MSc Chemistry
Atomic Scale Modelling of Physical, Chemical and Biomolecular Systems (AtoSiM)

Master Thesis

Epigenetics in 4D
The “living chromatin” model

by

Juan David Olarte Plata

10652981

July 2015
30 ECTS
2014-II

Supervisors:
Cédric Vaillant, PhD
Daniel Jost, PhD

Examiners:
Ralf Everaers, PhD
Evert Jan Meijer, PhD

Laboratoire de Physique, ENS de Lyon
# CONTENTS

1. **Introduction** | 1
  1.1 Epigenomic marks: the Ising spins | 1
      1.1.1 Mathematical model (Jost 2014) | 1
      1.1.2 Stability analysis and transition rates | 3
      1.1.3 Epigenomic mark patterns can affect the 3D structure | 3
  1.2 A polymer model for the chromatin (Jost et. al. 2014) | 4
      1.2.1 Block copolymer continuum model | 4
      1.2.2 Chromatin organization | 6

2. **Shifting to a polymer lattice model** | 8
  2.1 Lattice model | 8
      2.1.1 Monte Carlo algorithm | 8
  2.2 Equilibrium properties for the ideal polymer | 10
      2.2.1 End-to-end distance | 10
      2.2.2 Equilibration steps determination | 10
      2.2.3 Scaling of the RMS end-to-end vector | 11
      2.2.4 Radius of gyration | 12
  2.3 Dynamical properties for the ideal polymer: results from the Rouse model | 13
      2.3.1 g1 and g3 | 13
      2.3.2 Time correlation function for the end-to-end distance | 14
      2.3.3 Time correlation function per Rouse mode | 15
  2.4 Considering polymer flexibility | 16
      2.4.1 Equilibration considerations | 16
      2.4.2 Correlation between bond vectors | 16
      2.4.3 Size of the polymer | 18

3. **Considering excluded volume interaction** | 19
  3.1 Equilibration time determination | 20
  3.2 Scaling of the polymer size | 20
      3.2.1 Radius of gyration | 20
  3.3 Dynamical properties | 21
      3.3.1 g1 and g3 | 21
      3.3.2 Relaxation dynamics | 22
      3.3.3 Relaxation per Rouse mode | 22

4. **Considering interaction energy** | 23
  4.1 Characterizing the phase transition for a homopolymer | 23
      4.1.1 Equilibration step determination | 23
      4.1.2 Θ-collapse | 23
  4.2 Introducing more than one monomer type | 25
  4.3 Specific interaction energy between epigenetic states | 25
      4.3.1 The $A_{10}B_{10}$ system | 25
      4.3.2 Characterizing each phase | 26
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Phase diagram dependence on the configuration</td>
<td>37</td>
</tr>
<tr>
<td>5.1 Changing the size of the domains</td>
<td>37</td>
</tr>
<tr>
<td>5.1.1 ((A_5 B_5)_{12})</td>
<td>37</td>
</tr>
<tr>
<td>5.1.2 ((A_{20} B_{20})_3)</td>
<td>38</td>
</tr>
<tr>
<td>5.2 Introducing asymmetry in the size of the domains</td>
<td>39</td>
</tr>
<tr>
<td>5.2.1 Type 1 MPS</td>
<td>40</td>
</tr>
<tr>
<td>5.2.2 Type 2 MPS</td>
<td>41</td>
</tr>
<tr>
<td>5.2.3 Bistability in the ((A_5 B_{15})_6) system</td>
<td>41</td>
</tr>
<tr>
<td>5.3 Introducing a third epigenetic state</td>
<td>42</td>
</tr>
<tr>
<td>5.3.1 ((A_{10} B_{10} C_{10})_4)</td>
<td>42</td>
</tr>
<tr>
<td>6. The “living chromatin”</td>
<td>45</td>
</tr>
<tr>
<td>6.1 A Hamiltonian approach</td>
<td>45</td>
</tr>
<tr>
<td>6.2 State transitions in 3D</td>
<td>46</td>
</tr>
<tr>
<td>6.2.1 Algorithm</td>
<td>47</td>
</tr>
<tr>
<td>7. Conclusions</td>
<td>49</td>
</tr>
<tr>
<td>7.1 On the lattice model</td>
<td>49</td>
</tr>
<tr>
<td>7.2 On multistability in chromatin folding</td>
<td>49</td>
</tr>
<tr>
<td>7.3 On coupling 1D and 3D</td>
<td>49</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

First and foremost, I would like to thank my advisors, Cédric and Daniel, for all the help, fruitful discussions, the great environment of work and the catchy enthusiasm for the subject. I would also like to thank the Centre Blaise Pascal for the computational resources, and all the support personnel for their help.

To all the scientific committee of the programme involved in the academic life in Amsterdam, Rome and Lyon. It has been an incredible experience of learning. My warmest regards to the administrative staff for all the invaluable help, you have made our lives much easier. Without a doubt, this scholarship has played a key role in my present and future, many thanks to the AtoSIM Master programme for the financial support.

To my fellow simulators, for all the moments of suffering and joy. It has been a great ride! To all the friends that I’ve made in this 3 amazing cities, for sharing the moments that I will never forget.

To my family, for their love and encouragement from the distance.
1. INTRODUCTION

Living organisms often show different and stable phenotypes that derive from the same DNA sequence, and are maintained robustly in one of their states by transcriptional and epigenetic mechanisms. One of them is the modification of local chemical marks in the chromatin by modifying enzymes. These marks are distributed through the chromatin, generating linear domains. The interaction between the enzymes and the chemical marks results in the emergence of bistable states for each domain [3]. It has been shown that the pattern of states formed by the domains has a high correlation with the 3D structure of the chromatin, and that the 3D structure itself influences dynamically the markers and therefore the given state of each domain. In this project, we aim to better understand the cross talk between the local states of the domains in the chromatin with the large-scale 3D structure. For this, we will model the chromatin as a chain of Ising spins, to distinguish between active and inactive states. The 3D organization will be modeled by a lattice copolymer, where the interaction between monomers will depend on their nature. The cross-talk between the 3D structure and the distribution of epigenomic states will be modeled by transitions that depend on the number of neighbors of a given type, and a recruitment rate. This gives rise to a “magnetic polymer”-like model, the so called “living chromatin”.

In this chapter, we introduce important concepts in physical modeling of epigenetics, and review some previous attempts to model the 1D scale, and its influence on the 3D structure.

1.1 Epigenomic marks: the Ising spins

In multicellular organisms, cellular differentiation leads to the development of several tissues, which differ in gene expression but not in the genetic information they contain [21]. The modulation of gene expression is partly due to biochemical tags called epigenomic marks, that trigger different regulation mechanisms such as accessibility of the gene to transcription factors or enzymatic complexes [11].

Each tissue therefore corresponds to a particular epigenomic pattern [4], which is established during the cellular differentiation process by developmental signals. After differentiation, these signals disappear, but the epigenomic state is robustly maintained not only during the life of the particular cell, but also through generations. This robustness suggest a mechanism to maintain the state, even if the epigenomic marks are susceptible to environmental changes or dilution during cell division [14].

In this section, we address the formalism that describes the dynamics of epigenomic marks. This mathematical model, based on statistical physics and nonlinear dynamics, shows the emergence of coherent states, and explains their robustness.

1.1.1 Mathematical model (Jost 2014)

An epigenetically isolated DNA region with \( n \) nucleosomes is considered. The epigenomic state of each of the nucleosomes can have 3 values: unmarked (U), active (A) and inactive (I). The system can fluctuate from the active state to the inactive state passing through the intermediate, unmarked state: \( I \rightleftharpoons U \rightleftharpoons A \) [11]. A depiction of this system is shown in Figure 1.1.
Two mechanisms are considered for the modification of the nucleosomal state:

- Action of modifying enzymes, occurring at a rate $\epsilon_X \rho_X$, where $X$ refers to either A or I state, and $\rho$ is the local density of the corresponding state, $\rho_X = n_X/n$.
- Random transitions between states, occurring at a rate $k_0$.

A set of chemical reaction propensities for each of the state transitions is given by Equations 1.1 [11]. The transition $A \rightarrow U \rightarrow I$ is related to the enzyme action rate $\epsilon_I \rho_I$, while the transition $I \rightarrow U \rightarrow A$ is associated with the rate $\epsilon_A \rho_A$.

\[
\begin{align*}
U \rightarrow A, & \quad r_{u\rightarrow a} = (k_0 + \epsilon_A \rho_A)(n - n_A - n_I) \\
U \rightarrow I, & \quad r_{u\rightarrow i} = (k_0 + \epsilon_I \rho_I)(n - n_A - n_I) \\
A \rightarrow U, & \quad r_{a\rightarrow u} = (k_0 + \epsilon_I \rho_I)n_A \\
I \rightarrow U, & \quad r_{i\rightarrow u} = (k_0 + \epsilon_A \rho_A)n_I
\end{align*}
\]

Through a simple mass-action model, the differential equations governing the population of each of the states is given by Equations 1.2.

\[
\begin{align*}
\frac{d\rho_A}{dt} &= (k_0 + \epsilon_A \rho_A)(1 - \rho_A - \rho_I) - (k_0 + \epsilon_I \rho_I)\rho_A \\
\frac{d\rho_I}{dt} &= (k_0 + \epsilon_I \rho_I)(1 - \rho_A - \rho_I) - (k_0 + \epsilon_I \rho_I)\rho_A
\end{align*}
\]

A direct analogy between this system and the zero-dimensional Ising model can be readily made: each nucleosomal state can represent a spin: $I = -1$, $U = 0$, $A = +1$. The enzyme action rates $\epsilon_X$ represent the coupling between spins $J$, and the random transitions described in $k_0$ are related to thermal fluctuations. As in the Ising model, we can define a variable of interest in analogy with the magnetization: $m = \rho_A - \rho_I$.

The set of differential equations can be solved for the steady state case $\frac{d\rho_X}{dt} = 0$, to reveal three fixed points for the magnetization:

- $m_0 = 0 \forall \epsilon$.
- $m_\pm = \pm(k_0/\epsilon)^{1/2}(\epsilon/k_0 + 1)(\epsilon/k_0 - 3)$ for $\epsilon < 3k_0$.

Therefore, a critical point for the bifurcation of the system is identified as $\epsilon_c = 3k_0$. This is a well studied phase transition between ordered and disordered phases in the 1D Ising model [16]. In our isolated region of DNA, when the enzymatic action (recruitment) is strong enough, the system exhibits two stable fixed points, that represent coherent and global active or inactive states. Due to external perturbations, the global dominating configuration of the DNA region can jump from active to inactive or vice versa. If the enzymatic action is not strong enough, there will be no bifurcation of the stable state: the system will remain around $m_0$. 

Fig. 1.1: Transition model for the dynamics of the epigenomic states
1. Introduction

1.1.2 Stability analysis and transition rates

The differential Equations 1.2 can be rewritten in terms of the magnetization $m$ and a variable $s = \frac{n_A + n_I}{n}$. Under the assumption that the dynamics for $s$ is fast in comparison with $m$, a time-scale separation leads to the Fokker-Planck equation for the probability distribution $P(m)$ (Eq. 1.3a). [11]

$$\frac{\partial P}{\partial t} = -\frac{1}{2n} \frac{\partial}{\partial m} \left[ \left( \frac{w_+(m) + w_-(m)}{w_+(m) + w_-(m)} \right) P \right]$$

(1.3a)

Where $w_+ = \frac{(r_{u,a} + r_{r,a})}{n}$ and $w_- = \frac{(r_{u,i} + r_{r,a})}{n}$, the combined propensities to increase or decrease the magnetization. This corresponds to the Fokker-Planck equation for the diffusion of a particle in a unidimensional potential [17], and a steady state distribution (Eq. 1.4) and mean passage time (Eq. 1.5) can be found.

$$P_{steady}(m) = \frac{1}{Z} \exp \left[ \frac{2}{n} \int_{-\infty}^{m} dm' \frac{w_+(m') - w_-(m')}{w_+(m') + w_-(m')} \right]$$

(1.4)

$$\langle \tau \rangle \approx \frac{18\pi}{(\epsilon - \epsilon_c)\sqrt{3(\epsilon/k_0 + 3)}} \exp[V(0) - V(m_-)]$$

(1.5)

From Eq. 1.5, it is evident that the passage time depends on the height of the barrier for the steady-state epigenomic configuration [11], as it is commonly found in transition state theory [13]. In turn, $V(m) = -\log P_{steady}(m)$.

The bifurcation diagram (from the Fokker-Planck equation) as well as probability distribution function for $m$ (from stochastic simulations), and passage times (from both) as a function of $\epsilon/k_0$ are shown in Figure 1.2. The dependence of the passage time on the size of the system is also depicted.

![Fig. 1.2: Bifurcation diagram, magnetization probabilities and passage times in function of $\epsilon/k_0$ and n. [11]](image)

So far, the system has been treated as symmetric (the enzyme action rates that represent the coupling between spins are equal, $\epsilon_A = \epsilon_I$). The asymmetric case has also been investigated by Jost et al., and it reveals a well defined boundary between monostable and bistable regions (Fig. 1.3a). The magnetization as a function of the recruitment strengths is shown in Figure 1.3b.

The previous treatment not only helps elucidate the origin of the possible epigenomic states of an isolated DNA segment, but also, the natural emergence of bistability and transition rates between the states. The main parameters that give rise to such dynamics are the recruitment rates.

1.1.3 Epigenomic mark patterns can affect the 3D structure

A typical distribution of epigenomic marks along the chromatin is shown in Figure 1.4. It has been previously observed that the distribution of epigenomic marks along the chromatin can affect its structure. This resulting 3D configuration is not random, and the 1D compartmentalization results in 3D compartmentalization (Figure 1.5a).
1. Introduction

Fig. 1.3: a. Stability diagram with boundaries between monostable and bistable regions. Bistability is only present for strong and nearly-symmetric recruitment. b. Magnetization values in the asymmetric recruitment case. [11]

Fig. 1.4: Distribution of epigenomic marks along the chromatin of Drosophila [12].

One of the model systems for this type of study is the adaptation of the X-linked genes expression to the X chromosome copy number in C. elegans. This organism can either be male or hermaphrodite. In males, the single copy of the X chromosome is upregulated two-fold, while in hermaphrodites, the two copies are downregulated. The latter process is called dosage compensation (DC) and is achieved by the dosage compensation complex (DCC), a multiprotein complex that acts as a chromatin marker. This regulation not only affects the gene expression, but has also been reported to be correlated with the X chromosome structure. While compacted in a similar way, both cases show a very different localization of the X chromosome in the chromatin. [20]

For males, the upregulated X is located near the nuclear rim, while for hermaphrodites it is more core-located, as shown in Figure 1.5b. The influence of the distribution of the epigenomic marks on the large-scale structure is therefore evident. This is the motivation for the next scale of study.

1.2 A polymer model for the chromatin (Jost et. al. 2014)

In this section, we aim to review a model of the chromatin using a polymer model. The polymer of interest will be composed of isolated regions of DNA that will either be in a global active or inactive state. For now, this states are fixed, as it is also their distribution pattern through the chromatin.

1.2.1 Block copolymer continuum model

Previously, chromatin has been modeled as a self-avoiding bead-spring polymer containing N monomers, each one representing 10 kb of DNA [12]. The Hamiltonian of any conformation is made of two contributions: $H_{\text{chain}}$ for the self-avoiding Gaussian chain (Eq. 1.6), and $H_{\text{inter}}$, which accounts for short-range interactions between monomers (Eq. 1.7).
1. Introduction

Fig. 1.5: a. Electron microscopy of the spatial distribution of epigenomic domains in a nucleus. b. X chromosome localization on males and hermaphrodites of *C. elegans*. [20]

$$H_{chain} = \frac{3k_B T}{2l^2} \sum_n (X_n - X_{n-1})^2 + \sum_{n<m} U_{hc}(r_{nm})$$

$$H_{inter} = \sum_{n<m} E_{nm} \exp \left[ -\frac{r_{nm}^2}{2r_0^2} \right]$$

The first term of $H_{chain}$ represents a harmonic bonding potential with bond length $l$, and the second term is a truncated Lennard-Jones like potential. The $H_{inter}$ term is a Gaussian potential of a length-range $r_0$, with a term $E_{nm}$ that defines the strength of interaction, which differs among the monomers according to their nature. The copolymer model is depicted in Figure 1.6.

Fig. 1.6: The block copolymer model, depicting chromatin domains of different nature, which in turn define different interactions. [12]

As it is shown in Fig. 1.6, there are two possibilities for the interaction between monomers:

- A non-specific interaction, that is present between each pair of monomers: $U_{ns}$.
• A specific, epigenomic interaction between monomers that have the same epigenomic state: $U_s$.

Therefore, the strength of the short range interactions is given by $E_{nm} = U_{ns} + \delta_{mn}U_s$, with $\delta_{mn} = 1$ if the two monomers have the same epigenomic state, and 0 otherwise. The specific, epigenomic interaction is motivated by experimental evidence, that suggests an affinity between proteins that are recruited by an epigenomic state and that create physical bridges between isolated DNA regions [19].

### Molecular dynamics

The dynamics of this system can be naturally explored by molecular dynamics simulations, as it has been done previously by Jost et al. [12]. The dynamics of the system is modeled by a set of coupled equations of motion (Eq. 1.8).

$$m\frac{d^2X_n}{dt^2} = -\frac{\partial H}{\partial X_n} - \zeta \frac{dX_n}{dt} + \eta_n(t)$$

(1.8)

The first part of Eq. 1.8 can be solved using a velocity Verlet algorithm, with the last two terms representing the coupling of the system with a heat bath via the Andersen thermostat [7].

### The Gaussian self-consistent approximation

Another way to explore this system is by analyzing the Langevin equation that comes from the previous equations of motion. If on average, no acceleration takes place, the equations of motion simplify to a set of coupled Langevin equations. (Eq. 1.9).

$$\zeta \frac{dX_n}{dt} = -\frac{\partial H}{\partial X_n} + \eta_n(t)$$

(1.9)

The probability distribution function for $Y = X_n$ therefore obeys the Fokker-Planck equation (Eq. 1.10).

$$\frac{\partial P}{\partial t} = \zeta \sum_n \left[\frac{\partial}{\partial X_n} \left( P \frac{\partial H}{\partial X_n} \right) + k_BT \frac{\partial^2 P}{\partial X_n^2} \right]$$

(1.10)

This equation can be solved self-consistently, by approximating $P$ at each point by a multivariate Gaussian distribution: $P(Y, t) \approx \frac{1}{Z} \exp[-Y + C(t)Y/2]$, with $C(t) = \langle X_n \cdot X_m \rangle / 3$ (the covariance matrix). The initial guess for $P$ is evolved according to the Fokker-Planck equation, so that the Gaussian distribution that describes $P$ at a time $t + \delta t$ is found [12].

### 1.2.2 Chromatin organization

This type of modeling has revealed a complex phase diagram with multistability (Figure 1.7a), even for a toy example that consists on a chain of 120 blocks, that alternates between 10 active and 10 inactive blocks ($A_{10}I_{10}$) [12]. The parameters that are varied are the strengths of compaction $U_{ns}$, and the specific interaction strength $U_s$.

It is very important to note that one of the outcomes represented in Figure 1.7a is a region of multistability, where domains that correspond to the same epigenomic state are generated transiently. Depending on the size of the domains, the timescale and dynamics of its interactions can vary: small domains can have rapid dynamics in the multistable region, while bigger domains might be long-lived [12]. This situation of multistability is consistent with certain biological situations, as it has been observed experimentally. It has been suggested that this type of dynamics is responsible for the co-regulation of distant genes [6], and the response of the chromatin structure to developmental stimuli [15].

Comparison between experimental observation of the 3D chromatin structure (through Hi-C contact maps that experimentally measure the probability of contact between distal chromatin fragments) and simulation using this model shows excellent agreement. The multistable nature of the configurations is explored by considering different initial configurations (a coil configuration and a microphase separation configuration). The experimental result has patterns belonging to both dominating structures, evidencing bistability (Figure 1.7b).
Fig. 1.7: Left: phase diagram of the copolymer \((A_{10}I_{10})_6\). Different dominating structures are found, according to the values of the parameters. The small figures represent heat maps that depict the probability of contact between two monomers. [12] Right: A. Experimental contact map, B. calculated contact map from an initial coil configuration and C. from an initial microphase separation configuration. D. shows the evolution of distance between two genomic loci in time, evidencing bistability. [12]
2. SHIFTING TO A POLYMER LATTICE MODEL

2.1 Lattice model

The molecular dynamics framework applied to a Gaussian polymer is a very natural problem to tackle with a computer simulation. However, the changes of configuration in the polymer chain might occur on very different timescales. It has been observed [1] that the large scales involved on the global configuration of a polymer (in contrast with the short scale of a bond) can be studied in a more simple way. Even though MD simulations give the authentic image of the dynamics, they might be very costly in the long-polymer limit, and in this limit, the long polymeric chain exhibits a behavior that is independent of the chemical details of its monomers. In this section, we focus on a “coarse-grained” model for polymer physics: a Monte Carlo approach on a lattice model.

This framework is based on a previously developed model by Hugouvieux et al. [9]. In this model, a copolymer is modeled as a chain of \( N \) monomers of two types: active (A) and inactive (I). The pattern of distribution of the monomers along the chain is made of alternating blocks of A and I monomers. The number of monomers in the blocks of each type can be varied. This is denoted as \( (A_{B_A}I_{B_I})_n \). The total number of monomers is therefore \( N = n(B_A+B_I) \), where \( B_A \) is the number of monomers per active block, and \( B_I \) is the number of monomers per inactive block.

2.1.1 Monte Carlo algorithm

To replace the complicated dynamics that can be found on a long polymer by a coarse grained model, the space is discretized by dividing it regularly into cells. The centers of this cells form a regular lattice. The size of each cell is a free parameter, and has been previously defined as two monomer volumes [9]. In consequence, each cell can have an occupation number of 0, 1 or 2. The bond-length is also discretized, and can have values of \( 1/2l \) (for monomers resting on the same cell) or \( l \) (for monomers in two adjacent cells), where \( l \) is the regular lattice spacing. The molecular dynamics of the polymer is replaced by a nearest-neighbor hopping dynamics (with some restrains on the type of moves), and the interactions between monomers are limited to nearest neighbors.

This simple model has been shown to reproduce the continuum dynamics in the limit of large scales (larger than the lattice spacing). However, three important features have to be imposed:

a. Polymer connectivity This is achieved by restricting two connected monomers to either rest on the same cell, or on nearest-neighboring cells.

b. Excluded volume interaction Imposed by definition, since a cell can at most be occupied by two monomers.

c. Non-crossing of polymer strands Achieved by restricting the double occupancy of a cell to two monomers that are adjacent (i.e. chemically bonded). This results in the nearest distance between two strands to be equal to the lattice spacing.

The previous restrictions result in the cell polymer dynamics depicted in Figure 2.1. For the 2D case, a hexagonal lattice is used, while for the 3D case a periodic FCC lattice is used. The hexagonal or FCC lattices are preferred over the square or cubic lattices because they give a greater number of nearest neighbors (4 vs. 6 in the 2D case, and 6 vs. 12 in the 3D case), and add flexibility on the range of bonding angles [9].
2. Shifting to a polymer lattice model

Fig. 2.1: 2D lattice polymer dynamics, as defined from the restrictions imposed. $e_1$ and $e_2$ represent acceptable moves for end monomers, while $i_1$, $i_2$, and $i_3$ represent possible moves for internal monomers. $x_1$ and $x_2$ show forbidden moves due to the non-crossing restriction. [9]

Possible moves The acceptable moves fall into two categories: reptation ($e_1$ and $i_2$) and lateral displacement ($e_2$, $i_1$ and $i_3$).

Interactions The moves are not only rejected if they fail to fall on one of the previous categories, but also if they lead to a less energetically favorable conformation. If the new configuration leads to a pair of monomers of the same nature occupying the same site or two nearest-neighbor sites, there is an energy contribution $E_i < 0$. $E_A$, $E_I$ and $E_{A,I}$ can be defined according to the nature of the system, and for our purposes, in a similar fashion than that of the block copolymer continuum model. However, in the previous application of this model (a chain composed of hydrophilic and hydrophobic monomers), only a hydrophobic pair give rise to a contribution. This attractive energy was used to model indirectly the repulsion between the hydrophobic monomers and the solvent [9].

Acceptance ratio According to the Metropolis sampling scheme, an importance-guided random walk will be used to explore the configurations of the system. The relative probability of visiting a given configuration is proportional to its Boltzmann weight. This importance-guided random walk must not take the system out of equilibrium, and therefore, the average number of trial moves leaving one given state has to be equal to the average number of moves that get into that state. [7] This detailed balance condition is given in Eq. 2.1, where $\pi$ represents a transition probability.

$$N(o)\pi(o \rightarrow n) = N(n)\pi(n \rightarrow o) \tag{2.1}$$

The transition probability is given by Eq. 2.2.

$$\pi(o \rightarrow n) = \alpha(o \rightarrow n) \times acc(0 \rightarrow n) \tag{2.2}$$

We can suppose that the way we generate new configurations is symmetric ($\alpha$ is symmetric); therefore, the detailed balance condition reads:

$$N(o)acc(o \rightarrow n) = N(n)acc(n \rightarrow o) \tag{2.3}$$

A trial configuration is generated by randomly choosing a monomer and a type of move. If the move results in a configuration that respects the three criteria, and since the probability density of the new and old states is given by
2. Shifting to a polymer lattice model

their Boltzmann weights, the acceptance ratio of a trial move is given by Eq. 2.4, with the obvious condition that the acceptance probability cannot exceed 1, and where \( N(n)/N(o) = \exp(-\beta(H(n) - H(o))) \). If the move does not respect the three criteria, it is rejected.

\[
\text{acc}(o \rightarrow n) = \begin{cases} 
N(n)/N(o) & \text{if } N(n) < N(o) \\
1 & \text{if } N(n) \geq N(o)
\end{cases}
\] (2.4a)

\[
\text{acc}(o \rightarrow n) = \begin{cases} 
N(n)/N(o) & \text{if } N(n) < N(o) \\
1 & \text{if } N(n) \geq N(o)
\end{cases}
\] (2.4b)

**Ergodicity** The definition of the restrictions for this system has proven successful for avoiding ergodicity problems [9], since no locked-up conformations can be generated. In fact, even for the highest polymer densities, pure reptation can still occur, by moving monomers along the chain. This type of moves dominate the dynamics at melt densities, while lateral displacements dominate for dilute solutions, just as has been observed in experimental situations [1].

2.2 Equilibrium properties for the ideal polymer

The first step in the validation of the lattice model is to recover the well-known scaling laws for an ideal chain of monomers of the same chemical nature, without excluded volume interaction. This system plays the role of an “ideal gas” in polymer physics. It is composed of a freely jointed chain of N rigid segments of a fixed length, able to point in any direction. The ideal polymer also disregards interactions between monomers that are not chemically bonded. For this case, the conditions (b. Excluded volume interaction) and (c. Non-crossing of polymer strands) are omitted.

2.2.1 End-to-end distance

A first observable of interest is the end-to-end vector \( \vec{R} \), mean squared and averaged over all the conformations sampled, \( \langle \vec{R}^2 \rangle \). The end-to-end vector can be written as the sum of each bond vector, \( \vec{u}_i = \vec{x}_{i+1} - \vec{x}_i \).

\[
\vec{R} = \sum_{i=1}^{N} \vec{u}_i
\] (2.5)

The mean squared end-to-end vector is therefore:

\[
\langle \vec{R}^2 \rangle = \left\langle \left( \sum_{i=1}^{N} \vec{u}_i \right)^2 \right\rangle = \sum_{i=1}^{N} \langle \vec{u}_i^2 \rangle + 2 \sum_{1 \leq i < j \leq N} \langle \vec{u}_i \vec{u}_j \rangle
\] (2.6a)

Since the segment directions are not correlated, \( \langle \vec{u}_i \vec{u}_j \rangle = 0 \). In addition, \( \langle \vec{u}_i^2 \rangle = \frac{12}{13} \frac{b^2}{2} \), since are 12 possibilities of having the next monomer at a distance \( \frac{b^2}{2} \) (the 12 nearest neighbors in the FCC lattice), plus an extra possibility of having the next monomer in the same lattice site. The scaling law between the end-to-end distance and the length of the polymer is given in Eq. 2.7

\[
\langle \vec{R}^2 \rangle = N \frac{12}{13} \frac{b^2}{2}
\] (2.7)

2.2.2 Equilibration steps determination

A plot of the RMS end-to-end vector vs. the number of MC steps, without equilibration, is shown in Figures 2.2 and 2.3, for different polymer sizes. The equilibrated condition of the system is found when \( \langle \vec{R}^2 \rangle \) does not change significantly in time.
2. Shifting to a polymer lattice model

The equilibration time increases roughly with $N_{\text{chain}}^2$, and for the measurements, it is chosen as $4 \times N_{\text{chain}}^2$. Each equilibration step correspond to $N_{\text{chain}}$ substeps.

2.2.3 Scaling of the RMS end-to-end vector

A plot of $\langle \vec{R}^2 \rangle$ vs. $N - 1$ is shown in Figure 2.5. The measurements correspond to 200 different polymer strands per polymer size, each one with $4 \times N_{\text{chain}}^2$ equilibration steps.

The linear dependence with respect to $N$ is recovered. The slope is fitted to a value of 0.457, which corresponds closely to the theoretical value of $\frac{12}{13} = 0.462$. 
2. Shifting to a polymer lattice model

2.2.4 Radius of gyration

The radius of gyration is a quantity that gathers information not only from the first and last monomers of the polymer, but from every monomer position with respect to the center of mass. It is defined in Eq. 2.8.

\[ R_g^2 = \frac{1}{N_{\text{bonds}}} \left\langle \sum_i (\vec{S}_i)^2 \right\rangle \]  

(2.8)

Where \( \vec{S}_i = \vec{r}_i - \vec{r}_{cm} \). The following relation can be considered, to simplify this expression:

\[ \sum_{i,j} (\vec{S}_{ij})^2 = n \sum_i (\vec{S}_i)^2 + n \sum_j (\vec{S}_j)^2 - 2 \sum_{i,j} (\vec{S}_i \cdot \vec{S}_j)^2 \]  

(2.9)

The last summation of this expression is null, since by definition of the center of mass, \( \sum_i \vec{S}_i = 0 \). Therefore:

\[ R_g^2 = \frac{1}{2N^2m} \left\langle \sum_{i,j} (\vec{S}_{ij})^2 \right\rangle \]  

(2.10a)

\[ = \frac{1}{2N_{\text{bonds}}^2} \left\langle \sum_{i,j} (\vec{r}_i - \vec{r}_j)^2 \right\rangle \]  

(2.10b)

Writing \( \left\langle \sum_{i,j} (\vec{r}_i - \vec{r}_j)^2 \right\rangle \) as \( |n - m| b^2 \), and taking the limit of large \( N \), we get:

\[ R_g^2 \approx \frac{b^2}{2N^2} \int_0^N dn \int_0^N dm \, |n - m| \]  

(2.11a)

\[ = \frac{b^2}{N^2} \int_0^N dn \int_0^N dm (n - m) \]  

(2.11b)

\[ = \frac{Nb^2}{6} = \frac{\left\langle R_g^2 \right\rangle}{6} \]  

(2.11c)

A plot of \( \left\langle R_g^2 \right\rangle \) vs. \( N - 1 \) is shown in Figure 2.5. The linear dependence with respect to \( N-1 \) is recovered. The slope is fitted to a value of 0.077, which corresponds approximately to the theoretical value of \( \frac{12}{13} \frac{11}{26} = 0.0769 \).
2.3 Dynamical properties for the ideal polymer: results from the Rouse model

In this section, we address the dynamical properties of the ideal polymer chain. The first quantities to be analyzed are the mean squared displacement of the center of mass of the polymer and the mean squared displacement of the middle monomer. Another quantity of interest is the time correlation function for the end-to-end vector. We expect to recover the analytical results from the Rouse model, the simplest theory for polymer dynamics that has as starting point the Gaussian chain model.

2.3.1 \( g_1 \) and \( g_3 \)

By evaluating the mean squared displacement of the middle monomer (Equation 2.12) and of the center of mass of the polymer (Equation 2.13), we can validate that the lattice model reproduces the continuum dynamical properties.

\[
g_1(t) = \langle [\hat{R}_{N/2}(t) - \hat{R}_{N/2}(0)]^2 \rangle \tag{2.12}
\]

\[
g_3(t) = \langle [\hat{R}_{CM}(t) - \hat{R}_{CM}(0)]^2 \rangle \tag{2.13}
\]

A plot for \( g_1 \) and \( g_3 \) is shown in Figure 2.3.1. The expected behaviors from the Rouse model [8] are recovered:

\( g_1 (\text{MSD}_{N/2}) \) subdiffusion regime:

\[
g_1 \propto t^{0.5} \tag{2.14}
\]

\( g_1, g_3 (\text{MSD}_{CM}) \) diffusion regime:

\[
g_1, g_3 \propto t \tag{2.15}
\]
2.3.2 Time correlation function for the end-to-end distance

One of the most fundamental dynamical properties to be analyzed is how the end-to-end vector relaxes: \( \vec{R}(t) = \vec{r}(t, N) - \vec{r}(t, 0) \). The Rouse model predicts the following for the correlation function for the end-to-end vector [8]:

\[
\langle \vec{R}(t) \vec{R}(0) \rangle = \frac{8N a^2}{\pi^2} \sum_{p=1,3,5,...} \frac{1}{p^2} \exp \left( -\frac{tp^2}{\tau_1} \right) \tag{2.16}
\]

Where the relaxation time is defined as:

\[
\tau_1 = \frac{N^2 a^2 \zeta}{3\pi^2 T} \tag{2.17}
\]

\[
\tau_p = \frac{\tau_1}{p^2} = \frac{N^2 a^2 \zeta}{3\pi^2 T p^2} \tag{2.18}
\]

As seen from Eq. 2.16, the relaxation of the end-to-end vector autocorrelation function is exponential. This relationship is recovered for different polymer sizes, as shown in Fig. 2.6. The relaxation time is found to be \(1.147N^2 \pm 0.002N^2\).

![Fig. 2.6: log(\( \langle \vec{R}(0) \vec{R}(t) \rangle \)/log(\( \langle \vec{R}(0) \rangle^2 \))) vs t/N^2 for different polymer sizes.](image)

This model also predicts the a relationship between \( \langle \vec{R}(t) \vec{R}(0) \rangle \) and the correlation of each Rouse mode, as shown in Eq. 2.19 [8]. This relationship is recovered for different polymer sizes, as seen in Figure 2.3.2. The factor of
2. Shifting to a polymer lattice model

The proportionality found is $16.80 \pm 0.04$, a good agreement considering that the summation has been truncated up to the third odd Rouse mode.

$$\left\langle \vec{R}(t)\vec{R}(0) \right\rangle = 16 \sum_{1,3,5,...} \left\langle \vec{y}_p(t)\vec{y}_p(0) \right\rangle$$  \hspace{1cm} (2.19)

2.3.3 Time correlation function per Rouse mode

The relaxation time per Rouse mode is given in Eq. 2.18. Time correlation functions were also calculated for the first three modes. A normalized plot, scaled with $p^2$, is shown in Fig. 2.7, for different polymer sizes. The relaxation times are shown to be in the following orders:

- Mode 1: $\approx 1.0N^2$.
- Mode 3: $\approx 0.11N^2$.
- Mode 5: $\approx 0.04N^2$.

Which confirms the scaling with $1/p^2$, expressed in Eq. 2.16.

**Fig. 2.7:** $\log(\langle \vec{y}_p(0)\vec{y}_p(t) \rangle) / \log(\langle \vec{y}_p(0)^2 \rangle) \ vs \ t\frac{p^2}{N^2}$ for different polymer sizes.
2. Shifting to a polymer lattice model

2.4 Considering polymer flexibility

For a freely jointed chain with bending energy, \( \langle \vec{u}_i \vec{u}_i \rangle \neq 0 \), since the directions of bonds are now correlated. Since \( \langle \vec{u}_i \vec{u}_i \rangle \langle \cos \theta_{ij} \rangle \), the correlation that arises from considering chain flexibility can be described with the mean cosine of the bond angles.

The quantity \( \langle \cos \theta(s) \rangle \), which describes the mean cosine between bonds separated by length \( s \), has the following multiplicativity property:

\[
\langle \cos \theta(s + s') \rangle = \langle \cos \theta(s) \rangle \langle \cos \theta(s') \rangle \tag{2.20}
\]

Since the bond length in this model is a constant, the multiplicativity property can be rewritten as in Eq. 2.21. A function having this multiplicativity property can be expressed as an exponential decay, as written in Eq. 2.22. [8]

\[
\langle \cos \theta_{i,i+k} \rangle = (\cos \gamma)^k \tag{2.21}
\]

\[
\langle \cos \theta(s) \rangle = \exp(-s/\tilde{l}) \tag{2.22}
\]

Where \( \tilde{l} \) is a constant for a given polymer, denominated its persistent length. Using Eq. 2.21 and Eq. 2.22, the persistent length of the polymer can be found in terms of the mean value of the cosine of the angle between bonds [8]:

\[
\tilde{l} = \frac{b}{\sqrt{2} |\ln \cos \gamma|} \tag{2.23}
\]

The bending energy for a given angle can be expressed as in Eq. 2.24. The expected value of \( \cos \gamma \) can be found from its Boltzmann distribution, as shown in Eq. 2.25.

\[
U_{bend} = K_{bend}(1 - \cos \theta_{i,i+1}) \tag{2.24}
\]

\[
\langle \cos \gamma \rangle = \frac{\sum \theta \cos \theta \exp(-k(1 - \cos \theta))}{\sum \theta \exp(-k(1 - \cos \theta))} \tag{2.25}
\]

The loop (two chemically bonded monomers occupying the same lattice site) is not considered to contribute as forming a bond angle. In the FCC lattice, the discrete values of the angles are:

- \( \theta = 0 = 1 \) possibility.
- \( \theta = \pi/3 = 4 \) possibilities.
- \( \theta = \pi/2 = 2 \) possibilities.
- \( \theta = 2\pi/3 = 4 \) possibilities.
- \( \theta = \pi = 1 \) possibility.

2.4.1 Equilibration considerations

The criterium for equilibrated measurements is the same used in the previous section: measurement of the \( \langle \vec{R}^2 \rangle \) until it does not change significantly in time. A plot of \( \langle \vec{R}^2 \rangle \), for different polymer sizes, is shown in Fig. 2.8. The equilibration time is set to \( 4 \times N_{chain}^2 \).

2.4.2 Correlation between bond vectors

A plot of \( \langle \cos \gamma \rangle \) with respect to the bending constant, from both the numerical simulation and the Boltzmann distribution, is shown in Fig. 2.9, evidencing good agreement between the two quantities.
Fig. 2.8: $\langle R^2 \rangle$ vs $K_{\text{bend}}$ for a. $N=50$, b. $N=100$, c. $N=150$ and d. $N=200$, to determine the number of equilibration steps.

Fig. 2.9: $\langle \cos \gamma \rangle$ vs. bending constant, for numerical simulations of different polymer sizes, and its theoretical value from Eq. 2.25.
2.4.3 Size of the polymer

Since we now have correlation between bond vectors, the mean squared end-to-end distance can be written as:

\[
\langle \vec{R}^2 \rangle = \sum_{i=1}^{N} \langle \vec{u}_i^2 \rangle + 2 \sum_{1 \leq i < j \leq N} \langle \vec{u}_i \vec{u}_j \rangle \tag{2.26a}
\]

\[
= \frac{12}{13} \frac{Nb^2}{2} + \frac{12}{13} k^2 \sum_{1 \leq i < i+k \leq N} \langle \cos \theta_{i,i+k} \rangle \tag{2.26b}
\]

Substituting Eq. 2.21, the exact result for the mean squared end-to-end distance takes the form shown in Eq. 2.27.

\[
\langle \vec{R}^2 \rangle = \frac{12}{13} \frac{Nb^2}{2} \left[ \frac{1 + \cos \gamma}{1 - \cos \gamma} - \frac{2}{N} \cos \gamma \frac{1 + (\cos \gamma)^N}{(1 - \cos \gamma)^2} \right] \tag{2.27}
\]

A plot of \( \langle \vec{R}^2 \rangle \) with respect to the bending constant, from both the numerical simulation and the theoretical value from Eq. 2.27 is shown in Fig. 2.10.

---

Fig. 2.10: \( \langle \vec{R}^2 \rangle \) vs. bending constant, for numerical simulations of different polymer sizes, and its theoretical value from Eq. 2.27.
3. CONSIDERING EXCLUDED VOLUME INTERACTION

In this model, volume interactions are considered by imposing the non-crossing of polymer strands, through the condition that double occupancy is only allowed when two monomers are chemically bonded. The simplest manifestation expected from volume interactions is the swelling of the polymer coil. The spatial dimension of the polymer chain, defined through the mean squared end-to-end vector, can be related to its Gaussian counterpart by the swelling parameter $\alpha$, as shown in Eq. 3.1.

$$\alpha^2 = \frac{R^2}{R_0^2}$$  \hspace{1cm} (3.1)

Where $R_0$ refers to the end-to-end vector of the ideal polymer. The free energy of a swollen polymer coil can be written as the sum of two terms:

$$F(\alpha) = F_{el}(\alpha) + F_{int}(\alpha)$$ \hspace{1cm} (3.2)

For the freely jointed chain, the end-to-end vector obeys a Gaussian distribution [8]:

$$P_N(\vec{R}) \propto \exp \left[-\frac{3\vec{R}^2}{2N\bar{l}^2}\right]$$ \hspace{1cm} (3.3)

Where $\bar{l}$ corresponds to the persistent segment length. The entropy of a given configuration can be found from:

$$S = \ln P(\vec{R}) \hspace{1cm} (3.4a)$$

$$= \text{constant} - \frac{3\vec{R}^2}{2N\bar{l}^2} \hspace{1cm} (3.4b)$$

We can write the free energy as:

$$F(\vec{R}) = E + \frac{3T\vec{R}^2}{2N\bar{l}^2}$$ \hspace{1cm} (3.5)

The internal free energy can be expanded into a power series of the density of monomers (the virial expansion) [8], for low density systems:

$$E = TN(\rho B + \rho^2 C + ...)$$ \hspace{1cm} (3.6)

The density of monomers is of the order of $N/\|\vec{R}\|^3$. Truncating in the first term of the virial expansion, we can therefore write the total free energy as:

$$F(\vec{R})/T = \frac{N^2 B}{\vec{R}^3} + \frac{3T\vec{R}^2}{2N\bar{l}^2}$$ \hspace{1cm} (3.7)

The equilibrium free energy can be found by deriving with respect to the end-to-end vector, and setting equal to zero:

$$\frac{-3N^2 B}{\vec{R}^4} + \frac{3R}{N\bar{l}^2} = 0$$ \hspace{1cm} (3.8)

This reveals a relationship between the magnitude of the end-to-end vector and the polymer size:

$$R \propto N^{3/5}$$ \hspace{1cm} (3.9)

The exponent $\nu = 3/5$ is known as the Flory exponent. Although the previous derivation is only an approximation, the scaling law found is consistent with the result coming from the rigorous renormalization group theory: $\nu = 0.592$. [8]
3. Considering excluded volume interaction

3.1 Equilibration time determination

The same procedure of the ideal polymer is used to determine the number of steps before equilibration: measuring $\langle \vec{R}^2 \rangle$ until it does not evidence a significant change. A plot of $\langle \vec{R}^2 \rangle$ vs. MC steps is shown in Fig. 3.1. The equilibration time is fixed as $5N^2$ MC steps.

![Fig. 3.1: $\langle \vec{R}^2 \rangle$ vs MC steps for different polymer sizes with excluded volume interaction.](image)

3.2 Scaling of the polymer size

A plot of $\langle \vec{R}^2 \rangle$ vs. polymer size is shown in Figure 3.2, using 200 different initial configurations. The scaling exponent recovered is 0.5895, in agreement with the theoretical values 0.6 (Flory model) and 0.588 (renormalization group theory).

![Fig. 3.2: $\langle \vec{R}^2 \rangle$ vs. size of the polymer, with excluded volume interaction.](image)

3.2.1 Radius of gyration

Since now the scaling of the distance between monomers is $2\nu$, the squared radius of gyration can be expressed as follows:
\[ R_g^2 = \frac{b^2}{2N^2} \int_0^N dn \int_0^N dm |n - m|^{2\nu} \]  
\[ = \frac{b^2}{N^2} \int_0^N dn \int_0^N dm (n - m)^{2\nu} \]  
\[ = \frac{b^2}{(2\nu + 1)(2\nu + 2)} N^{2\nu} \]

A plot of the mean squared radius of gyration for different polymer sizes is shown in Fig. 3.3. The mean squared radius of gyration also shows the same scaling, with an exponent of 0.5905. The prefactor is estimated as 0.072, in close agreement with the theoretical value of \( \frac{12}{13(2(2\nu + 1)(2\nu + 2))} = 0.0656 \).

![Graph showing \( \langle R_g^2 \rangle \) vs. size of the polymer, with excluded volume interaction.](image)

**Fig. 3.3:** \( \langle R_g^2 \rangle \) vs. size of the polymer, with excluded volume interaction.

### 3.3 Dynamical properties

#### 3.3.1 \( g_1 \) and \( g_3 \)

As for the ideal Gaussian chain, the mean squared displacement of the middle monomer and center of mass are explored. A plot of \( g_1 \) and \( g_3 \) is shown in Figure 3.3.1.
3.3.2 Relaxation dynamics

The dynamics of relaxation are also explored. A logarithmic plot of the correlation function \( \langle \vec{R}(0)\vec{R}(t) \rangle \) with respect to time scaled with \( N^2 \) is shown in Fig. 3.4. The dynamics appear to follow an exponential decay, similar to the one found in the Rouse model. However, the relation times are increased roughly by a factor of 5. The relaxation time is now estimated as \( 6.031N^2 \pm 0.006 \).

![Fig. 3.4: \( \langle \vec{R}(0)\vec{R}(t) \rangle / \langle \vec{R}^2(0) \rangle \) vs. time/\( N^2 \), with excluded volume interaction.](image)

3.3.3 Relaxation per Rouse mode

The relaxation time per Rouse mode, with excluded volume interaction, also scales with \( 1/p^2 \) (as in the Gaussian chain case). This is evident in Figure 3.5 for different polymer sizes.

![Fig. 3.5: \( \log(\langle \vec{y}_p(0)\vec{y}_p(t) \rangle) / \log(\langle \vec{y}_p(0)^2 \rangle) \) vs. \( t\frac{p^2}{N^2} \), with excluded volume interaction, for different polymer sizes.](image)
4. CONSIDERING INTERACTION ENERGY

In this chapter, we introduce the interaction energy between monomers. First, we recover a well-known relationship that characterizes the phase transition between a coil and a globule due to an attractive interaction energy between monomers. The coil-globule transition, known as the Θ-collapse, can also be achieved by confining the polymer to a high density. In a previous attempt to characterize the influence of the epigenetic marks on the 3D structure of the chromatin, a non-specific interaction energy was used to emulate higher densities. For the present exercise, the Θ-collapse will be achieved through confinement of the polymer at high densities. In the second part of this section, we explore different phases of the toy model (A_{10}B_{10})_6, using as parameters the specific interaction energy between monomers of the same type, and the density of the system.

4.1 Characterizing the phase transition for a homopolymer

This problem can be also treated with a mean field approximation, forgetting the details about the chemical connectivity and supposing we have a "gas" of N monomers in a ball of radius R. The internal energy is now expanded up to the second virial coefficient, taking into account that the first virial coefficient is negative (reflecting the attractive interaction).

\[
\frac{F(\vec{R})}{T} = -\frac{N^2 B}{R^3} + \frac{N^3}{R^6} + \frac{3T\vec{R}^2}{2N\vec{l}^2} \tag{4.1}
\]

The equilibrium free energy can be found by deriving with respect to the end-to-end vector, and setting equal to zero:

\[
\frac{3N^2 B}{R^3} - \frac{N^3}{R^6} + \frac{3R}{N\vec{l}^2} = 0 \tag{4.2a}
\]

\[
\frac{3N^2 B}{R^5} - \frac{N^3}{R^8} + \frac{3}{N\vec{l}^2} = 0 \tag{4.2b}
\]

The third term decays with the size of the polymer. Neglecting it reveals a relationship between the magnitude of the end-to-end vector and the polymer size:

\[R \propto N^{1/3} \tag{4.3}\]

Therefore, there is a phase transition to a globular conformation with a scaling characterized by \(\nu = 1/3\). Close to the transition point, the polymer behaves like an ideal, Gaussian chain, and therefore, \(\nu = 1/2\).

4.1.1 Equilibration step determination

A system composed of a polymer chain with only one type of monomer, which exhibits an attractive interaction between monomers, was explored. The equilibration time was estimated using the procedure of stabilization of the mean squared end to end vector, as shown in Fig. 4.1. The equilibration time is fixed as 5N^2.

4.1.2 Θ-collapse

The mean squared radius of gyration with respect to the interaction energy was explored, for 200 different initial configurations. The phase transition is evident at an interaction energy of around 0.16\(k_B T\), where \(\nu = 1/2\) independent of the polymer size. Finite size effects are also evident: the phase transition is sharper for longer polymers.
Fig. 4.1: $\langle \vec{R}^2 \rangle$ vs $E_{int}$ for a. $N=50$, b. $N=100$, c. $N=150$ and d. $N=200$, to determine the number of equilibration steps.

Fig. 4.2: $\langle \vec{R}_g^2/N \rangle$ vs. interaction energy, for different polymer sizes.
4.2 Introducing more than one monomer type

A system composed of a polymer chain with two types of monomer was explored. Monomers of type A have a specific attractive interaction between each other, while monomers of type B do not have any interaction. The monomer configuration is intercalated: 3 monomers of type A followed by 3 monomers of type B. The phase transition is still existent, although it is shifted from its original value, as it is shown in Fig. 4.3.

![Figure 4.3: \( \langle R_g^2/N \rangle \) vs. interaction energy in the copolymer \((3_A3_B)_N/3\), for different polymer sizes.](image)

This type of system has been previously explored by Hugovieux et. al. [9], in the context of amphiphilic multiblock copolymers. These copolymers consist of polar and apolar monomers, that, depending on the solvent, have different behaviors. The specific interaction is introduced to model hydrophobicity in the apolar monomers, in a non-explicit fashion. The characteristic configuration found is a hydrophobic core surrounded by a hydrophilic shell, which agrees with experimental observations [23]. Depending on the proportion of hydrophilic to hydrophobic monomers, several different structures, such as tubular or layered configurations, have been found. This gives insight on the dependence of phase behavior on the polymer configuration.

4.3 Specific interaction energy between epigenetic states

For the case of study of this project, the polymer configuration will represent the epigenetic marks distribution along the chromatin, with a key difference: monomers of the same type will interact between themselves. This type of specific interaction is motivated by experimental observations, which suggest an affinity between same epigenetic states [19]. This affinity might be promoted by physical bridging of the proteins associated to maintaining a certain epigenetic state [2] [10].

4.3.1 The \( A_{10}B_{10} \) system

A toy model of \( N=120 \), composed by intercalated domains of two different types of monomers, with each domain composed of 10 monomers, was explored. The parameters for phase space exploration are the specific interaction strength (in a symmetrical fashion) and the density of the system. Periodic boundary conditions were used to mimic crowding in highly dense states. Density was varied by changing the size of the box, with the limitation of having only certain values of density to be explored, due to the discrete nature of the system. An increased equilibration time was used, \( 100N^2 \), to assure that the measurements for each point in phase space are in equilibrium. To accelerate phase space exploration, the same initial condition was used for different energies, varying gradually in 0.1 \( k_BT \) steps. Two different initial conditions were used to assure that the results are consistent with an equilibrium condition: coil and globular. The globular conformation was first obtained by equilibrating the system using a non-specific interaction strength, just above the \( \Theta \)-collapse condition.

A complex phase diagram

The phase diagram for the \( (A_{10}B_{10})_6 \) is shown in Figure 4.4. For a system without interaction energy, the expected phase transition from coil to globule due to confinement is found, at a density close to 0.09. For high interaction...
energies, a microphase separation is observed. For intermediate regimes, a bistable region is recognized. This bistable region has been previously observed in the toy model by Jost et. al. [12], using a block copolymer off-lattice model with a non-specific interaction energy to mimic high densities, as previously mentioned. Coil and intermediate microphase separation phases are also found, which are characterized by a mycelle-chain like formation, with varying degrees of compaction.

![Phase diagram](image)

**Fig. 4.4:** Phase diagram for the \((A_{10}B_{10})_6\), using density and specific interaction energy strength as phase space parameters.

### 4.3.2 Characterizing each phase

The following properties were used to characterize each of the points in phase space as belonging to one of the phases identified:

- **Contact map:** a logarithmic plot of the contact probability between pairs of monomers.
- **Distance map:** a logarithmic plot of the mean squared distance between pairs of monomers.
- **Distribution of radius of gyration:** a histogram showing the relative size distribution of the whole polymer, and the A or B domains.
- **Scatter plots of radius of gyrations:** bihistogram of the quantities \((R_g, R_{g,A})\), \((R_g, R_{g,B})\) and \((R_{g,A}, R_{g,B})\).
- **\(MSD_{monomer}\):** average of the mean squared displacement of individual monomers.
- **\(MSD_{CM}\):** mean squared displacement of the center of mass of the polymer, and of each type of monomer (g3).

**Coil phase**

The coil phase is found for low densities and low specific interaction energies. As seen in Figure 4.5, the interaction between regions within a short genomic distance is prevailing, and it decays quickly for regions farther apart from the diagonal.

The distribution of the radius of gyration is wide, as seen in Figure 4.6. Both A and B monomers share the same distribution as the general radius of gyration, which implies that no subdomains are formed, as expected. Scatter
Fig. 4.5: a. Contact probability and b. RMS distance between different regions of the polymer, in logarithmic scale, for the coil phase of the \((A_{10}B_{10})_6\) configuration.

plots for \((R_g, R_{g,A})\) and \((R_{g,A}, R_{g,B})\) are shown in Figure 4.7, which show as well a wide distribution of sizes, with no characteristic pattern formation.

Fig. 4.6: Histogram of the radius of gyration for all monomers, and for each monomer type, in the coil phase of the \((A_{10}B_{10})_6\) system.

Finally, the mean squared displacements of each type of monomer, and the G3 quantity, show normal diffusive behavior (Figure 4.8) for both the general and monomer-specific quantities, which implies, again, that no preferential domains are formed.

Globule phase

The main difference between the coil phase and the globular phase lies in the distribution of radius of gyration. As seen in Figure 4.9, there is a very narrow distribution of sizes for the globule. Since this phase is found due to confinement, the radius of gyration is limited to that of a box with periodic boundary conditions, \(L^2/4\). The scatter plot of the quantities \((R_g, R_{g,B})\) and \((R_{g,A}, R_{g,B})\) (Figure 4.10) also shares this characteristic.
Fig. 4.7: a. Scatter plot of \((R_g, R_{g,A})\) and b. \((R_{g,A}, R_{g,B})\), for the coil phase of the \((A_{10}B_{10})_6\) configuration.

Fig. 4.8: a. Mean-squared displacement of individual monomers and b. of the center of mass for the coil phase of the \((A_{10}B_{10})_6\) configuration.

Fig. 4.9: Histogram of the radius of gyration for all monomers, and for each monomer type, in the globule phase of the \((A_{10}B_{10})_6\) system.
Since now the polymer is confined to a smaller space, the probability of contact between distant monomers is increased, and this results in more uniform contact and distance maps between monomers, as seen in Figure 4.11. This difference is clearly seen when the contact probability is plotted in function of the genomic distance (Figure 4.12). For both cases, an power-law decay is found. However, for the globular phase, the probability of contact remains constant in a non-zero value after a genomic distance of around 10 monomers.

**Microphase separation (MPS)**

The effect of having a strong, specific interaction energy is clearly seen in the contact and distance maps for the MPS (Figure 4.14), which show how two clear, distinct domains form: one for the A monomers and one for the B monomers. The mean distances between monomers remain low since the polymer is collapsed, with smaller values inside the subdomains. This effect is also evident in the distribution of radius of gyration (Figure 4.15), with very sharp distributions for all cases, but centered around a smaller value for the A or B subtypes. The probability of contact in function with genomic distance is therefore oscillating (Figure 4.13).

---

**Fig. 4.10**: a. Scatter plot of \((R_g, R_{g,A})\) and b. \((R_{g,A}, R_{g,B})\), for the globule phase of the \((A_{10}B_{10})_6\) configuration.

**Fig. 4.11**: a. Contact probability and b. RMS distance between different regions of the polymer, in logarithmic scale, for the globule phase of the \((A_{10}B_{10})_6\) configuration.
4. Considering interaction energy

Fig. 4.12: Probability of contact in function of genomic distance for a. coil and b. globule phases of the $(A_{10}B_{10})_6$ configuration.

Fig. 4.13: Probability of contact in function of genomic distance, for the MPS in the $(A_{10}B_{10})_6$ system.

A very interesting characteristic of the dynamical properties of this phase is identified in Figure 4.16. While the mean squared displacement of individual monomers is the same, the mean squared displacement of the center of mass of the subdomains is larger than that of the center of mass of the whole polymer. This suggests that the increased interaction between monomers of the same type results in highly mobile globules of the same type, perhaps one of the mechanisms of the long-ranged interactions between genomically distant domains belonging to the same epigenetic state.
4. Considering interaction energy

![Image](a) Density=0.087, E_specific=0.700

![Image](b) Density=0.087, E_specific=0.700

**Fig. 4.14:** a. Contact probability and b. RMS distance between different regions of the polymer, in logarithmic scale, for the microphase separation of the \((A_{10}B_{10})_6\) configuration.

![Image](Density=0.087, E_specific=0.700)

**Fig. 4.15:** Histogram of the radius of gyration for all monomers, and for each monomer type, in the microphase separation of the \((A_{10}B_{10})_6\) system.

**Coil MPS**

For low densities and specific interaction energies with values of 0.2 and 0.3 \(k_B T\), an intermediate regime is found. This is characterized by an increased probability of contact between monomers of the same type, but not a full collapse into subdomains, as seen in Figure 4.17. This increased contact probability is not reflected in the mean distance map, which suggests short-lived contacts. The effect is also visible on the distribution of the radius of gyration (Figure 4.18), which is roughly the same for the general polymer and for each type of monomer.
4. Considering interaction energy

Fig. 4.16: a. Mean-squared displacement of individual monomers and b. of the center of mass for the MPS of the \((A_{10}B_{10})_6\) configuration.

Fig. 4.17: a. Contact probability and b. RMS distance between different regions of the polymer, in logarithmic scale, for the coil MPS of the \((A_{10}B_{10})_6\) configuration.

Fig. 4.18: Histogram of the radius of gyration for all monomers, and for each monomer type, in the coil MPS of the \((A_{10}B_{10})_6\) system.
Intermediate MPS

The main difference between the coil MPS and the intermediate MPS is an effect of the confinement. As it can be seen in Figure 4.19, at lower densities, the polymer tends to have an increased probability of contact only between genomically close domains. This could be thought as a sort of mycelle-chain like configuration. For higher densities, the probability of contact is more uniform, regardless of the genomic distance. This is a very similar characteristic to the one that helps differentiate between the coil and globular phases, and in fact, the boundary between phases corresponds to the same density.

![Contact probability map](image)

(a) Density=0.087, E_specific=0.200
(b) Density=0.030, E_specific=0.300

Fig. 4.19: Comparison of the contact probability map between the a. intermediate MPS and b. coil MPS.

Bistability

Perhaps the most striking feature of the phase space exploration of this toy model is a region of bistability. This bistability has been previously observed by Jost et. al. [12], as mentioned in the first chapter, although a different phase space parameter has been used: in this case density instead of non-specific interaction energy. The clearest evidence of the bistability regime is seen in Figure 4.20. While the distribution of sizes for the polymer remains unimodal, the radius of gyration for each type of monomer shows two peaks. This implies that either the intradomain or the interdomain distances can assume swollen or collapsed configurations.
4. Considering interaction energy

Fig. 4.20: Histogram of the radius of gyration for all monomers, and for each monomer type, in the coil MPS of the \((A_{10}B_{10})_6\) system.

The contact probability and mean distance map (Figure 4.21) are very similar to the microphase separation case, and this is because they are averaged in time. To further characterize this transition between states, a time series of the distance between the centers of mass of different domains belonging to the same epigenetic state is shown in Figure 4.22 and Figure 4.23. The first one depicts the distance between the centers of mass of domains genomically distanced 1 domain (\(ABA\)), 3 domains (\(ABABA\)) or 5 domains (\(ABABABA\)), imposed over each other. The second one shows more genomic distances, shifted one order of magnitude.

Fig. 4.21: a. Contact probability and b. RMS distance between different regions of the polymer, in logarithmic scale, for the bistable region of the \((A_{10}B_{10})_6\) configuration.

From the time series, it is evident that the distance between centers of mass of different domains is effectively jumping between two values. An increased interaction energy results in longer-lived collapsed states. Figure 4.25 shows two characteristic 3D configurations of the polymer: either a big A phase surrounded by two smaller B phases, or the contrary. This helps to identify that the bistability (or multistability) results in a microphase separation, with 2, 3 or more microphases composed of a varying number of A or B domains.
4. Considering interaction energy

Fig. 4.22: Time series of distance between A domains genomically distanced 1, 3 or 5 domains, for two different interaction energies.

Fig. 4.23: Time series of distance between A domains genomically distanced 1, 3, 5, 7 or 9 domains, for two different interaction energies, shifted by orders of magnitude.
Fig. 4.24: Characteristic configurations for a. coil, b. glouble and c. microphase separation for the \((A_{10}B_{10})_6\) configuration.

Fig. 4.25: Two characteristic configurations belonging to the bistability region for the \((A_{10}B_{10})_6\) configuration.
5. PHASE DIAGRAM DEPENDENCE ON THE CONFIGURATION

The richness of the phase diagram for the (A\textsubscript{10}B\textsubscript{10})\textsubscript{6} configuration with respect to the homopolymer suggests that the distribution of the types of monomer along the chain (or the distribution of the different epigenetic states along the chromatin) suggests that the latter is a very important parameter in the phase behavior. In this chapter, we first explore two different configurations that retain the symmetrical distribution, but varying the size of the domains. This will reveal that the size of the domains affects the width of the bistability region in phase space. In a second exercise, asymmetry in the distribution of the epigenetic states is introduced, which will be shown to result in a second type of bistability. Lastly, a third epigenetic state is introduced, and two different patterns of distribution are explored.

5.1 Changing the size of the domains

Two different configurations with domain sizes of 5 and 20 were used. The polymer size was kept in \( N = 120 \), so the resulting polymers are \((A_5B_5)_{12}\) and \((A_{20}B_{20})_3\).

5.1.1 \((A_5B_5)_{12}\)

Following the characterizations reviewed in the previous chapters, a phase diagram for the polymer with small domains was produced (Figure 5.1). This configuration shows a wide bistability region, persistent through all energies explored. Only at very high density, the polymer assumes a full microphase separation. Again, the bistability consists of a bimodal distribution of the sizes of A and B domains, while the general size of the polymer assumes a unimodal distribution, as evidenced in Figure 5.2.

![Fig. 5.1: Phase diagram for the \((A_5B_5)_{12}\), using density and specific interaction energy strength as phase space parameters.](image)
5. Phase diagram dependence on the configuration

The striking feature of this bistable regime is its persistence. With increasing energies, the population of the fully collapsed state does not increase until a full microphase separation is reached, but it is more equally distributed between swollen and collapsed states. Scatter plots of \((R_g, R_{g,A})\) and \((R_{g,A}, R_{g,B})\) are shown in Figure 5.3. Two clear attractors are identified for the quantities, basically showing complementarity: when the A domains are fully collapsed, the B domains remain swollen, and vice versa.

In turn, the system with the bigger domains shows a narrower bistability region, as evidenced in its phase diagram (Figure 5.4. Although bimodal, the distribution of sizes of the domains does not show the deep valley as in the previous configuration, as depicted in Figure 5.5. This suggests that domains with increased size help stabilize the full collapse of the microphase.
5. Phase diagram dependence on the configuration

Fig. 5.4: Phase diagram for the \((A_{20}B_{20})_3\), using density and specific interaction energy strength as phase space parameters.

Fig. 5.5: Histogram of the radius of gyration for all monomers, and for each monomer type, in the bistability region of the \((A_{20}B_{20})_3\) system.

The plots for \((R_g, R_{g,A})\) and \((R_{g,A}, R_{g,B})\) are very scattered, and show no clear attractors. This essentially suggests that, unlike in the small domain case, the polymer is not constrained to assume either one of two configurations: the individual microphases can be either swollen or collapsed in an increased number of combinations.

5.2 Introducing asymmetry in the size of the domains

We have clear evidence of the dependence of the phase diagram and of the region of bistability, on the size of the epigenetic domains. To further explore this, a system consisting of small domains for the type A monomers, and big domains for the type B monomers was used. The resulting configuration is \((A_5B_{15})_6\). A phase diagram for this system is shown in Figure 5.7, with a very interesting new characteristic: two types of microphase separation.
5. Phase diagram dependence on the configuration

Fig. 5.6: a. Scatter plot of $(R_g, R_{g,A})$ and b. $(R_{g,A}, R_{g,B})$, for the bistability regime of the $(A_{20}B_{20})_3$ configuration.

![Scatter plots](image)

Fig. 5.7: Phase diagram for the $(A_5B_{15})_6$, using density and specific interaction energy strength as phase space parameters.

![Phase diagram](image)

5.2.1 Type 1 MPS

From the distribution of radius of gyration (Figure 5.8), this phase is characterized by a microphase separation consisting of collapsed B domains and swollen A domains. This first MPS supports the idea that bigger domains help stabilize the microphases.
5. Phase diagram dependence on the configuration

5.2.2 Type 2 MPS

This phase is only found at very high densities, much in the spirit of the full microphase separation of the \((A_5B_{15})_{12}\) system. In this case, both the A type and B type domains are fully collapsed (Figure 5.9).

5.2.3 Bistability in the \((A_5B_{15})_6\) system

The bistability region found in this configuration is due to the small A domains, as seen in Figure 5.10. As seen previously in the \((A_5B_5)_{12}\) system, the bistability of the small domains is persistent, even at very high interaction energies.
5. Phase diagram dependence on the configuration

Fig. 5.10: Histogram of the radius of gyration for all monomers, and for each monomer type, in the bistability region of the \((A_5B_{15})_6\) system.

Scatter plots of the quantities \((R_g, R_{g,A}), (R_g, R_{g,B})\) and \((R_{g,A}, R_{g,B})\) are shown in Figure 5.11. While the B monomer radius of gyration assumes a unimodal distribution, the A monomers have two attractors. This also results in two dominant states for the radius of gyration of the whole polymer.

5.3 Introducing a third epigenetic state

In the previous exercises it has been seen that the pattern of epigenetic states in the polymer results in very rich behaviors. Introducing a third type of monomer is consequent with the type of system that we want to model in reality: the epigenetic states are not limited to active and inactive. In eukaryotes, at least 4 principal types of epigenetic marks have been identified [18]. Euchromatin, the active epigenetic state, is characterized by being less condensed, while heterochromatin, highly condensed, is inactive. Heterochromatin can be further divided into at least 3 subtypes: ultra-repressive chromatin, constitutive chromatin (associated to the HP1 protein) and facultative protein (Polycomb). [5] [22]

5.3.1 \((A_{10}B_{10}C_{10})_4\)

The state distribution along the polymer is done in an intercalated way, with equal sizes of domains. The phase diagram for this system is shown in Figure 5.12. The bistable region is now smaller than in the \((A_{10}B_{10})_6\) case, even though the domain sizes are the same.
5. Phase diagram dependence on the configuration

Fig. 5.11: a. Scatter plot of \((R_g, R_{g,A})\), b. \((R_g, R_{g,B})\) and b. \((R_{g,A}, R_{g,B})\), for the bistability regime of the \((A_5B_{15})_6\) configuration.

Fig. 5.12: Phase diagram for the \((A_{10}B_{10}C_{10})_4\), using density and specific interaction energy strength as phase space parameters.
The contact and distance maps (Figure 5.13) show the expected formation of three distinct domains, in the MPS regime. In the bistability region, all three monomers show bimodality in their radius of gyration, while the overall is unimodal (Figure 5.14).

Fig. 5.13: a. Contact probability and b. RMS distance between different regions of the polymer, in logarithmic scale, for the MPS of the \((A_{10}B_{10}C_{10})_4\) configuration.

Fig. 5.14: Histogram of the radius of gyration for all monomers, and for each monomer type, in the bistability phase of the \((A_{10}B_{10}C_{10})_4\) system.
6. THE “LIVING CHROMATIN”

In the previous chapter, we understood clearly that the 1D distribution of states along the polymer can have striking effects on its behavior. But in living organisms, epigenetics is a highly dynamical process, which has to be either very robust (to maintain the same gene expression in specialized cells) or flexible (to turn on or off genes in response to environmental cues). The 3D structure itself is suspected to be one of the mechanisms to induce the formation of domains of the same epigenetic state. This chapter aims to define a methodology for treating the cross talk between 1D and 3D dynamics in the chromatin. Coupling the state transition model (jumping between epigenetic states) with the copolymer lattice model is what composes the “living chromatin” model.

\[
\begin{align*}
\Delta E_{A \rightarrow B} &= \epsilon_B * N_{B\text{neighbors}} - \epsilon_A * N_{A\text{neighbors}} \\
\Delta E_{B \rightarrow A} &= \epsilon_A * N_{A\text{neighbors}} - \epsilon_B * N_{B\text{neighbors}}
\end{align*}
\]

There are two possibilities for this scheme. One is to evaluate both energy terms (interaction strength and coupling) in the same Hamiltonian, to accept or reject a trial move. A second possibility is to evaluate the spatial trial moves using one Hamiltonian (considering the interaction strength), while the state transitions are evaluated with a second Hamiltonian, defined from Eq. 6.1. This means that the system has two temperatures: one that defines the 3D structure of the polymer, and other that defines the 1D pattern of the states. This separation can be useful in exploring different timescales for the polymer dynamics with respect to the spin dynamics.

The system preliminarily explored has interaction energies of \( E_{AA} = 0.5 \), \( E_{AB} = 0.0 \). The recruitment strengths are defined as \( \epsilon_A = 0.5 \) and \( \epsilon_B = 0.0 \). This will imply that neighbors of A type favor the change in spin from A to B, while changing from B to A is not energetically favorable. The time evolution of the contact frequency is shown in Fig. 6.2. It shows indeed a transition from a microphase separation to a single globule, with predominant red spins due to the asymmetric recruitment. For the same initial configuration, 1000 different measurements, each with 1000 MC steps in between are taken, with \( 10N^2 \) initial equilibration time steps. Each time frame has 50 trial spin flips, followed by \( N^2 \) intermediate equilibration steps.

This type of system can be explored with symmetric recruitment rates and symmetric interaction strengths, as parameters of phase space, to understand the conditions under which either very robust or very dynamical epigenetic
domains are formed. A very robust domain formation results in the epigenetic state maintenance through many generations, seen in specialized tissues. A rapidly responding domain can help explain the quick response to the activation or inactivation of group of genes, to environmental cues.

6.2 State transitions in 3D

The state transition framework presented in the introduction chapter is one-dimensional. A three-dimensional state transition can be built by defining two types of recruitment rates:

- $\epsilon_{\text{cis}}\rho_{\text{cis}}$: a transition rate due to the presence of a chemically bonded monomer (along the chain) of a different state.
- $\epsilon_{\text{trans}}\rho_{\text{trans}}$: a transition rate due to the presence of a non-chemically bonded monomer (but a near-neighbor in 3D) of a different state.

This type of model is depicted in Figure 6.3.
Cis recruitment  This type of recruitment is one-dimensional. For a monomer, transition rates are calculated according to the number of chemically bonded monomers belonging to an epigenetic state (0, 1 or 2).

Trans recruitment  An extension of the cis recruitment to three dimensions. The transition rates are calculated according to the number of monomers in the vicinity.

6.2.1 Algorithm

The state transition scheme in the Monte Carlo simulation can be taken as a new type of trial move. Therefore, there are two different phenomena in consideration: polymer dynamics and spin dynamics. These two dynamics can have different rates, this is, an attempt on a monomer spatial or epigenetic state move can be done with varying probabilities, and is an extra parameter for the system. If the move chosen is spatial, the polymer lattice model is invoked. For epigenetic state transitions, the following scheme, inspired by a theoretical analysis of epigenetic cell memory by nucleosome modification [3], is proposed:

First step  Select randomly a monomer.

Second step  Calculate the probabilities for each type of move:

- Cis: $\epsilon_{cis}p_{cis}$.
- Trans: $\epsilon_{trans}p_{trans}$.
- Noisy : $k_0$.

Select the type of move accordingly.

Third step  Select a cis-neighboring or a trans-neighboring monomer, according to the outcome of step 2. If a noisy move is selected, change the state to unmarked.
Fourth step

- If the two monomers belong to the same type, no change is performed.
- If the second monomer is unmarked, no change is performed.
- If the first monomer is unmarked, change it to the state of monomer 2.
- If the two monomers are different, change the first monomer to unmarked.

This type of modeling for coupling 1D and 3D is currently being explored by Cédric Vaillant and Daniel Jost, with promising results on how the interaction strength induces a phase transition, even for low recruitment rates (unpublished data). If no Trans recruitment is considered, the system never undergoes a phase transition: a clear analogy with the Ising model, for which the lower critical dimension is 1. Therefore, there is no formation of stable domains (both spatially and epigenetically). This suggests that the cross-talk mechanism between the distribution of epigenetic marks and the 3D structure of the polymer is indeed essential for the formation and maintenance of epigenetic domains.
7. CONCLUSIONS

7.1 On the lattice model

With a simple, coarse grained model for the chromatin consisting of a homopolymer in an FCC lattice, we were able to recover size scaling laws and dynamical properties that are found either theoretically or in full-space simulations. The idea of using a lattice model instead of the full, continuum model, is mainly due to efficiency. Therefore, the lattice model provides a perfect framework to implement the influence of the 1D pattern in the 3D structure, without being computationally expensive. Excluded volume interactions are imposed in a very simple manner, by preventing the DNA strands from crossing. Interaction energies are added when the condition of having a near neighbor of the same epigenetic state is fulfilled, instead of calculating expensive energetic functions in a continuum space. The system is also valid to take into account bending energies, although they were not further explored in the present work, since it is assumed that the size of a bead is larger than the persistence length of the chromatin.

7.2 On multistability in chromatin folding

Previous attempts to study the genome has essentially treated it as a unidimensional object, while its 3D structure through homopolymer physics. Experimental evidence of the formation of topologically-associated domains (TADs) through Hi-C experiments has inspired physicists to develop models that couple chromatin structure and function. Block copolymer models have been particularly successful to describe chromatin organization in TADs, the microphases treated in this work. This simplified model that puts the copolymer on a lattice also recovers multistable conformations, for which experimental evidence has emerged. The present work aimed to understand better the physical properties of multistability in chromatin folding. We have successfully identified the dependence of the width of the multistability region (using as parameters the specific interaction energy and the density of the system) on the size of domains, and general pattern of distribution of epigenetic states. A manuscript on the subject is soon to be submitted.

7.3 On coupling 1D and 3D

The modularity and simplicity of this model makes it ideal to implement the cross-talk between 1D and 3D, by simply introducing a new type of trial move: changing states. The two approaches that have been preliminarily studied show that the 3D structure helps form and maintain 1D domains, which in turn are known to affect the dynamics of the folding. This type of modelling is very promising, since all previous attempts of describing chromatin folding have treated all epigenetic states as static. Many questions are yet to be answered, and the recent development of new experimental techniques, such as single cell Hi-C experiments, can help identify cell-to-cell variability in chromosome structure as a natural consequence of multistability and feedback between the epigenomic pattern and the chromating folding structure.


