, along the extrusion process. The results are provided in figure 7.4 for the condensed and decondensed initial cases (panel a and b). First, at fixed chromosome set, the gyration radius is larger in the decondensed case. This is due to the different compaction of the initial configurations. Second, it is interesting to see what do the chromosome structures behave under loop extrusion. Do they get compact or expanded? From the results in figure 7.4, we see that it depends on the initial relaxation of the designed mitotic arrangement. Indeed, the system has an initial expansion in case a that then is progressively reduced, while in case b the gyration radius has a decreasing trend. We will further investigate this aspect in the continuation of the work. Finally, we observe that the

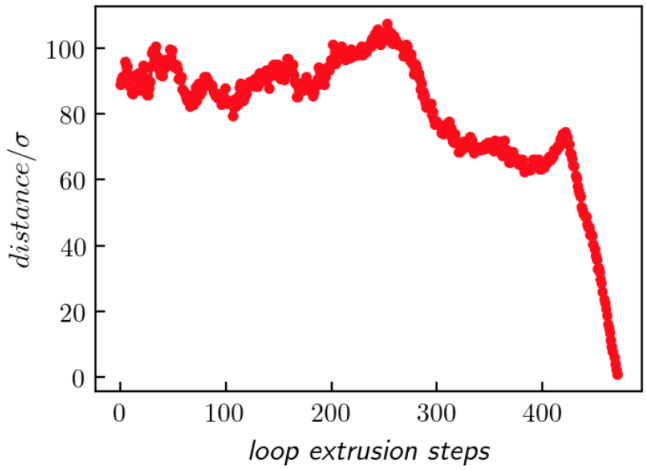


FIGURE 7.3: Example of distance between two target beads at the base of a loop as a function of the extrusion timesteps.

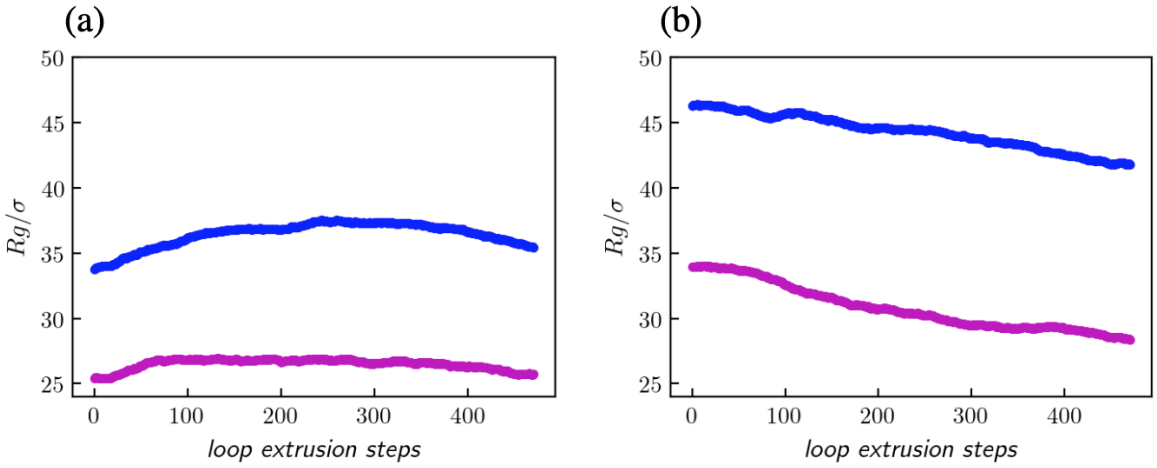


FIGURE 7.4: Average gyration radius for the two sets of chromosome copies, with random loop length (blue) and constant loop length (purple), (a) for the condensed and (b) decondensed initial configurations.

trend of the gyration radius is not affected by the different loop length choice for the two sets (random chosen loop length for the set in blue and constant loop length for the set in purple in figure 7.4).

7.2.2 Contact matrix

Next, we built the simulated intra-chromosomal contact maps. We measured the physical contacts among intervals of length $300kbp$. According to our model, using this resolution corresponds to probe chromosome intervals of 100 beads, with 197 intervals spanning the whole chromosome.

Two intervals of $300kbp$ are defined to be in contact if their cartesian distance is larger than the threshold value 5σ, that corresponds to 5 chain beads. We finally obtained the average number of contacts by averaging over the chromosome copies, separately considering the set of three chromosomes with random or constant loop length.

In figure 7.5 there are the contact maps at $300kbp$ resolution, after extrusion from the condensed initial case (panel a and b, respectively for the set of chromosome copies with random and constant loop length) and the decondensed initial case (panel c and d). The maps show that the physical contacts are mostly limited along the map diagonal. This is typical of the linear organization of the mitotic chromosomes [103, 122]. Finally, we do not report appreciable differences between the cases of random and constant loop length.

7.3 Future work and perspectives

I presented our ongoing study of the modeling of human chromosome 19 via coarse grained molecular dynamics simulations.

We studied the effect of the loop extrusion process on mitotic-like initial conformations, and we found that the intra-chromosomal contact maps show that the physical contacts among chromosome sequences are mostly localised at the matrix diagonal. Thus, we were able to reproduce the mitotic-like chromosome arrangement.

The next step of this study consists in the introduction of the phase separation folding process, using a block-copolymer model in order to induce a compartmental interaction. We will do this starting at different levels of the simulation: first, from the initial condensed and decondensed mitotic structures and second, from the final extruded configurations in figure

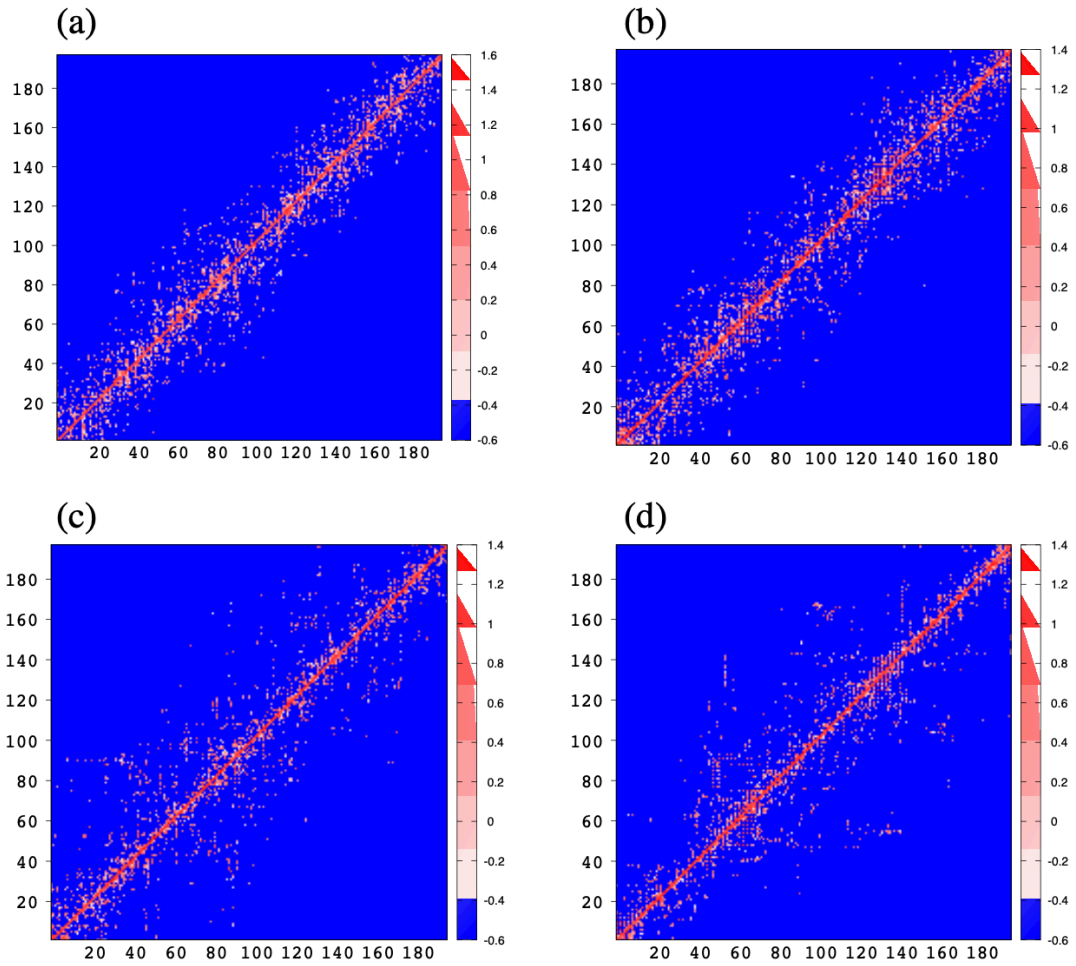


FIGURE 7.5: Simulated contact maps after loop extrusion, at resolution 300 kb. (a) Contact map for the set of three chromosomes with random loop length, having an initial condensed configuration; (b) the same for the set of three chromosomes with constant loop length. (c) Contact map for the set of three chromosomes with random loop length, having an initial decondensed configuration; (d) the same for the set of three chromosomes with constant loop length.

7.2b and d. These two protocols will allow us to explore the effect of different combinations of folding processes. Respectively: first, only phase separation process, and second, phase separation acting after loop extrusion.

The long term goal is to understand which combination of folding processes would allow us to better reproduce the experimental chromatin structural features, studying also to which extent the formation of the known chromatin interphase structures is pre-conditioned by the pre-existent mitotic physical contacts.

Bibliography

[1] Ritort, F. Single-molecule experiments in biological physics: methods and applications. *Journal of Physics: Condensed Matter* **18**, R531 (2006).

[2] Binnig, G., Quate, C. F. & Gerber, C. Atomic force microscope. *Physical review letters* **56**, 930 (1986).

[3] Moreno-Herrero, F., Colchero, J., Gomez-Herrero, J. & Baró, A. Atomic force microscopy contact, tapping, and jumping modes for imaging biological samples in liquids. *Physical Review E* **69**, 031915 (2004).

[4] Bustamante, C., Rivetti, C. & Keller, D. J. Scanning force microscopy under aqueous solutions. *Current opinion in structural biology* **7**, 709–716 (1997).

[5] Kellermayer, M. S. Visualizing and manipulating individual protein molecules. *Physiological measurement* **26**, R119 (2005).

[6] Ashkin, A. Optical trapping and manipulation of neutral particles using lasers. *Proceedings of the National Academy of Sciences* **94**, 4853–4860 (1997).

[7] Bryant, Z. *et al.* Structural transitions and elasticity from torque measurements on dna. *Nature* **424**, 338–341 (2003).

[8] Merkel, R., Nassoy, P., Leung, A., Ritchie, K. & Evans, E. Energy landscapes of receptor–ligand bonds explored with dynamic force spectroscopy. *Nature* **397**, 50–53 (1999).

[9] Gordon, M. P., Ha, T. & Selvin, P. R. Single-molecule high-resolution imaging with photobleaching. *Proceedings of the National Academy of Sciences* **101**, 6462–6465 (2004).

[10] Forkey, J. N., Quinlan, M. E., Shaw, M. A., Corrie, J. E. & Goldman, Y. E. Three-dimensional structural dynamics of myosin v by single-molecule fluorescence polarization. *Nature* **422**, 399–404 (2003).

[11] Yildiz, A. *et al.* Myosin v walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *science* **300**, 2061–2065 (2003).

[12] Krog, J. *et al.* Stochastic unfolding of nanoconfined dna: Experiments, model and bayesian analysis. *The Journal of chemical physics* **149**, 215101 (2018).

[13] Khorshid, A. *et al.* Dynamic compression of single nanochannel confined dna via a nanodozer assay. *Physical review letters* **113**, 268104 (2014).

[14] Smith, S. B., Finzi, L. & Bustamante, C. Direct mechanical measurements of the elasticity of single dna molecules by using magnetic beads. *Science* **258**, 1122–1126 (1992).

[15] Rubinstein, M., Colby, R. H. *et al. Polymer physics*, vol. 23 (Oxford university press New York, 2003).

[16] Huang, L. R. *et al.* A dna prism for high-speed continuous fractionation of large dna molecules. *Nature Biotechnology* **20**, 1048–1051 (2002).

[17] Tegenfeldt, J. O. *et al.* The dynamics of genomic-length dna molecules in 100-nm channels. *Proceedings of the National Academy of Sciences* **101**, 10979–10983 (2004).

[18] Reisner, W. *et al.* Statics and dynamics of single dna molecules confined in nanochannels. *Physical Review Letters* **94**, 196101 (2005).

[19] Reisner, W., Pedersen, J. N. & Austin, R. H. Dna confinement in nanochannels: physics and biological applications. *Reports on Progress in Physics* **75**, 106601 (2012).

[20] White, M. A., Eykelenboom, J. K., Lopez-Vernaza, M. A., Wilson, E. & Leach, D. R. Non-random segregation of sister chromosomes in escherichia coli. *Nature* **455**, 1248–1250 (2008).

[21] Ghosh, S. K. & Jost, D. How epigenome drives chromatin folding and dynamics, insights from efficient coarse-grained models of chromosomes. *PLoS computational biology* **14**, e1006159 (2018).

[22] Niewieczerzal, S., Niemyska, W. & Sulkowska, J. I. Defining and detecting links in chromosomes. *Scientific reports* **9**, 1–10 (2019).

[23] Marko, J. F. Linking topology of tethered polymer rings with applications to chromosome segregation and estimation of the knotting length. *Physical Review E* **79**, 051905 (2009).

[24] Liu, Z., Deibler, R. W., Chan, H. S. & Zechiedrich, L. The why and how of dna unlinking. *Nucleic acids research* **37**, 661–671 (2009).

[25] Liu, Z., Zechiedrich, L. & Chan, H. S. Action at hooked or twisted–hooked dna juxtapositions rationalizes unlinking preference of type-2 topoisomerases. *Journal of molecular biology* **400**, 963–982 (2010).

[26] Micheletti, C., Marenduzzo, D. & Orlandini, E. Polymers with spatial or topological constraints: Theoretical and computational results. *Physics Reports* **504**, 1–73 (2011).

[27] Marenduzzo, D., Micheletti, C. & Orlandini, E. Biopolymer organization upon confinement. *Journal of Physics: Condensed Matter* **22**, 283102 (2010).

[28] Dai, L., Renner, C. B. & Doyle, P. S. The polymer physics of single dna confined in nanochannels. *Advances in colloid and interface science* **232**, 80–100 (2016).

[29] Whittington, S. G. & Soteros, C. E. Polymers in slabs, slits, and pores. *Israel journal of chemistry* **31**, 127–133 (1991).

[30] Micheletti, C. & Orlandini, E. Knotting and metric scaling properties of dna confined in nano-channels: a monte carlo study. *Soft Matter* **8**, 10959–10968 (2012).

[31] Dai, L., van der Maarel, J. R. & Doyle, P. S. Effect of nanoslit confinement on the knotting probability of circular dna. *ACS Macro Letters* **1**, 732–736 (2012).

[32] Poier, P., Likos, C. N. & Matthews, R. Influence of rigidity and knot complexity on the knotting of confined polymers. *Macromolecules* **47**, 3394–3400 (2014).

[33] Beaton, N. R., Eng, J. W., Ishihara, K., Shimokawa, K. & Soteros, C. E. Characterising knotting properties of polymers in nanochannels. *Soft matter* **14**, 5775–5785 (2018).

[34] Micheletti, C. & Orlandini, E. Knotting and unknotting dynamics of dna strands in nanochannels. *ACS Macro Letters* **3**, 876–880 (2014).

[35] Baiesi, M. & Orlandini, E. Universal properties of knotted polymer rings. *Physical Review E* **86**, 031805 (2012).

[36] Amin, S., Khorshid, A., Zeng, L., Zimny, P. & Reisner, W. A nanofluidic knot factory based on compression of single dna in nanochannels. *Nature communications* **9**, 1–10 (2018).

[37] Klotz, A. R., Soh, B. W. & Doyle, P. S. Motion of knots in dna stretched by elongational fields. *Physical review letters* **120**, 188003 (2018).

[38] Wasserman, S. A. & Cozzarelli, N. R. Biochemical topology: applications to dna recombination and replication. *Science* **232**, 951–960 (1986).

[39] Sumners, D. W. Untangling dna. *The Mathematical Intelligencer* **12**, 71–80 (1990).

[40] Ashley, C. & Lee, J. S. A triplex-mediated knot between separated polypurine–polypyrimidine tracts in circular dna blocks transcription by escherichia coli rna polymerase. *DNA and cell biology* **19**, 235–241 (2000).

[41] Portugal, J. & Rodríguez-Campos, A. T7 rna polymerase cannot transcribe through a highly knotted dna template. *Nucleic acids research* **24**, 4890–4894 (1996).

[42] Deibler, R. W., Mann, J. K., De Witt, L. S. & Zechiedrich, L. Hin-mediated dna knotting and recombining promote replicon dysfunction and mutation. *BMC molecular biology* **8**, 44 (2007).

[43] Virnau, P., Mirny, L. A. & Kardar, M. Intricate knots in proteins: Function and evolution. *PLoS computational biology* **2** (2006).

[44] Virnau, P., Mallam, A. & Jackson, S. Structures and folding pathways of topologically knotted proteins. *Journal of Physics: Condensed Matter* **23**, 033101 (2010).

[45] Potestio, R., Micheletti, C. & Orland, H. Knotted vs. unknotted proteins: evidence of knot-promoting loops. *PLoS computational biology* **6** (2010).

[46] Lou, S.-C. *et al.* The knotted protein uch-l1 exhibits partially unfolded forms under native conditions that share common structural features with its kinetic folding intermediates. *Journal of molecular biology* **428**, 2507–2520 (2016).

[47] Wu, Z.-S., Shen, Z., Tram, K. & Li, Y. Engineering interlocking dna rings with weak physical interactions. *Nature Communications* **5**, 1–10 (2014).

[48] Adams, C. C. *The knot book* (American Mathematical Soc., 1994).

[49] Rolfsen, D. *Knots and links*, vol. 346 (American Mathematical Soc., 2003).

[50] Alexander, J. W. Topological invariants of knots and links. *Transactions of the American Mathematical Society* **30**, 275–306 (1928).

[51] Archibald, J. *The multivariable Alexander polynomial on tangles.* Ph.D. thesis, University of Toronto (2010).

[52] Caraglio, M., Micheletti, C. & Orlandini, E. Physical links: Defining and detecting inter-chain entanglement. *Scientific reports* **7**, 1–10 (2017).

[53] Amici, G., Caraglio, M., Orlandini, E. & Micheletti, C. Topologically linked chains in confinement. *ACS Macro Letters* **8**, 442–446 (2019).

[54] De Gennes, P.-G. & Gennes, P.-G. *Scaling concepts in polymer physics* (Cornell university press, 1979).

[55] Daoud, M. & De Gennes, P. Statistics of macromolecular solutions trapped in small pores. *Journal de physique* **38**, 85–93 (1977).

[56] Odijk, T. Scaling theory of dna confined in nanochannels and nanoslits. *Physical Review E* **77**, 060901 (2008).

[57] Wang, Y., Tree, D. R. & Dorfman, K. D. Simulation of dna extension in nanochannels. *Macromolecules* **44**, 6594–6604 (2011).

[58] Werner, E., Cheong, G., Gupta, D., Dorfman, K. & Mehlig, B. One-parameter scaling theory for dna extension in a nanochannel. *Physical review letters* **119**, 268102 (2017).

[59] Brochard-Wyart, F., Tanaka, T., Borghi, N. & De Gennes, P.-G. Semiflexible polymers confined in soft tubes. *Langmuir* **21**, 4144–4148 (2005).

[60] Jung, Y. *et al.* Ring polymers as model bacterial chromosomes: confinement, chain topology, single chain statistics, and how they interact. *Soft Matter* **8**, 2095–2102 (2012).

[61] Benkova, Z. & Cifra, P. Simulation of semiflexible cyclic and linear chains moderately and strongly confined in nanochannels. *Macromolecules* **45**, 2597–2608 (2012).

[62] Benkova, Z. & Cifra, P. Comparison of linear and ring dna macromolecules moderately and strongly confined in nanochannels (2013).

[63] Alizadehheidari, M. *et al.* Nanoconfined circular and linear dna: Equilibrium conformations and unfolding kinetics. *Macromolecules* **48**, 871–878 (2015).

[64] Jun, S. & Wright, A. Entropy as the driver of chromosome segregation. *Nature Reviews Microbiology* **8**, 600–607 (2010).

[65] Jung, Y., Kim, J., Jun, S. & Ha, B.-Y. Intrachain ordering and segregation of polymers under confinement. *Macromolecules* **45**, 3256–3262 (2012).

[66] Minina, E. & Arnold, A. Entropic segregation of ring polymers in cylindrical confinement. *Macromolecules* **48**, 4998–5005 (2015).

[67] Cook, P. R. & Marenduzzo, D. Entropic organization of interphase chromosomes. *Journal of Cell Biology* **186**, 825–834 (2009).

[68] Marenduzzo, D. & Orlandini, E. Topological and entropic repulsion in biopolymers. *Journal of statistical mechanics: theory and experiment* **2009**, L09002 (2009).

[69] Arnold, A. & Jun, S. Time scale of entropic segregation of flexible polymers in confinement: implications for chromosome segregation in filamentous bacteria. *Physical Review E* **76**, 031901 (2007).

[70] Kremer, K. & Grest, G. S. Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *The Journal of Chemical Physics* **92**, 5057–5086 (1990).

[71] Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. Tech. Rep., Sandia National Labs., Albuquerque, NM (United States) (1993).

[72] Tubiana, L., Orlandini, E. & Micheletti, C. Probing the entanglement and locating knots in ring polymers: A comparative study of different arc closure schemes. *Progress of Theoretical Physics Supplement* **191**, 192–204 (2011).

[73] Virnau, P., Kantor, Y. & Kardar, M. Knots in globule and coil phases of a model polyethylene. *Journal of the American Chemical Society* **127**, 15102–15106 (2005).

[74] Grosberg, A. Y. & Rabin, Y. Metastable tight knots in a wormlike polymer. *Physical review letters* **99**, 217801 (2007).

[75] Tubiana, L., Rosa, A., Fragiacomo, F. & Micheletti, C. Spontaneous knotting and unknotting of flexible linear polymers: Equilibrium and kinetic aspects. *Macromolecules* **46**, 3669–3678 (2013).

[76] Krajina, B. A., Zhu, A., Heilshorn, S. C. & Spakowitz, A. J. Active dna olympic hydrogels driven by topoisomerase activity. *Physical review letters* **121**, 148001 (2018).

[77] Liu, K. *et al.* Detecting topological variations of dna at single-molecule level. *Nature communications* **10**, 1–9 (2019).

[78] Milchev, A. Single-polymer dynamics under constraints: scaling theory and computer experiment. *Journal of Physics: Condensed Matter* **23**, 103101 (2011).

[79] Chen, J., Rauch, C. A., White, J. H., Englund, P. T. & Cozzarelli, N. R. The topology of the kinetoplast dna network. *Cell* **80**, 61–69 (1995).

[80] Baiesi, M., Orlandini, E., Trovato, A. & Seno, F. Linking in domain-swapped protein dimers. *Scientific reports* **6**, 1–11 (2016).

[81] Branco, M. R. & Pombo, A. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol* **4**, e138 (2006).

[82] Nelson, D. L., Lehninger, A. L. & Cox, M. M. *Lehninger principles of biochemistry* (Macmillan, 2008).

[83] Cremer, T. & Cremer, C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature reviews genetics* **2**, 292–301 (2001).

[84] Lieberman-Aiden, E. *et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *science* **326**, 289–293 (2009).

[85] Nishino, Y. *et al.* Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. *The EMBO journal* **31**, 1644–1653 (2012).

[86] Beagrie, R. A. *et al.* Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* **543**, 519–524 (2017).

[87] Dekker, J., Rippe, K., Dekker, M. & Kleckner, N. Capturing chromosome conformation. *science* **295**, 1306–1311 (2002).

[88] Boettiger, A. N. *et al.* Super-resolution imaging reveals distinct chromatin folding for different epigenetic states. *Nature* **529**, 418–422 (2016).

[89] Ou, H. D. *et al.* Chromemt: Visualizing 3d chromatin structure and compaction in interphase and mitotic cells. *Science* **357** (2017).

[90] Bintu, B. *et al.* Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* **362** (2018).

[91] Nir, G. *et al.* Walking along chromosomes with super-resolution imaging, contact maps, and integrative modeling. *PLoS genetics* **14**, e1007872 (2018).

[92] Wang, S. *et al.* Spatial organization of chromatin domains and compartments in single chromosomes. *Science* **353**, 598–602 (2016).

[93] Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).

[94] Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the x-inactivation centre. *Nature* **485**, 381–385 (2012).

[95] Cremer, T. & Cremer, M. Chromosome territories. *Cold Spring Harbor perspectives in biology* **2**, a003889 (2010).

[96] Rosa, A. & Everaers, R. Structure and dynamics of interphase chromosomes. *PLoS Comput Biol* **4**, e1000153 (2008).

[97] Halverson, J. D., Smrek, J., Kremer, K. & Grosberg, A. Y. From a melt of rings to chromosome territories: the role of topological constraints in genome folding. *Reports on Progress in Physics* **77**, 022601 (2014).

[98] Fudenberg, G. *et al.* Formation of chromosomal domains by loop extrusion. *Cell reports* **15**, 2038–2049 (2016).

[99] Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. & Mirny, L. A. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proceedings of the National Academy of Sciences* **115**, E6697–E6706 (2018).

[100] Brackley, C. *et al.* Extrusion without a motor: a new take on the loop extrusion model of genome organization. *Nucleus* **9**, 95–103 (2018).

[101] Sanborn, A. *et al.* Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *APS* **2016**, R39–011 (2016).

[102] Goloborodko, A., Marko, J. F. & Mirny, L. A. Chromosome compaction by active loop extrusion. *Biophysical journal* **110**, 2162–2168 (2016).

[103] Naumova, N. *et al.* Organization of the mitotic chromosome. *Science* **342**, 948–953 (2013).

[104] Jost, D., Carrivain, P., Cavalli, G. & Vaillant, C. Modeling epigenome folding: formation and dynamics of topologically associated chromatin domains. *Nucleic acids research* **42**, 9553–9561 (2014).

[105] Marti-Renom, M. A. & Mirny, L. A. Bridging the resolution gap in structural modeling of 3d genome organization. *PLoS computational biology* **7**, e1002125 (2011).

[106] Rao, S. S. *et al.* A 3d map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).

[107] Dixon, J. R. *et al.* Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331–336 (2015).

[108] Lonfat, N., Montavon, T., Darbellay, F., Gitto, S. & Duboule, D. Convergent evolution of complex regulatory landscapes and pleiotropy at hox loci. *Science* **346**, 1004–1006 (2014).

[109] Valton, A.-L. & Dekker, J. Tad disruption as oncogenic driver. *Current opinion in genetics & development* **36**, 34–40 (2016).

[110] Pombo, A. & Dillon, N. Three-dimensional genome architecture: players and mechanisms. *Nature reviews Molecular cell biology* **16**, 245–257 (2015).

[111] Haering, C. H., Farcas, A.-M., Arumugam, P., Metson, J. & Nasmyth, K. The cohesin ring concatenates sister dna molecules. *Nature* **454**, 297–301 (2008).

[112] Stigler, J., Çamdere, G. Ö., Koshland, D. E. & Greene, E. C. Single-molecule imaging reveals a collapsed conformational state for dna-bound cohesin. *Cell reports* **15**, 988–998 (2016).

[113] Davidson, I. F. *et al.* Rapid movement and transcriptional re-localization of human cohesin on dna. *The EMBO journal* **35**, 2671–2685 (2016).

[114] Guo, Y. *et al.* Crispr inversion of ctcf sites alters genome topology and enhancer/promoter function. *Cell* **162**, 900–910 (2015).

[115] de Wit, E. *et al.* Ctcf binding polarity determines chromatin looping. *Molecular cell* **60**, 676–684 (2015).

[116] Narendra, V. *et al.* Ctcf establishes discrete functional chromatin domains at the hox clusters during differentiation. *Science* **347**, 1017–1021 (2015).

[117] Jost, D., Vaillant, C. & Meister, P. Coupling 1d modifications and 3d nuclear organization: data, models and function. *Current opinion in cell biology* **44**, 20–27 (2017).

[118] Bolzer, A. *et al.* Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biol* **3**, e157 (2005).

[119] Dekker, J., Marti-Renom, M. A. & Mirny, L. A. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nature Reviews Genetics* **14**, 390–403 (2013).

[120] Meaburn, K. J. & Misteli, T. Chromosome territories. *Nature* **445**, 379–381 (2007).

[121] Anokhin, B., Hemmrich-Stanisak, G. & Bosch, T. Karyotyping and single-gene detection using fluorescence in situ hybridization on chromosomes of hydra magnipapillata (cnidaria: Hydrozoa). *Comparative Cytogenetics* **4**, 97 (2010).

[122] Paulson, J. R. & Laemmli, U. The structure of histone-depleted metaphase chromosomes. *Cell* **12**, 817–828 (1977).

[123] Maeshima, K. & Laemmli, U. K. A two-step scaffolding model for mitotic chromosome assembly. *Developmental cell* **4**, 467–480 (2003).

[124] Maeshima, K., Eltsov, M. & Laemmli, U. K. Chromosome structure: improved immunolabeling for electron microscopy. *Chromosoma* **114**, 365–375 (2005).

[125] Liang, Z. *et al.* Chromosomes progress to metaphase in multiple discrete steps via global compaction/expansion cycles. *Cell* **161**, 1124–1137 (2015).

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

[127] Ganji, M. *et al.* Real-time imaging of dna loop extrusion by condensin. *Science* **360**, 102–105 (2018).

[128] Nasmyth, K. Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annual review of genetics* **35**, 673–745 (2001).

[129] Di Stefano, M., Rosa, A., Belcastro, V., di Bernardo, D. & Micheletti, C. Colocalization of coregulated genes: a steered molecular dynamics study of human chromosome 19. *PLoS Comput Biol* **9**, e1003019 (2013).

[130] Schram, R. D., Rosa, A. & Everaers, R. Local loop opening in untangled ring polymer melts: a detailed feynman test of models for the large scale structure. *Soft Matter* **15**, 2418–2429 (2019).