SISSA
Scuola
Internazionale
Superiore di
Studi Avanzati

Scuola Internazionale Superiore di Studi Avanzati

Doctoral Thesis

Bending rigidity, supercoiling and knotting of ring polymers: models and simulations

Author:
Lucia Coronel

Supervisor:
Prof. Cristian Micheletti

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

September 22, 2019

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

"*Inprimisque hominis est propria veri inquisitio atque investigatio.*"

Marcus Tullius Cicero

Contents

Introduction

Polymers are macromolecules made up by repeated simple units called monomers that are strung together with covalent bonds. They can be classified in several ways: homopolymers, composed by the same type of monomers, or heteropolymers when monomers are chemically diverse; they can be linear or circular, as well as branched. Polymers are ubiquitous in everyday life and, like macroscopic ropes, can become spontaneously entangled. In the case of linear polymer chains, knots can be untied by a suitable motion of the filaments and this non permanent entanglement is referred to as physical knots. On the other hand, if the polymer ends are joined, the geometrical self-entanglement becomes trapped permanently in the form of a proper mathematical knot, whose topology cannot be changed by any geometrical rearrangement of the polymer except by cutting it.

Spontaneus knots have been reported in biopolymers too. Although no knots have yet been found in naturally occurring RNAs [1], they have been documented in proteins [2–10] and in DNA filaments, both *in vivo* and *in vitro* [11–23]. Indeed, it has long been known that genomic DNA can be highly entangled due to the high packing degree that it attains in all organisms, from eukaryotic chromosomes [24] down to minimalistic viral genomes [15, 25]. For instance, the μm-long genome of bacteriophages is packed inside capsids having diameter of about 50 nm. Such tight packing has been shown to introduce abundant and complex DNA knots [26, 27].

In addition, it is also known that the incessant DNA processing involved by replication, transcription and recombination can lead to the (statistically inevitable) emergence of DNA knots (and links too) [28, 29].

On the other hand bacterial DNA, which is circular, has to be kept unknotted and unlinked at the very end of the replication; failing to do so would be fatal for the cell as the replicated genetic material would not be separable into the two daughter cells.

More in general, because of the intimate connection between structure and function in biomolecules [30], there is the biological necessity to maintain a tight control on the degree of entanglement of these biomolecules *in vivo*. Therefore, many efforts have been spent to understand both the mechanisms leading to the emergence of knots in biomolecules as well as their impact on biological functionality.

In bacteria and eukaryotic cells, the detrimental consequences of DNA entanglement are avoided via active mechanisms, based on topoisomerase enzymes, to remove DNA knots and links [13, 28, 31–34].

Detecting DNA topological state is most commonly achieved via gel electrophoresis that can efficiently separate charged polymers having different length, molecular weight, level of supercoiling and topology [35–37]. Note however that such topological profiling is limited of DNA lengths $\leq$ 12 Kbp. Recent breakthroughs in single-molecule techniques have pushed this upper bound to 150Kbp, though they are not standarly used yet [38].

The experimental efforts that, over the decades, have provided such a comprehensive picture of biopolymer knotting, have been complemented by an equally intense theoretical and computational activity. This has been aimed at understanding both the incidence of knots in specific proteins or DNA systems, as well as the more general underlying physical principles.

For the latter, we recall that several studies have systematically investigated how the intrinsic properties of polymers affect their equilibrium knotting probability.

Arguably, the first and best known of such studies is the seminal work of Frisch, Wasserman and Delbruck [39], that first conjectured how the knotting probability increases with a polymer's contour length. Later studies have addressed the impact on entanglement of spatial confinement as well as of the thickness of the polymer itself (an effect relevant for DNA where it can be modulated by the concentration of counterions) [21, 40–43].

In addition to this, several other works have taken a complementary standpoint, and examined how the presence of knots influences, in turn, the physical properties of polymers, such as the metric and dynamical ones [44, 45].

Interestingly, one general aspect that has so far eluded an in-depth investigation is

how the knotting probability is affected by a polymer's bending rigidity at fixed contour length.

This open questions has motivated the first part of the work that is presented in this thesis where, starting from a general model of flexible chains I systematically consider how the knotting probability depends the bending rigidity. As I discuss, the results is quite surprising: as one goes from the fully flexible to the rigid case, the decreasing compactness of the chains is not accompanied by an analogous decrease of their knotting probability. Consequently, the equilibrium incidence of knots is strongly non-monotonic.

The second part of my thesis is motivated by the intriguing interplat of DNA knots and supercoiling. These, in fact, can both be simultaneously present in bacterial DNA in vivo, and as already mentioned, can impact the funtionality of the cell. In earlier work [46] this interplay has been addressed for the simplest knot types. Here, I will extend consideration to more complex knots, and in addition use molecular dynamics simulations and oxDNA [47], an accurate mesoscopic DNA model, to get a detailed structural and dynamical insight too. In this case too, the results were unexpected: essential tangles of the supercoiled DNA knot are very long-lived, and remain persistent over timescales that are much longer than the metric relaxation times.

The overall organization of the thesis is as follows:

In chapter 1, I introduce a general coarse-grained model for semiflexible rings, as well as the Monte Carlo method that I used to sample its canonical conformational space. Secondly, I describe the oxDNA model and the Langevin dynamics used to simulate it.

In chapter 2, I provide a brief introduction to knot theory, which sets a reference for concepts used in the thesis, describe how we detect knot types and locate knotted portions in closed chains. I will also briefly survey concepts related to link theory and on which we rely to profile of properties of supercoiled dsDNA.

In chapter 3, I present a computational and theoretical study to explore how the knotting probability of ring polymers depends on its bending rigidity. As anticipated, I

found a non-monotonic profile for the knotting probability, as well as for the length of the knotted region. I will also discuss how thermodynamic reweighting techniques can be used to compute the free energies of knotted and unknotted rings, their entropic and energetic contributions and to explore how the enthalpic-entropic interplay is regulated by the bending rigidity. The insight will give a better understand of the non-monotonic behaviour from a physical perspective.

In chapter 4, I will build on the results of chapter 3, and extend them with an approximate analytical approach. To do so, I will introduce and use a heuristic mapping between semi-flexible chains and self-avoiding polygons, another widely used polymer model. As I discuss, the mapping confirms and extends the results on knotting probability for previously unexplored combination of chain lengthh and bending rigidities. In particular up to length ≥ 20000 the knotting probability behaviour switches from unimodal to bimodal.

In chapter 5, I move on the second main topic and discuss the effect of entanglement and supercoiling in DNA rings. The analysis relies in part on the capability of identify plectonemically-wound DNA region, and I will accordingly describe the algorithmic approach that I developed to thhis purpose.As I discuss, the presence of topological constraints, together with supercoiling, due to torsional stress, affects very significantly the dynamics of the ring. In particular, when both are present they freeze, for time-scale larger then our simulation time, the diffusion of knotted and plectonemically-wound regions along the ring, an effect with possible implications for topoisomerases action.

In chapter 6, I study another type of interplay between supercoiling and knots., motiveted by the recent breakthrougs results [24], that showed that eukaryotic DNA can be knotted *in vivo*. The result also motivated a collaborative project with the group of prof. Roca of Structural Biology Unit of Barcelona, with aiming of studying the effect of positive supercoling on the knotting of yeast minichromosomes. We show, via computer simulation and using a simple model, that the level of compactness comparable with that expectedly introduced by positive supercoiling, can enhance

the knotting probability and complexity similarly to the experimental results.

Finally, in chapter 7, I will overview the aforementioned results for a common perspective and outline possible directions for future studies.

The above results have been presented in the following publications, on which this thesis is based:

- L. Coronel, E. Orlandini, C. Micheletti.
  Non-monotonic knotting probability and knot length of semiflexible rings: the competing roles of entropy and bending energy
  Soft matter, 13 (23), 4260-4267
- L. Coronel, A. Suma, C. Micheletti.
  Dynamics of supercoiled DNA with complex knots: large-scale rearrangements and persistent multi-strand interlocking
  Nucleic Acids Research, Volume 46, Issue 15, 7533-7541
- E. Uehara, L. Coronel, C. Micheletti, T. Deguchi.
  Bimodality in the knotting probability of semiflexible rings suggested by mapping with self-avoiding polygons
  REACT, Volume 134, Pages 141-149
- A.Valdés, L. Coronel, B. Martínez-García, J.Segura, S.Dyson, O.Diaz-Ingelmo, C. Micheletti, Joaquim Roca.
  Transcriptional supercoiling boosts topoisomerase II-mediated knotting of intracellular DNA
  Nucleic Acids Research, Volume 47, Issue 13, 26 July 2019, Pages 6946-6955.

Models and methods

1.1 Introduction

In this chapter I describe the main models and methods that are used in the remainder of the thesis. The material is organized as follows: first, I present a general coarse-grained model for semiflexible chains and the Monte Carlo method used to sample its canonical conformational space, as discussed in Chapter 3. Next, I will overview oxDNA, an accurate mesoscopic DNA model and the molecular dynamics simulation that have been used for the study presented in Chapter 5.

1.2 Coarse-grained model for semiflexible rings

1.2.1 Model

To explore the impact of a polymer's bending rigidity on its equilibrium knotting probability, we resorted to a general model of rings on N touching spherical beads. We will indicate with $\vec{b}_i$ the bond vector connecting the centers of consecutive touching beads, see Fig. 1.1, and with b their length, which is equal to the beads diameter. The potential energy of the ring consists of two terms:

$$U = U_{e.v} + U_{bend} \tag{1.1}$$

The first term, $U_{e.v}$, accounts for the excluded volume interactions: we assign infinite energy to configuration when overlaps of beads occur and zero-energy otherwise.

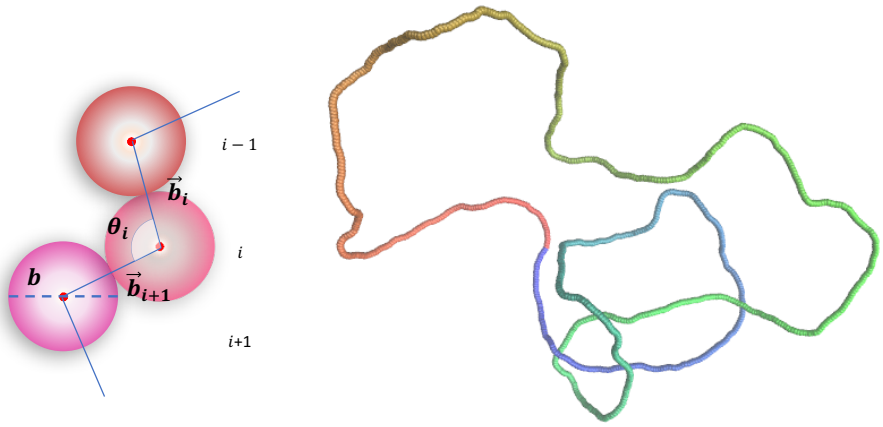


FIGURE 1.1: Typical snapshot of a ring of beads and insert with detail of the beads.

The second term is the bending energy and is defined as:

$$U_{bend} = \kappa_b B = \kappa_b \sum_i \left(1 - \frac{\vec{b}_i \cdot \vec{b}_{i+1}}{b^2}\right) = \kappa_b \sum_i (1 - \cos(\theta_i)) \tag{1.2}$$

where κ_b is the bending rigidity coefficient, B the total bending, and θ_i is the angle between two consecutive bond vectors, see Fig. 1.1.

The model corresponds to a discretised self-avoiding worm-like chain or, equivalently, a self-avoiding Kratky-Porod chain. Despite its simplicity it is not analytically-soluble, because of the non-local effects introduced by the excluded-volume interactions.

If we neglect the latter, and retain only the local bending potential, the model reduces to a freely-jointed chain (FJC). This model is exactly soluble and, as we discuss hereafter, it allows for an exact calculation of the bond-orientation correlation function. In the FJC model, such correlation is obtained by computing the expectation value of the scalar product of two consecutive bonds:

$$\langle \vec{b}_i \cdot \vec{b}_{i+1} \rangle = \frac{1}{Z} \int \prod_{i=1}^{N-1} d\vec{b}_i d\vec{b}_{i+1} \vec{b}_i \cdot \vec{b}_{i+1} e^{-\beta U_{bend}} \tag{1.3}$$

where $\beta = 1/K_B T$, with K_B the Boltzmann constant. It is possible to define the partition function as:

$$Z = \int \prod_{i=1} d\cos(\theta_i) e^{-\beta\kappa_b(1-\cos(\theta_i))} = \prod_{i=1} \int_{-1}^{1} d\cos(\theta_b) e^{-\beta\kappa_b(1-\cos(\theta_b))} = \prod_{i=1} Z_i \tag{1.4}$$

$$Z_i = e^{-\beta\kappa_b}\frac{2\sinh(\beta\kappa_b)}{\beta\kappa_b} \tag{1.5}$$

Therefore, considering eq. 1.5 and eq. 1.3:

$$\langle\vec{b}_i\cdot\vec{b}_{i+1}\rangle = b^2\langle\cos(\theta_i)\rangle = \frac{b^2}{Z}\int\prod_{i=1}^{N-1} d(\cos\theta_i)\cos\theta_i e^{-\beta\kappa_b(1-\cos(\theta_i))} \tag{1.6}$$

$$= \frac{b^2}{Z_i}\int_{-1}^{1} d(\cos\theta_i)\cos\theta_i e^{-\beta\kappa_b(1-\cos(\theta_i))} \tag{1.7}$$

$$= \frac{b^2}{Z_i}\left(\frac{\partial Z_i}{\partial(\beta\kappa_b)} - Z_i\right) \tag{1.8}$$

$$= b^2\left(\frac{\partial log(Z_i)}{\partial(\beta\kappa_b)} - 1\right) \tag{1.9}$$

Finally, using the Langevin function $L(\beta\kappa_b) = \coth(\beta\kappa_b) - \frac{1}{\beta\kappa_b}$, eq. 1.3 can be written as:

$$\langle\vec{b}_i\cdot\vec{b}_{i+1}\rangle = b^2 L(\beta\kappa_b) \tag{1.10}$$

Because the bending angles are statistically independent, the correlation function of two bonds at sequence separation l can be factorised, yielding an exponential decay:

$$\langle\vec{b}_i\cdot\vec{b}_{i+l}\rangle = b^2(L(\beta\kappa_b))^l \approx b^2 e^{-l\cdot b/l_p} \tag{1.11}$$

The decay length, l_p, is termed the *persistence length*, and it is equal to:

$$l_p = -b/\log\left[\coth\left(\frac{\kappa_b}{K_BT}\right) - \frac{K_BT}{\kappa_b}\right] , \tag{1.12}$$

which, in the limit of $\kappa_b \gg K_BT$, simplifies to

$$l_p = b\kappa_b/K_BT . \tag{1.13}$$

The persistence length is often used to define the Kuhn length, $l_{Kuhn} \equiv 2\,l_p$, which is the bond length of a, equivalent random walk having the same contour length and mean square gyration radius (see below) of the FJC.

1.2.2 Metric Properties

Gyration Radius The quantity that we used to characterize the typical size of the polymer is the mean square radius of gyration, defined as:

$$R_g^2 \equiv \frac{1}{N+1} \sum_i (\vec{r_i} - \vec{r}_{CM})^2 \tag{1.14}$$

where $\vec{r}_{CM} \equiv \frac{1}{N} \sum_j \vec{r_j}$, is the position of the center of mass. In general, R_g^2 scales asimptotically as N^ν, where ν is called Flory exponent, which is equal to 0.5 for random walks and is approximately 0.59 for self-avoiding ones in three-dimensions [48].

1.3 Monte Carlo

We model ring polymers as semi-flexible rings of beads, described by the potential energy 1.1. In order to study the influence of stiffness on the topological properties, we explore different combination of length N and bending rigidity κ_b. This allows to study the system from fully-flexible to rigid polymers. For each combination, we generate a set of independent and equilibrated configurations using a Monte-Carlo technique, as detailed below.

Sampling The Monte-Carlo (MC) method allows one to generate a series of microstates of a statistical ensemble. Starting from a given state Γ_A with energy E_A, a new state Γ_B is stochastically chosen. The new configuration will be accepted or rejected following a criterion satisfying detailed balance:

$$\Pi_{A\to B} P_A = \Pi_{B\to A} P_B \tag{1.15}$$

where $P_{A(B)}$ is the canonical probability of occurrence of state $A(B)$ and $\Pi_{A\to B}$ is the transition rate from A to B. In actual numerical implementations of the Monte Carlo sampling scheme the transition rate subsumes two distinct algorithmic steps: a stochastic modification of state A to generate the new state, B, also termed the trial state, and the stochastic acceptance or rejection of the newly-generated state.

Accordingly, the transition rate can be written as:

$$\Pi_{A\to B} = t(A \to B) \cdot a(A \to B) \ , \tag{1.16}$$

where $t(A \to B)$ is the probability to generate trail state B from A, and $a(A \to B)$ is the probability to accept it as the new system state. The detailed balance condition of eq. 1.16 becomes:

$$\frac{P_B}{P_A} = \frac{t(A \to B)a(A \to B)}{t(B \to A)a(B \to A)} \ . \tag{1.17}$$

For systems in canonical equilibrium, as in our case, P_A is given by

$$P_A = \frac{e^{-\beta E(A)}}{\mathcal{Z}} \tag{1.18}$$

where $\beta = \frac{1}{K_B T}$, $E(A)$ is the energy of state A and $\mathcal{Z}$ is the partition function. Hence, eq. 1.17 yields the constraint that has to be satisfied by the acceptance rates, a, in order to respect detailed balance:

$$\frac{a(A \to B)}{a(B \to A)} = \frac{P_B t(B \to A)}{P_A t(A \to B)} = \frac{t(B \to A)}{t(A \to B)} e^{-\beta(E_B - E_A)} \ . \tag{1.19}$$

This expression simplifies if, as in our case, the stochastic deformation rules used to generate state B are independent of A. Under such conditions, $t(B \to A) = t(A \to B)$, and the detailed balance condition reduces to:

$$\frac{a(A \to B)}{a(B \to A)} = e^{-\beta(E_B - E_A)} \ . \tag{1.20}$$

The above relationship leaves some freedom of choice for the acceptance rates between two states, as it only specifies their ratios. A common choice is the well-known Metropolis criterion:

$$a(A \to B) = \min\left(1, \frac{t(B \to A)P_B}{t(A \to B)P_A}\right) \tag{1.21}$$

which means that if the ratio is bigger than one the new state B is accepted, otherwise is accepted with probability equal to the ratio of the canonical probabilities. The latter condition is implemented by extracting a uniformly-distributed random number

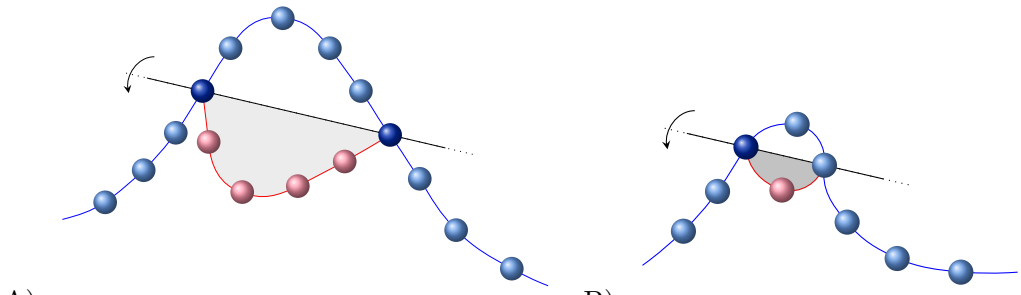


FIGURE 1.2: Example of crankshaft (A) and local crankshaft (B) in a ring of beads. In this representation, the beads are smaller than necessary and disconnected for the only purpose of visual clarity.

$r \in [0, 1]$ and accepting the trial configuration if $r < e^{-\beta(E_B - E_A)}$. Otherwise, the previous configuration is retained, and the Monte Carlo time is incremented.

Crankshaft Moves Because we are dealing with closed rings, our deformation rules consist of the so-called crankshaft moves or, better, of a combination of global and local crankshaft moves, see Fig. 1.2. These moves, in fact, preserve bond lengths and guarantee the ergodic sampling of conformational.

The global moves involve picking randomly any two distinct beads of the ring, and rotating one of the resulting arcs by a random angle $\alpha \in [0, 2\pi]$ around the axis joining the picked beads' centers. The local moves differ for the choice of the beads, which is restricted to beads at sequence separation equal to 2. For every global move we perform a sweep of N local moves. The combination of local and global moves allows for an efficient sampling of conformational space as it balances the pros and cons of the two types. In fact, global moves are advantageous because they can alter significantly a conformation, which is desirable for the breadth of visited conformational space, but at the same time it can excessively increase the energy and hence be frequently rejected; the converse is true for local moves [49]. We apply these moves, within the Metropolis Monte Carlo scheme, starting from a perfectly circular ring conformation.

Note that, consistently with the ergodicity condition, crankshaft moves do not necessarily preserve the initial topology (knot type) of the rings, as they allow for bond crossings (though the sampled conformations are necessarily self-avoiding).

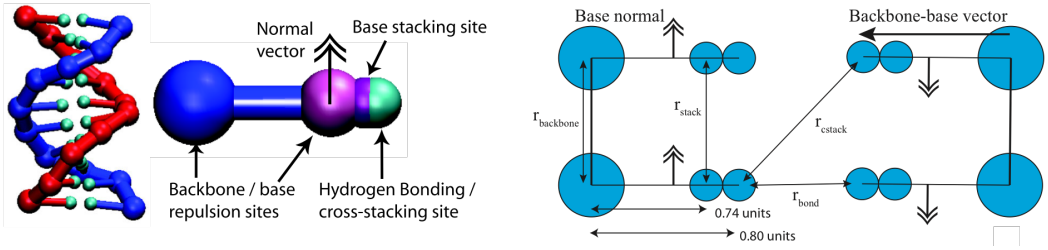


FIGURE 1.3: Model representation of a duplex. Image from [52].

Autocorrelation Time The duration of the Monte Carlo simulation was set from the target number of *independent* conformations we wished to sample. For the characteristic time-interval separating independent conformations, we used the autocorrelation time of the radius of gyration, which we computed in preliminary runs. We recall that the autocorrelation function of a general observable x is given by:

$$C_x(\tau) = \frac{\langle (x(t) - \overline{x})\,(x(t+\tau) - \overline{x})\rangle}{\sigma^2} \approx e^{-\tau/\tau_c} \tag{1.22}$$

where $\overline{x}$ is the average on the ensemble of x, and the variance is $\sigma^2 = \langle (x(t)\overline{x})^2\rangle$. The characteristic autocorrelation time, τ_c, is given by the value where C_x drops to $1/e$.

1.4 Coarse-grained model for double strands DNA rings

In the second part of the thesis we studied the interplay of supercoling and knotting in DNA rings. For such investigation we used oxDNA [47, 50, 51], an accurate mesoscopic model that allows for representing each of the two strands in dsDNA filament. This, in turns, makes it possible to account for torsional effects in a transparent and straightforward manner (unlike when e.g. a single chain of beads is used to model dsDNA).

The oxDNA is a mesoscopic coarse-grained model of DNA where each nucleotide is represented by three beads, one for the backbone and the other two for the base and sugar groups, see Fig. 1.3. The relative positions of the three interaction centers are fixed and hence the beads form a rigid unit with precise geometry.

Potential Energy Properties such as chain connectivity, bending rigidity, pairwise, stacking and screened electrostatic interactions are all included in the potential

energy:

$$V = \sum_{nn} \left(V_{backbone} + V_{stack} + V_{exc}^{'} + \right) + \sum_{otherpairs} \left(V_{HB} + V_{c_{stack}} + V_{exc.} \right) \qquad (1.23)$$

In this expression the index nn is related to the nearest-neighbours on a strand. Quoting verbatim from ref. [50], the potential energy terms are described as follows:

> $V_{backbone}$ is a finitely extensible non linear elastic (FENE) spring [...], with an equilibrium length of 6.4Å representing the covalent bonds which hold nucleotides in a strand together.
> V_{stack} represents the tendency of bases to form coplanar stacks: it is a smoothly cut-off Morse potential between base- stacking sites, with a minimum at 3.4Å. It is modulated by angular terms which favour the alignment of normal vectors, and the alignment of the normal vectors with the vector between stacking sites. As such, the interaction encourages coplanar stacks, separated by a shorter distance than the equilibrium backbone length, leading to helical structures. Right-handed helices are imposed through an additional modulating factor which reduces the interaction to zero for increasing amounts of left-handed twist.
> V_{exc} and $V_{exc}^{'}$, representing the excluded volume of nucleotides, prevent the crossing of chains and provide stiffness to unstacked single strands. The lack of explicit angular or dihedral potentials along the backbone allows single strands to be extremely flexible. For non-nearest neighbours, smoothly cut-off (and purely repulsive) Lennard-Jones interactions are included between all repulsion sites on the two nucleotides. For nearest neighbours, the backbone/backbone site interaction is not included because the distance between sites is regulated by the FENE spring.
> V_{HB}, representing the hydrogen bonds which lead to base pairing, is a smoothly cut-off Morse potential between hydrogen-bonding sites, modulated by angular terms which favour the anti alignment of normal vectors and a collinear alignment of all four backbone and hydrogen-bonding sites. V_{HB} is set to zero unless the two bases are complementary (AT or GC). Together with V_{stack}, V_{HB} causes the formation of antiparallel, right-handed double helices for complementary strands.
> $V_{c_{stack}}$ represents cross-stacking interactions between a base in a base pair and nearest-neighbour bases on the opposite strand, providing additional stabilization of the duplex. We incorporate it through smoothed, cut-off quadratic wells, modulated by the alignment of base normals and backbone- base vectors with

the separation vector in such a way that its minimum is approximately consistent with the structure of model duplexes.

The parameters of the potential were optimally set so to reproduce measurements from thermal denaturation, unzipping and stretching experiments. In particular, they reproduce the pitch ($10.34bp$ per turn), and persistence length ($\sim 50nm$) of B-form DNA at room temperature [50].

1.5 Molecular dynamics

The dynamical evolution of oxDNA rings was integrated numerically within a Langevin dynamics framework.

We recall that the Langevin dynamics of a particle is described by the following stochastic differential equation:

$$m\ddot{\vec{r}} = -\gamma\dot{\vec{r}} - \nabla U + \vec{\eta} \tag{1.24}$$

where m is the mass of the particle, $\vec{r}$ is its vector position, γ is the friction coefficient, and $\vec{\eta}$ is a Gaussian noise term. The friction and noise terms both account for the solvent, which is not modelled explicitly. On the one hand, the particle motion through the solvent is dissipative (an effect captured by the friction term) and, on the other hand, the solvent acts as a heat-bath for the particle (an effect captured by the noise terms). The connection between these two different solvent-induced effects is embodied by the fluctuation-dissipation relationships that γ and $\vec{\eta}$ have to satisfy:

$$\langle \eta_{i,\alpha}(t) \rangle = 0 \tag{1.25}$$

$$\langle \eta_\alpha(t)\eta_\beta(t\prime) \rangle = 2K_B T\gamma\delta_{\alpha\beta}\delta(t-t') \tag{1.26}$$

where α, β represent the Cartesian coordinates; $\delta(t-t')$ the Dirac delta which means that the force at time t is uncorrelated with the force at time $t' \neq t$. For a multi-particle system, the above equations are complemented by the condition that the noise terms of different particles are independent at all times.

The integration of the Langevin dynamics for the oxDNA system was carried out

within the LAMMPS simulation package using the default model parameters for the particle's masses and the friction coefficient [53, 54].

A compendium of Knot Theory

2.1 Introduction

In this section I provide a basic introduction to knot theory, focusing of those aspects that are of direct relevance to the topics presented in this thesis. We are all familiar with knots that spontaneously arise in ropes, cables and filaments. Such entanglements affect, in turn, the physical properties of the strings - such as the resistance to rupture or sliding friction - and, in fact, they are purposedly used for various tasks, from binding multiple ropes, to producing textile structures or simply used as ornaments.

2.2 Definition of a knot

In topology, a knot is an embedding of a closed curve in 3-dimensional Euclidean space. Figure 2.1 shows a series of knots of increasing complexity starting from trivial one, a planar circle, which is termed the unknot.

All knots in Figure 2.1 belong to different topological classes, as they cannot be converted from one type to another through a series of *ambient isotopies*. The latter are manipulations of the space in which the knot is embedded and correspond to continuous deformation of a knotted string except for cutting it, passing it through itself, see Fig. 2.3 or that shrink the curve to a point, see Fig. 2.2.

Instead of considering curves in 3-dimensional space, is is often convenient to restrict considerations to their planar representations, the so-called *knot diagrams*, where crossings are graphically represented by breaks in the underpassing portions.

Of particular importance are the *minimal diagrams*, which are obtained when the number of crossings is the smallest possible for the given knot topology. All knots in

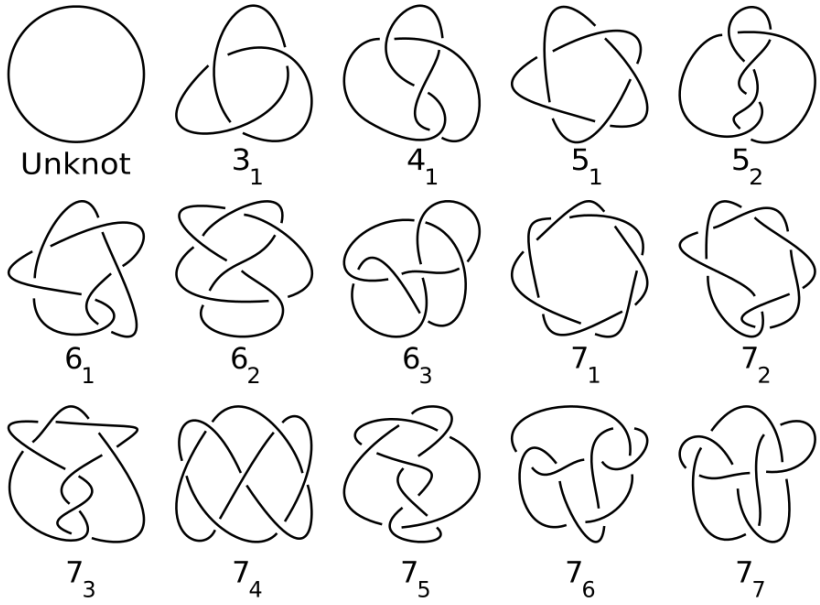


FIGURE 2.1: Knot table. Image taken from wikipedia.org

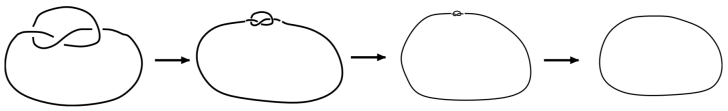


FIGURE 2.2: Image taken from ref. [55].

Fig. 2.1 are represented by minimal diagrams.

It is also possible to assign an orientation to a knot diagram and hence associated a conventional sign to each crossing based on the right-hand rule applied to the oriented over- and under-passing segments, see Fig. 2.4.

Reidemeister moves The deformations of a knotted curve through ambient isotopies are clearly reflected as deformations of the projected planar diagram. It can be shown that all such diagrammatic deformations can be described as a succession of three types of (topology-preserving) elementary moves, termed *Reidemeister moves*:

- twist an arc in either direction to produce a loop
- move one arc over another
- move a strand completely over/under a crossing ,

as shown in Fig. 2.5.

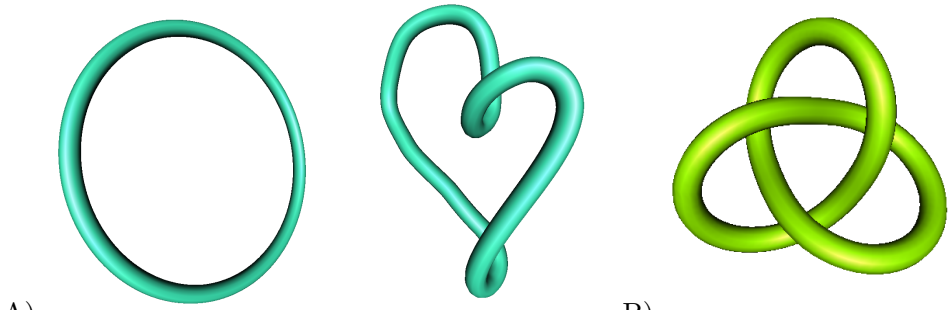


FIGURE 2.3: A). Trival knots 0_1: they are both unknots, it is evident that the ring can be easly obtain by deforming the second one. The B) panel shows the first non-trivial knot, with non trivial crossings, that is called trefoil knot, 3_1. It is not possible to obtain the unknot from the trefoil without cutting the curve. Most of the figures of knots of this chapter are made using the Knotplot software [56]

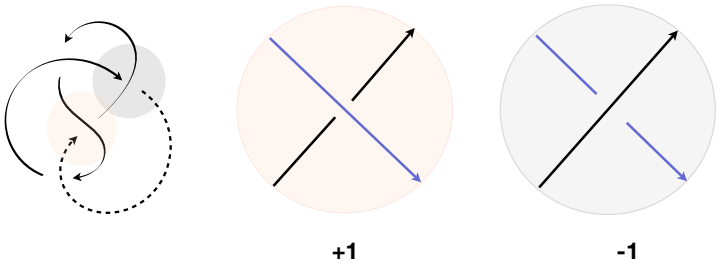


FIGURE 2.4: Positive and negative crossings defined with the right-hand rule.

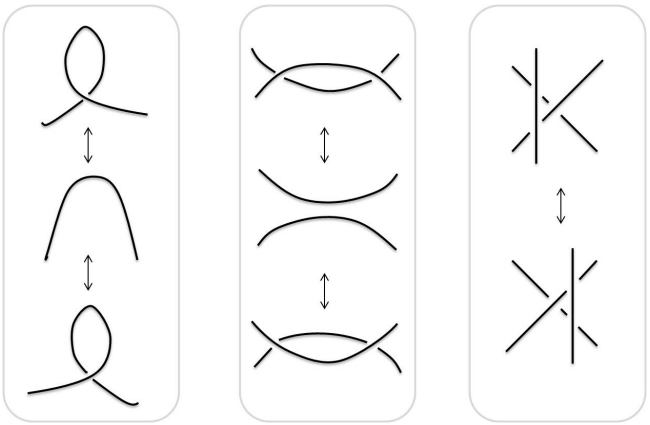


FIGURE 2.5: Reidemeister moves, in order from left to right: 1) Twist: for a twist in the knot its possible to put in or take out the twist; 2) poke/unpoke: its possible to add two crossings or remove two crossings; 3) slide: its possible to slide a strand of the knot from one side of a crossing to the other side of the crossing.

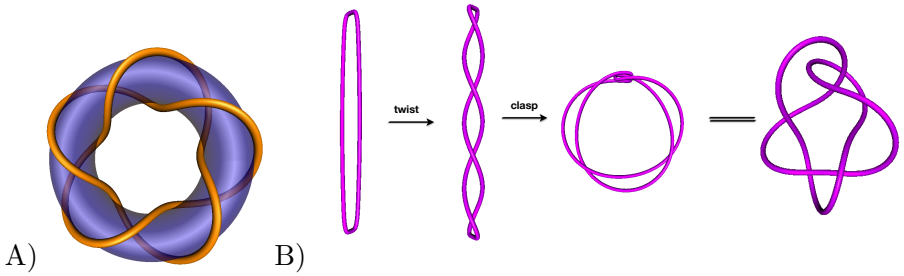


FIGURE 2.6: 5_1 and 5_2 knots shows as A) torus knot, and B) Twist knot, respectively.

Classification of knots Customarily, the number of crossings in the minimal diagram, the *crossing number*, is almost universally as the primary index of knot complexity. All knots in Fig. 2.1 are, in fact, ranked for increasing crossing number. In the standard nomenclature, knots are labelled by two integers, C_i, where C is the crossing number and while i is a conventional enumerative index introduced to distinguish different knots with the same nominal complexity, i.e. for a knot of $C = 5$ crossings we have two topologically different knot, which are indicated as $5_1, 5_2$, see Fig. 2.1.

The 5-crossings knots shows also how knot of same complexity can be classified in different families, for instance, based on notable or distinctive features of their minimal representation: the *torus* knots, which include all prime knots that can be drawn as a simple closed curve on the surface of a torus, and the *twist* knots, which includes all knots that can be obtained by repeatedly twisting an unknotted ring and then clasping the two ends together, see Fig. 2.6.

Writhe We introduced the diagrammatic representation of knots, the concept of oriented curve and signed crossing. Another informative geometric observable is the writhe, Wr, which is the sum of the sign over all the crossings and averaged over all projections.

Heuristically, the writhe is a measure of how much a curve wraps around itself. It is a geometric quantity because it depends on the specific geometry of the curve at hand, and not only on its knot topology. Indeed, consider that for a given projection, the sum of signed crossings is unaffected Reidemeister moves Type II and Type III, but Reidemeister moves of Type I increases or decreases it by 1. The rigorous definition

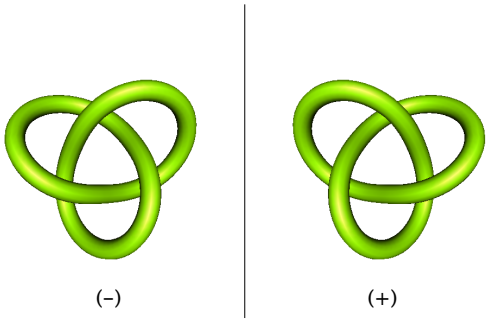


FIGURE 2.7: A trefoil knot 3_1 and its mirror image. The two enantiomers are left-handed and right-handed

.

of writhe is give by the Gaussian integral of a curve C:

$$Wr = \frac{1}{4\pi} \int_C \int_C d\vec{r_1} \times d\vec{r_2} \cdot \frac{\vec{r_1} - \vec{r_2}}{|\vec{r_1} - \vec{r_2}|^3} \tag{2.1}$$

where $\vec{r_1}, \vec{r_2}$ are points on C. In Chapter 6, where we model yeast minichromosomes with a beads-on-a-string model, we shall compute the writhe from the average sum of the crossings sign. Instead, in Chapter 5, where we use an explict double-stranded representation for DNA, we will naturally compute the writhe through an approximation of the Gauss integral done by Langowski and Klenin [57].

Chirality Knots can also be chiral: a property for which there exist no deformation that brings an embedding of the knot to its mirror image. The two images of the knot are called enantiomers, see Fig. 2.7. If such ambient isotopy exists, the knot is called amphichiral, an example is shown in Fig. 2.8. Amphichiral knots are relatively rare, indeed among the ones show in Fig. 2.1 the $0_1, 4_1, 6_3$ are amphichiral.
For prime knots up to 7 crossings, the enantiomers can be distinguished by computing the balance of left- and right-handed crossings in the minimal diagram. Notice that the standard knot nomenclature does not distinguish between the enantiomers of chiral knots.

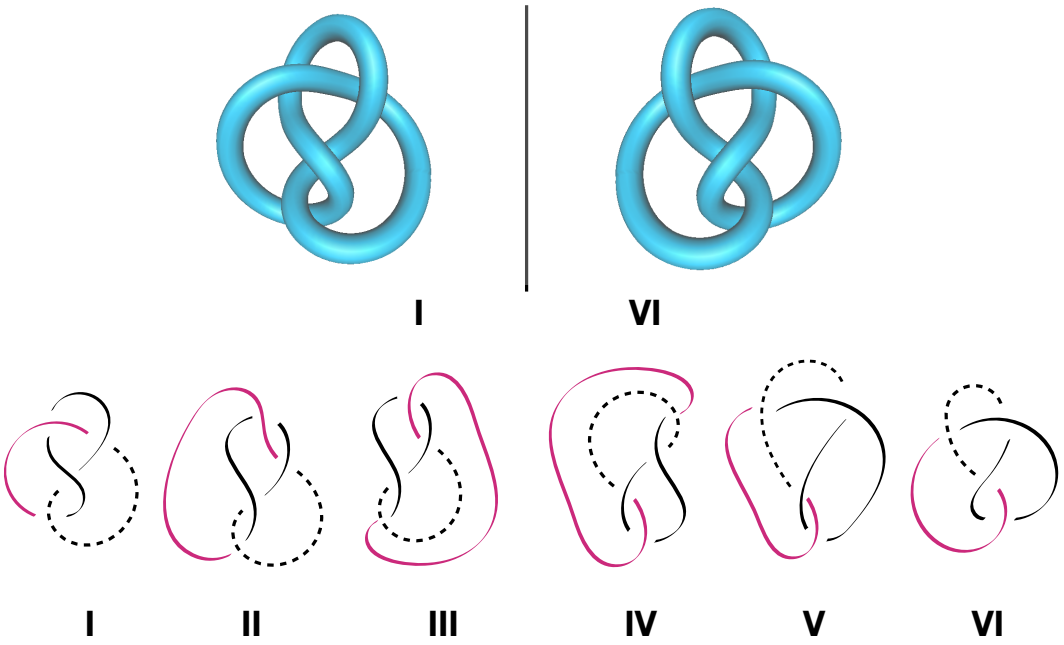


FIGURE 2.8: Example of amphichiral knot 4_1. Step I $\rightarrow$ II: deformation of the purple arc; step II$\rightarrow$ III: flip of the purple arc; step III$\rightarrow$ IV: rotation of 180°; step IV$\rightarrow$V: deformation of the dotted line and black one; step V $\rightarrow$ VI: manipulation of all the curves.

2.3 Topological invariants

In principle, the knotted curve could be classified by using ambient isotopies in three-dimensions (or Reidemeister moves on its diagram) to manipulate it and establish its topological equivalence with a set of reference knot types, such as those in Fig. 2.1. Such procedure, however, is non-deterministic and hence of limited, or no practical use. Most commonly, knots are classified by computing topological invariants.

Topological invariants are fingerprints of knots, they are the same for all knots of a given class of equivalence, and they represent a powerful tool to analyse efficiently the topology of thousands of configurations generated in numerical contexts. In particular, we mainly use the *Alexander polynomial* as topological invariant.

Alexander polynomials The Alexander polynomial is defined for a knot diagram in terms of a single variable t and can be computed using the following algorithm:

- Assign an orientation to the diagram and for each crossing establish the sign following the right-hand rule.

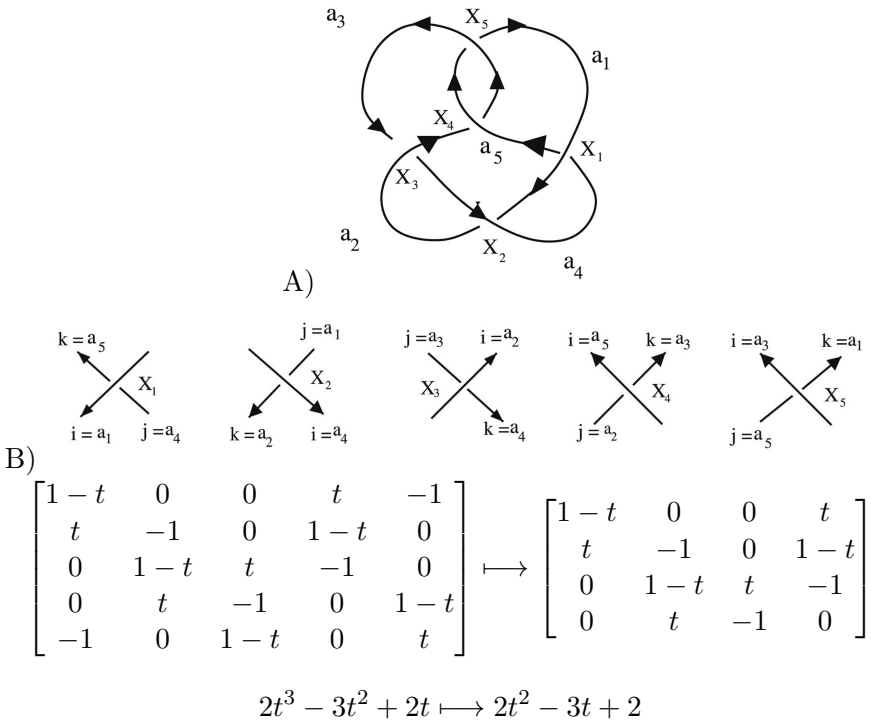
- Assign a progressive numbering index to the n arcs, going from an under- crossing to the next one, and to the n crossings in the diagram, see Figure 2.9.

- Define the matrix M with n rows corresponding to the number of crossings and n columns in accord to the number of arcs. To compute the entries of the matrix consider the xth crossing and rename the three arcs involved as i, j and k in the assumption that the ith arc passes over the others, then:
  if the crossing x is positive: $M(x,i) = 1-t, M(x,j) = -1, M(x,k) = t$, and the remaining elements equal to zero;
  if the crossing x is negative: $M(x,i) = 1-t, M(x,j) = t, M(x,k) = -1$, and the remaining elements equal to zero;

- Iterate for all the crossings.

Consider now the minor of M: any minor of order $n-1$ of the matrix M obtained by deleting any one of the rows and columns. This minor of M is the so-called *Alexander polynomial* $\Delta(t)$. An example is shown in Fig. 2.9.
Since the size of the matrix and hence the determinant depend on the number of crossings and therefore on the details of a given diagrammatic representation, the Alexander polynomial is not uniquely defined for a given knot type. However, the polynomials of two diagrams of a knot can differ by a quantity $\pm t^m$,$m \subset Z$. Therefore, multiplying by $\pm t^m$ will define the (irreducible) Alexander polynomial $\Delta(t, C_i)$.
The irreducible Alexander polynomial is independent of the knot diagram used to calculate it. As in many computational studies, to identify the knot type, we compute the Alexander polynomial values in $t = -1$ and $t = -2$, $\Delta(-1; C_i)$ and $\Delta(-2; C_i)$. Note that that these two values depend on the knot representation respectively by a factor $(-1)^m$ and $(-2)^m$ (as a consequence of the $\pm t^m$ dependence mentioned above). Finally, it is important to stress that exchanging positive and negative crossings for all crossings, does not change the resulting Alexander polynomial, which is unable to distinguish between two enantiomers of a chiral knot.

Notice that, even if, the Alexander polynomial is the same for each knot of the class, so no move can change it, there is no guarantee that knots that are genuinely different must have a different polynomial. Indeed, such degeneracy does occur for the Alexander polynomial and other types of invariants too.



$$\begin{bmatrix} 1-t & 0 & 0 & t & -1 \\ t & -1 & 0 & 1-t & 0 \\ 0 & 1-t & t & -1 & 0 \\ 0 & t & -1 & 0 & 1-t \\ -1 & 0 & 1-t & 0 & t \end{bmatrix} \longmapsto \begin{bmatrix} 1-t & 0 & 0 & t \\ t & -1 & 0 & 1-t \\ 0 & 1-t & t & -1 \\ 0 & t & -1 & 0 \end{bmatrix}$$

$$2t^3 - 3t^2 + 2t \longmapsto 2t^2 - 3t + 2 \tag{2.2}$$

FIGURE 2.9: A) Each X_i is drawn separately in order to define a sign with the right-hand rule, and compute the elements of the matrix. B) Deleting one row and one column we obtain the Alexander matrix and consequently the Alexander polynomial given by the determinant of the matrix. Image adapted from ref. [55]

2.4 Knots in open chain and length of the knotted portion

So far, given the definition of knot, we only considered closed knotted curve. Indeed, for a knot in a open curve, no matter how complex the knot can be, it is always possible to define a set of moves, not involving strand crossings, that transform it into a linear, hence unknotted, segment. However, even if topologically trivial, knots tied on an open chain can have mechanical effects not entirely negligible, for instance, knot in a rope can lower the mechanical resistance. Hence, these kind of knots, even thought they lack a strict mathematical definition, are relevant and are usually referred to as physical knots.

Intuitively, to characterize the degree of entanglement of an open chain it is possible to join the two ends with an auxiliary arc (in this way the topology is fixed and the physical knot trapped in the ring) and measure the Alexander polynomial. The closure procedure is a key element and it must be chosen appropriately and robustly

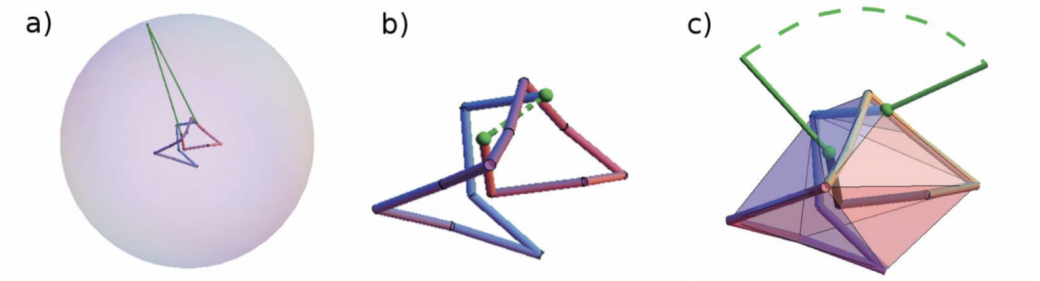


FIGURE 2.10: examples of the closure schemes. Image from [58]. (a) Stochastic closure at infinity, (b) direct bridging, (c) minimally-interfering closure.

to ensure that it does not itself introduce spurious entanglement to the chain, which could offset the complexity of the detected physical knot.

Here we recall two main ways of closure: the stochastic closure to infinity and the minimally-interfering closure [58].

In the first procedure, the stochastic closure to infinity, the two ends are joined together, with two segments, on a random point taken on a sphere. The sphere is centered around the chains center of mass and has radius much larger than its gyration radius, see Fig. 2.10a. Different random points will generally yield different types of knots, so for this reason one can take the most common topology out of the many closures as the representative one. This procedure provides a more complete statistics on the possible topologies of the chain, but needs the calculation of many knot invariants.

The second procedure, called minimally-interfering closure, compares the arc lengths introduced by two different closure schemes. The first scheme joins the two ends with a straight line, and is hence called direct bridging, see Fig. 2.10b. The second scheme prolongs the ends to their respective closest point at the convex hull and then joins them together with an arc at infinity, see Fig. 2.10c.

The scheme with the shortest total arc length (excluding the closure at infinity for the second scheme) is selected, in order to introduce the least amount of spurious crossings. We use the latter method for computational effectiveness.

Length of the knotted portion Alongside with the identification of the knot type, we are interested also in locating the knotted portion of a given chain. We define the length of knot as the shortest portion of a curve that, after a suitable closure has the same topology of the entire ring.

The tool we use is given by ref. [58], and it can be done via a *bottom up* or *top-down* search of the knotted region.

In the first case, the search starts with dividing the curve in short portions. For each of them, it is checked that: (i) the topological state is the same as the whole chain, and (ii) that the complementary closing arc is unknotted. If the criteria are not satisfied, the length of the considered portion is increased, until the aforementioned criteria are satisfied.

The top-down approach, instead, considers all the possible arcs of given length, starting from the maximum possible one (the whole chain), and identifies the arcs that have the same topology as the entire chain. Only the arcs satisfying this criterion will be then subdivided in smaller arcs of length decreased of a unit, while the others will be discarded. The new candidate arcs will be again checked with the same criterion, and the arcs satisfying it will be selected and further subdivided.

The procedure ends if, after a subdivision, no one of the arcs has the same topology of the entire chain. Then, the arc that survived the previous iteration will be considered the shortest continuously-knotted portion of the ring.

We use, in our case, the bottom-up approach. The algorithm for knot location used in this thesis has been implemented in the Kymoknot software package [59].

2.5 Links

In this thesis, we focus on knots, that occur in DNA rings, and their properties. However there are some concepts, that emerges in the context of links that can be useful to characterized properties of dsDNA rings, where two strands are wound around each other. To this aim, we need some tools that are beyond the one described for knots.

We define link a set of knotted loops all tangled up together [60]. Some simple examples are shown in Fig. 2.11. Two links are considered to be the same if we can

deform one link into the other link without ever having any one of the loops intersect itself or any of the other loops in the process tied together. For our purpose, we use links of two components, each being on of the strands in dsDNA ring. A way to quantify how linked up two components are is to count how many time one curve wraps around the other. More precisely, is to compute the so called Gaussian linking number.

Linking number The linking number, L_k, is an invariant that describes the linking of two closed curves in three-dimensional space.
L_k is defined as the sum of the sign over all the crossing between two oriented curves average over all possible projections. The sign of each crossing is defined accordingly with the right-hand rule of Fig. 2.4.
The outcome will always be an integer number, either positive or negative depending on the orientations on the two components. Notice that the only considered crossings are the ones between the two curves and not the ones that one component does with itself. However, the self-crossing of one curves can influence the resulting linking number: for example, consider the Whitehead link shown in Fig. 2.11c, it has four crossings. Because of the crossing present in one component, each underpass is followed by an overlapp and the resulting linking number is 0. It is not isotopic to the unlink, but it is *link homotopic* to the unlink, which means that is allowed to move, through regular homotopy, each component of the extended link, knotting or unknotting itself, but is not allowed to move through other component.
The linking number was firstly introduced by Gauss in the form of the linking integral [61]: given two non-intersecting differentiable curves γ_1, γ_2:

$$L_k = \frac{1}{4\pi} \oint_{\gamma_1} \oint_{\gamma_2} d\vec{r_1} \times d\vec{r_2} \cdot \frac{\vec{r_1} - \vec{r_2}}{|\vec{r_1} - \vec{r_2}|^3} \tag{2.3}$$

The linking number remain constant for ambient isotopies of the curves, but the spatial displacement of the two components does have consequences in the geometrical properties of the link. In particular, the coiling and twisting of the link component: writhe Wr, which we have defined before, and twist Tw. They are related to the

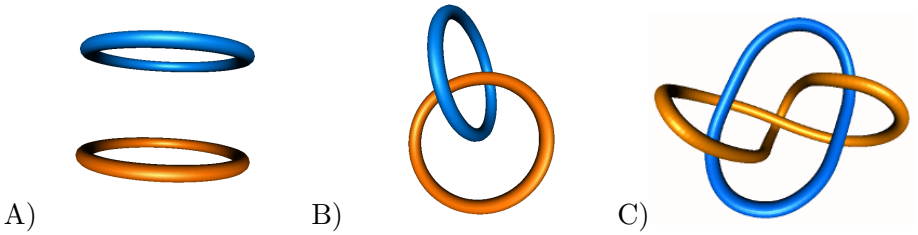


FIGURE 2.11: Simple links. A) the unlink. B) the Hopf link. C) Whitehead link.

linking number in the following way:

$$L_K = Tw + Wr \tag{2.4}$$

This equation implies that any twist added to the curve, will be exactly balance in changes of writhe.

Knots and Linking number Let consider now the centerline of the two curves of a given link, and consider the case in which the centerline is topologically unknotted. When placing the centerline on a plane, the writhe is equal to zero and the twist will be equal to the linking number according to eq. 2.4. However, if the curve is knotted it cannot possibly attain perfectly planar configurations due to its essential crossings and, in fact, its writhe will typically be different from zero. This intrinsic writhe must be properly discounted when using the linking number to compute the amount of excess or deficit twist in a knotted DNA ring.

Knotting probability of semiflexible rings: numerical results

Non monotonic dependence on bending rigidity

3.1 Introduction

Since the '70s [12, 14, 62], the spontaneous knotting of biological polymers has been studied both from mathematical and physical point of view. In particular, the description of the types of the entanglements, their abindance and formation mechanism are questions of direct biological relevance since several studies have found knotted DNA in all organism, from eukaryotes, to bacteria, to viruses [11, 15].

Knots can be detrimental for the cell [63]: when occur in DNA or RNA, they affect the correct *in vivo* processing of the biopolymer, such as replication or transcription, as well as its translocation through biological pores. Indeed, knots introduce strong constraints which affect the properties of polymers, such as their metric properties, their mechanical response or their dynamics. Therefore, with the aim of understanding these effects, several knot-related equilibrium properties have been systematically studied over the past years using theoretical and computational means.

For instance, one physical property that affects knotting probability, P_k, is the length of the polymer. In this case, we already recalled in the introduction, the Frisch, Wasserman, and Delbruck (FWD) conjecture that can be applied to self-avoiding lattice walks, for which can be proved that *almost all sufficiently long self-avoiding walks on the 3D simple cubic lattice contain a knot* [64]. In further studies [65], numerical simulations have shown that this result holds also in random polymers, in particular,

the asymptotic probability that the chain is unknotted, $P_{unknotted} = 1 - P_k$, decreases exponentially with the number of polymers segments, see Fig. 3.1a.

Other elements that strongly affect P_k are the confinements of the polymer [21], the quality of the solvent [66, 67], the applied mechanical load at the chain ends [68].

Nevertheless, some features are still underexplored, for instance the role played by polymer bending rigidity. To our knowledge no systematic survey of the interplay of chain length, bending rigidity and knotting has been reported so far. The need of such systematic study emerges also through the difficulty of rationalising results available in literature under a simple, unifying framework.

For example, for the semi-flexible chain regime we recall the study on phage P4 DNA by Rybenkov *et al.* [43]. They investigated, experimentally and numerically, the abundance of knots produced by random cyclization. In particular, considering P4 DNA molecules of 10Kbp, cyclized in 1 M of NaCL, the knotting probability is $\sim 5.5\%$. They also modelled the P4 DNA with ring of cylinders, and reproduced the experimental data via computer simulations. In particular, they fixed the thickness equal to $2.5nm$, the contour length equal to $3.4\mu m$, and included, via the bending rigidity, the persistence length, which is 50nm for DNA.

Consider now the case of fully-flexible chains. Indeed, in the case of zero rigidity, one can explore the knotting probability of polymers modelled as fully-flexible ring of cylinders, as it as been done, for instance, in the study of ref. [69]. In the latter the unknotting probability was found to decrease, as expected, with the number of cylinders, and to increase with the cylinder's radius. The combined effect yields knotting probability profile of Fig. 3.1b.

Consider the contour length and thickness of the P4 DNA, since the persistence length of the DNA is 50nm the Kuhn length is equal to 100nm (see chapter 1). The genome of P4 thus consists of 34 Kuhn segment, and can be mapped to rings of 34 cylinders of unit length and radius equal to 0.01. The estimate probability of unknotting is $\geq 97\%$ as shown in Fig. 3.1b.

The resulting knotting probability for *fully-flexible* rings is $\sim 3\%$, which is unexpectedly *lower* than aforementioned knotting probability of equivalent (P4 DNA) rings which was $\sim 5.5\%$. Therefore, on one hand the aforementioned comparison leads to the overall suggestion that the knotting probability increases with rigidity. On the

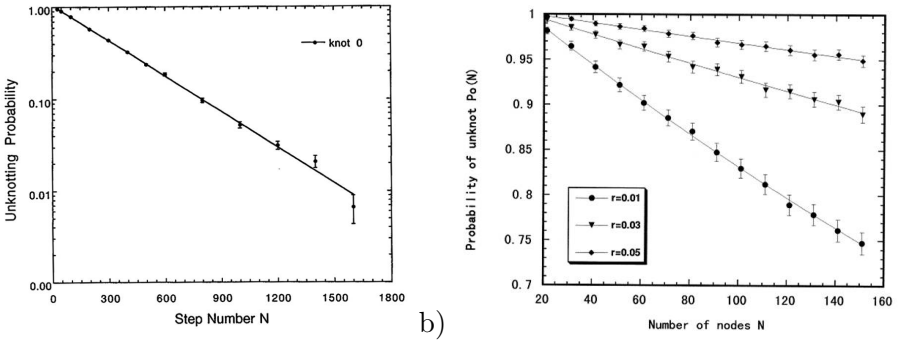


FIGURE 3.1: (a) Unknotting probability of random polygons versus the step number N, in semi-log scale. The unknotting probability decays exponentially. Image from [65]. (b) Unknotting probability versus the step number N for self-avoiding polygons with different values of radius r, in semi-log scale. The unknotting probability grows with the thickness of polygon.Image from [69].

other hand is known that, when rigidity goes to infinity, the probability of knotting goes to zero.

This apparently contradictory result has not been addressed, and suggests that the behaviour of knotting probability for semiflexible rings is more complex than expected. The closest related investigations are the study of ref. [70] on lattice rings with corner penalties, and the one of ref. [71] off-lattice open phantom and self-avoiding chains of 150 segments. Given the absence of systematic results, we decided to undertake the elucidation of the problem.

In this chapter, we describe how the knotting probability of a semi-flexible polymer depends on its bending rigidity. In particular, we systematically address this problem by applying Monte Carlo sampling and thermodynamic reweighting techniques to semiflexible rings of up to 1000 beads, for different bending rigidities.

The content of this chapter is mainly based on the work published in ref. [72].

3.2 Model and Methods

3.2.1 Model

We use a minimalistic, coarse-grained polymer model that incorporates only salient physical properties of a polymer, such as chain connectivity, excluded volume effects and bending rigidity. The model consists in ring of N identical and spherical touching

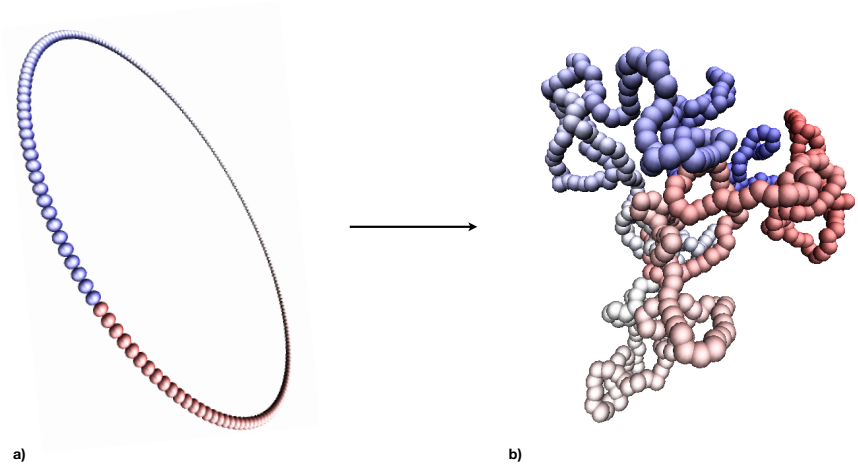


FIGURE 3.2: (a) Starting configuration; circular ring. (b) Evolved configuration of $N = 500$ beads.

beads, as described in chapter 1.2.1. We considered ring in equilibrium at temperature T so that the nominal persistence length is $l_p = b\kappa_b/K_BT$. Notice that in the following discussion K_BT and the bond length, b, are taken as unit of energy and length, respectively, and all the shown quantities are expressed in these units or their suitable combinations.

Total bending and total curvature The model is characterised by the bending energy $U = \kappa_b B$, where κ_b is the bending rigidity and B is the total bending, as defined in chapter 1.2.1. This physical quantity is linked to the total curvature. Among the studies that investigate the connection of knotting properties and chain curvature, we wish to recall to works of Fary (1949) [73] and Milnor (1950) [74]; they have independently proved that *the total curvature of a knot in three-dimensional space must be* $\geq 4\pi$, meaning that the curve must complete at least two loops in order to be knotted. This result sets the lower bound for the total curvature, and hence, total bending of a knot.

3.2.2 Monte Carlo

The sampling of the conformational space is done with the Monte Carlo method described in chapter 1.3. We explore different contour lengths, considering rings of $N = 100, 200, 500, 1000$ beads and vary the bending rigidity in the $0 < \kappa_b < 40$ range

with typical increments equal to 2: from fully-flexible to rigid polymers.

The rings, initially prepared in an exactly circular arrangement of touching beads, are evolved using a Metropolis Monte Carlo scheme based on unrestricted crankshaft moves, see Fig. 3.2. In particular, the global crankshaft is picked with probability $\sim 90\%$.

For each combination of N and κ_b we collect no less than 10^5 uncorrelated configurations. To define the number of MC steps needed to have two independent configurations, we use the gyration radius as observable to establish the autocorrelation time τ_c, defined in Chapter 1.3, and we save one configuration each $10N$ and check *a posteriori* that the configurations are uncorrelated.

Of the various chain lenghts, we particularly focus on chains with $N = 500$ beads. This choice is motivated by the necessity of keeping a reasonable computational time and yet collect enough statistics of knotted rings. We sample at least 10^6 configurations varying in the same range of bending rigidity with typical increments equal to 0.5.

The topology of each sampled configuration is determined by computing the Alexander polynomial. For rings of non-trivial topology we identified and measured the length of the knotted portion, l_k as described in Chapter 2.

Free-Energy and Entropy Calculation To have better understanding of the mechanism that underpins the system behaviour, we need to study independently how the energy and the entropy of the system behaves in function of κ_b. This is done by analysing the collected data with a thermodynamic reweighting technique, as detailed below.

The κ_b-dependent free energies, F, of knotted and unknotted rings and their entropic, $T \cdot S$, and energetic, U, contributions can be computed via the following

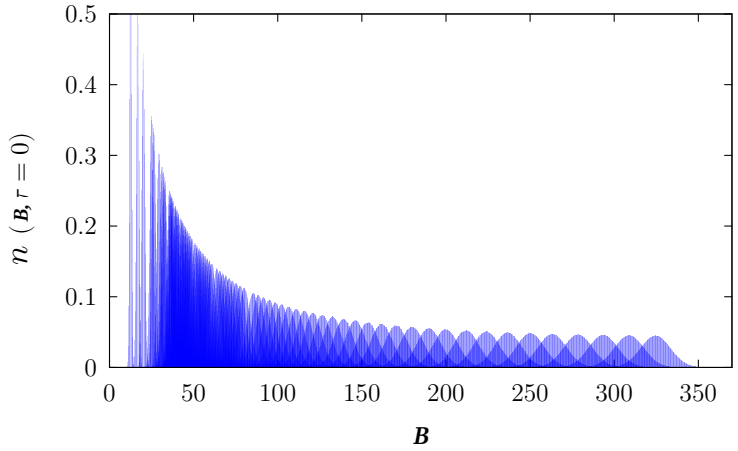


FIGURE 3.3: Number of configuration at given value of total bending, topology and temperature, for different value of bending rigidity κ_b, with decreases from left to right.

equations:

$$F(\kappa_b,\tau) = -\log(\sum_i W(B_i,\tau)\, e^{-\kappa_b B_i}) \;, \tag{3.1}$$

$$U(\kappa_b,\tau) = \frac{\sum_i \kappa_b B_i\; W(B_i,\tau)\, e^{-\kappa_b B_i}}{\sum_i W(B_i,\tau)\, e^{-\kappa_b B_i}} \;, \tag{3.2}$$

$$T\cdot S(\kappa_b,\tau) = U(\kappa_b,\tau) - F(\kappa_b,\tau) \;, \tag{3.3}$$

where τ labels the topology, i labels the discretised values of B and, $W(B_i,\tau)$ is the density of states with total bending B_i and topology τ.

3.2.3 Thermodynamic reweighting techniques

The method presented here is the weighted histogram technique of Ferrenberg and Swendsen [75], which is customary used to compute the density of states $W(E)$ with given energy E. This quantity, in turn can be used to compute the average number of states $n(E)$ as:

$$n(E) = \frac{NW(E)e^{-\beta E}}{\sum_E W(E)e^{-\beta E}} = \frac{NW(E)e^{-\beta E}}{Z_T(E)} \tag{3.4}$$

where $\beta = 1/K_BT$ and K_B is the Boltzmann constant, and, the total number of states is:

$$N = \sum_E n(E) \tag{3.5}$$

by inverting the 3.4, the density of states as function of E is:

$$W(E) = \frac{n(E)}{N} e^{\beta E} \tag{3.6}$$

To combine the density of state at different temperatures, the eq. 3.4, can be rewritten as:

$$n_j(E) = N_j W_j(E) e^{-\beta_j (E - F(T_j))} = \frac{N_j W_j(E) e^{-\beta_j E}}{Z_{T_j}(E)} \tag{3.7}$$

and the density of states can be obtained by combination of the data $W_j(E)$ provided by the simulations at temperature T_j:

$$W(E) = \Sigma_j p_j(E) W_j(E) = \Sigma_j p_j(E) \frac{n_j(E)}{N_j} e^{\beta_j [E - F_j(T_j)]} \tag{3.8}$$

where $\Sigma_j p_j(E) = 1$. For each value of E is possible to assign a weight p_j that depends on the accuracy of the information of the histogram W_j. The criteria is given by the least squares method, and leads to:

$$p_i = \frac{\frac{1}{{\varepsilon_r}^2(W_j)}}{\Sigma_j {\varepsilon_r}^2(W_j)} \tag{3.9}$$

Variance To estimate the variance, $\varepsilon_r^2(W_j)$, we consider the number of configurations at given E, $n_j(E)$ as events of a Poission distribution and therefore $\varepsilon_r^2(n_j(E)) = n_j(E)$, and it's possible to compute:

$$\varepsilon_r^2(W_j) = \frac{\varepsilon_r^2(n_j(E))}{N_j^2} e^{2\beta_j [E - F(T_j)]} = \tag{3.10}$$

$$= n_j(E) \left[\frac{e^{2\beta_j [E - F(T_j)]}}{N_j} \right]^2 \tag{3.11}$$

Substituting the last result in 3.9 and combining 3.8 and 3.7, we obtain that the approximation of the density of states is:

$$W(E) = \frac{\Sigma_i N_i e^{-\beta_i(E-F(T_i))}}{\Sigma_j \frac{N_j}{n_j(E)} e^{-\beta_j(E-F(T_j))} N_j e^{-\beta_j(E-F(T_j))}} = \tag{3.12}$$

$$= \frac{\Sigma_i \frac{n_i(E)}{W(E)}}{\Sigma_j \frac{1}{W(E)} N_j e^{-\beta_j(E-F(T_j))}} = \tag{3.13}$$

$$= \frac{\Sigma_i n_i(E)}{\Sigma_j N_j e^{-\beta_j(E-F(T_j))}} \tag{3.14}$$

Once the density of states is calculated, up to a multiplicative constant, it is possible to obtain the free energy which is established up to an additive constant. For instance, our purpose is to compute the density of states to obtain the free-energy profile, in particular:

$$e^{-F(T_j)/Tj} = \sum_E W(E) e^{-\beta_j E} \tag{3.15}$$

Notice that, this result, is easily improved by adding more data in the regions of E in which the overlap is lacking.

Multidimensional case The method can be extended to multidimensional cases: consider a system described by its energy and topology: τ. The density of states is $W(E,\tau)$, and $\sum_\tau W(E,\tau) = W(E)$. Since $\Sigma_E W(E) = 1$, also $W(E,\tau)$ is normalized:

$$\sum_\tau W(E,\tau) = \left(\sum_{E'} W(E') e^{-\beta_i E'} \right) \frac{\sum_\tau n(E,\tau,T_i)}{N_i} e^{\beta_i E} \tag{3.16}$$

$$\frac{\sum_\tau n(E,\tau,T_i)}{N_i} = \frac{n(E,T_i)}{N_i} = \frac{W(E) e^{-\beta_i E}}{\sum_{E'} W(E') e^{-\beta_i E'}} \tag{3.17}$$

$$\sum_\tau W(E,\tau) = W(E) \tag{3.18}$$

Collecting data from the n_t simulations at different temperatures, the density of states is:

$$W(E,\mu) = \frac{1}{n_t} \sum_i \left[\left(\sum_{E'} W(E') e^{-\beta_i E'} \right) \frac{n(E,\tau,T_i)}{N_i} e^{\beta_i E} \right] \tag{3.19}$$

where $1 \leq i \leq n_t$.

Errors The errors on these quantities are estimated from their semi-dispersion computed by doubling and halving the bin size, as well as by using only the first and second half of the sampled conformations at each considered value of κ_b.

Application We applied this framework in our study to compute the density of states of rings of $N = 500$ beads. Specifically, we computed the density of states, $W(B, \tau)$ as a function of two properties of the rings: total bending, B, and topology, τ. The total bending B, was discretised in steps of width 0.5, while the inherently discrete topological classes were encoded in a binary form, with $\tau = 0$ or 1 for unknotted and knotted rings, respectively.

For each value of κ_b we computed the two dimensional histogram, $n(B, \tau)$, see fig. 3.3. Clearly $n(B, \tau)$ is proportional to the sought $W(B, \tau)$ and to the Boltzmann weight $\exp(-\kappa_b B)$. This relationship allows for recovering $W(B, \tau)$, up to a multiplicative constant, from $n(B, \tau)$ by inverting the Boltzmann weight. Due to inevitable sampling limitations, the recovered W will be reliable in the neighbourhood of the most probable values of B. By combining, via eq. 3.19, the reweighted histograms collected at different κ_b's, one can obtain the density of states, $W(B, \tau)$, throughout the relevant parameter range.

3.3 Results

3.3.1 Knotting probability

The probabilities of non-trivial knots, collected for $N = 100, 200, 500$ and 1000 show tricking behaviour as function of bending rigidity, see Fig. 3.4a. We see that as κ_b increases the knotting probability grows, reaching a peak in the range $4 < \kappa_b < 10$, and then it decreases to zero. It is possible to notice that the enhancement of Pk from the fully flexible case, $\kappa_b = 0$, to the peak value is very large: more than a factor of 40 for the longest considered chains, $N = 500$ and 1000.
This result is unexpected and counter-intuitive; in particular it leads to the fact that is more probable to have a knot in a polymer with a finite value of rigidity than in a fully-flexible one.

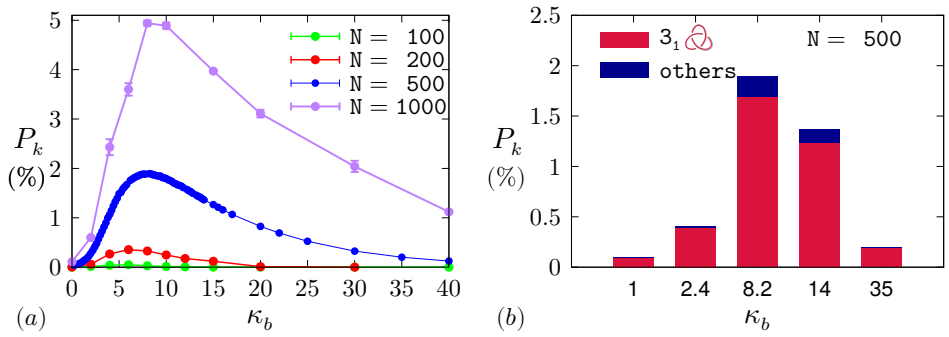


FIGURE 3.4: (a) Knotting probability as a function of κ_b for rings of different lengths. Estimated statistical errors (shown) are typically smaller the symbol size. (b) Pile up representation of the abundance of knots with 3_1 or other non-trivial topologies in rings of $N = 500$ beads for different values of κ_b.

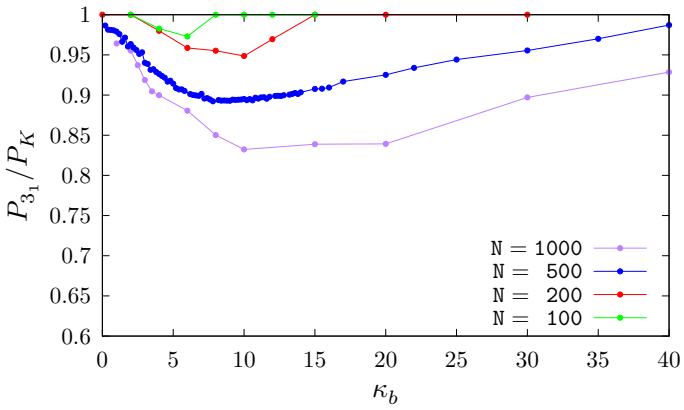


FIGURE 3.5: Percentage of 3_1 knot for all the rings of N beads for different values of κ_b.

To confirm that the non-monotonic behaviour is a genuine effect which does not arise from a complex interplay between the types of knots, we compute the abundance of knots with 3_1 and other non-trivial topologies in rings of $N = 500$ beads for different values of κ_b, in Fig 3.4b . Indeed, we find that trefoil knots are by far the dominant non-trivial topology. In fact, for any combination of κ_b and N they account for no less than 80% of non-trivial knots as shown in Fig. 3.5

3.3.2 Metric Description

To have a comprehensive description of the previous result we study the behaviour of metric properties; in particular, averages of the gyration radius, $\langle R_g \rangle$ (defined as eq. 1.14 in Chapter 1.1), and the total bending, $\langle B \rangle$ (see Chapter 3.2.1), both computed on the entire set of sampled rings.

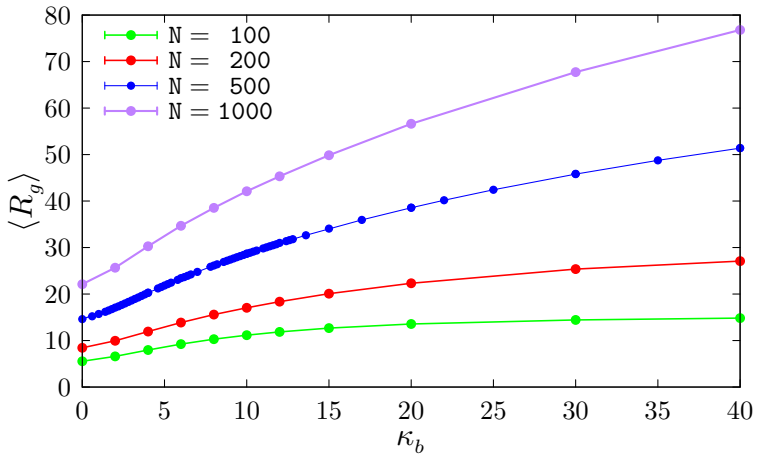


FIGURE 3.6: The average gyration radius of rings of different lengths is shown as a function of κ_b. Estimated statistical errors are typically smaller than the symbol size.

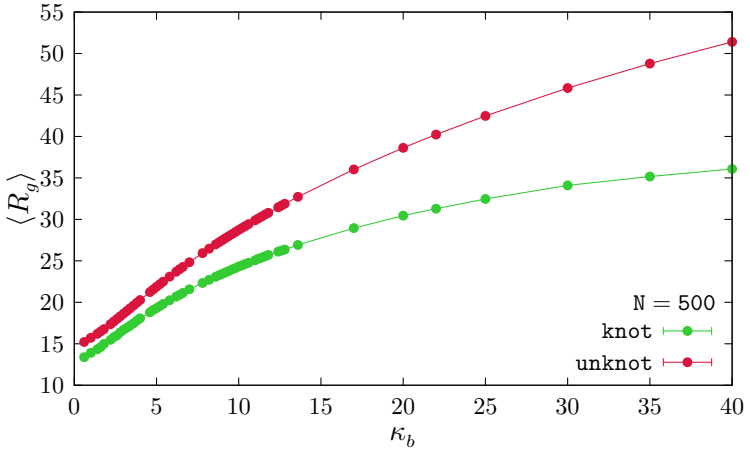


FIGURE 3.7: Average gyration radius of knotted and unknotted rings as a function of κ_b.

Gyration Radius In Fig 3.6 it is shown the behaviour of $\langle R_g \rangle$ as function of κ_b: the size of the polymers monotonically increases with bending rigidity. However, even considering large value of rigidity, there still enough conformational freedom to sample some knotted configurations.

We next proceed looking at the difference between the knotted rings and the unknotted one.

As mentioned before, we limit the analysis to rings of $N = 500$ beads. Firstly, we established that the average $\langle R_g \rangle$ grows differently with κ_b for the two topology, see Fig. 3.7. Unknotted rings are larger than knotted ones, especially for large values of rigidity. This is because a portion of the contour of knotted rings is used up to maintain their non-trivial topology, more precisely to form crossings in the ring, a portion of it has to fold to form the crossing. This reduce the possible extension of

ring, and therefore the average size at fixed ring-length is smaller for knotted ring. In addition, increasing the rigidity, knots will require a even larger portion to be bent, this is clear looking that the distance between the two curve of $\langle R_g \rangle$ increases with κ_b, see Fig. 3.7.

Knot's length The second observable of interest in the characterization of the system is the length of the knotted region, l_k. Surprisingly, the behaviour as function of bending rigidity is non-monotonic. For $\kappa_b \geq 2.4$ the behaviour follow the intuition: a progressive chain stiffening implies that the knots grow in contour length, up to one third or more of the ring contour. However, for small values of bending rigidity, $\kappa_b < 2.4$, the knotted region expands as the ring becomes more flexible.

This non-monotonicity is also found in the ratio of $\langle R_g \rangle$ for knotted and unknotted rings, see Fig. 3.8a. The ratio is unimodal with κ_b and with the maximum at $\kappa_b < 2.4$ which is less than one third of the κ_b-value of the P_k maximum (marked with a blue line in the figure). Comparing the two results shown in Fig 3.8a, it is clear that there exists an anti-correlation between $\langle l_k \rangle$ and $\langle R_g \rangle$.

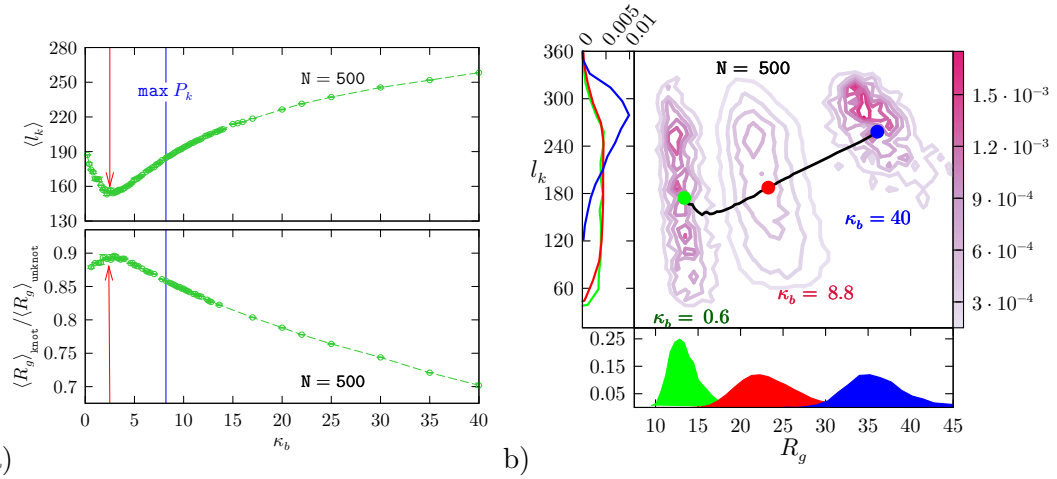


FIGURE 3.8: (a) UP: Average gyration radius of knotted and unknotted rings as a function of κ_b. BOTTOM: Ratio of the average gyration radii of knotted and unknotted rings and (c) average knot length as function of the bending rigidity. Data refer to rings of $N = 500$ beads. The vertical red and blue lines mark respectively the position of the l_k minimum and P_k maximum for $N = 500$ (see Figure 3.4). Statistical errors (shown) are typically smaller than the symbol size.
b) Central panel: Contour maps of the probability distribution $P(R_g, l_k)$ of knotted rings of $N = 500$ beads for three selected values of κ_b: 0.6, 8.8 and 40.0. The coloured filled circles mark the average values of $\langle R_g \rangle$ and $\langle l_k \rangle$ at the selected values of κ_b. The black curve is the locus of points corresponding to the averages of $\langle R_g \rangle$ and $\langle l_k \rangle$ at all considered values of κ_b. The side-panels show the marginal distributions of R_g and l_k.

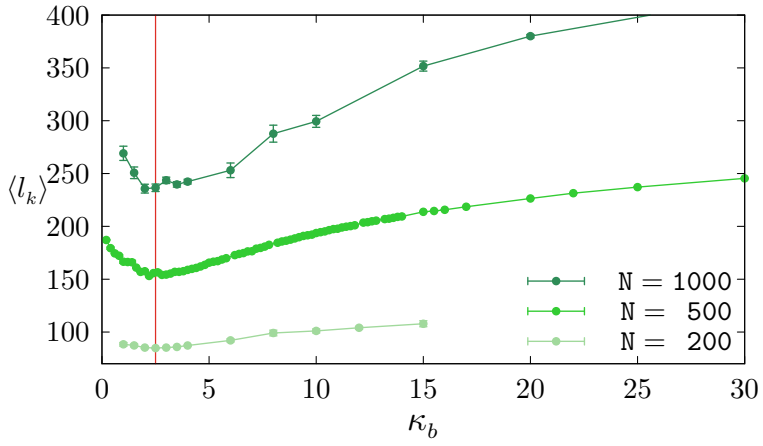


FIGURE 3.9: Average knot length as function of the bending rigidity for rings of $N = 200$, 500 and 1,000 beads. The vertical red line marks the position of the l_k minimum for $N = 500$. Statistical errors (shown) are typically smaller than the symbol size.

Anticorrelation between gyration radius and knot length The same figure, in panel b, shows the interplay between l_k and R_g and bending rigidity. The central panel shows the $P(R_g, l_k)$ for three value of κ_b, while the side panels, provide the corresponding marginalized probability distributions of l_k and R_g.

At fixed κ_b, the contour lines of $P(R_g, l_k)$ reflect the anticorrelation of knot length and ring size.

This effect is arguably the cause for the initial decrease of l_k with κ_b observed at small bending rigidities in Fig. 3.8a. In fact, an increase of κ_b always yields a larger gyration radius, see Fig. 3.6, which in turn translates in a decrease of l_k. Note that as κ_b increases, the degree of anticorrelation between l_k and R_g is reduced and the marginalized distribution of l_k narrows and shifts towards larger values, see the blue line on the side panel of Fig. 3.8b.

This property was previously established for fully-flexible chains without excluded volume interactions, where it was rationalized as follows: reducing the knot length is equivalent to increasing the length of the complementary unknotted arc and, because this enjoys a larger conformational freedom, it ultimately reflects in a larger gyration radius of the ring [76].

Finally, we also profile l_k for other ring lengths, namely $N = 200, 500$ and 1000 beads to verify that the non-monotonicity of l_k exists for these chain length too, see

Fig. 3.9. Indeed, the minimum of $< l_k >$ exists and is not appreciably displaced by varying N unlike the location of the P_k peak, which visibly depends on N.

3.3.3 Knotting and bending energy

The intriguing aspect of the non-monotonic behaviour lies in the expectation that a progressive decrease of bending rigidity translates into a decrease of bending energy. Conversely bending a rigid ring, in order to form a knot, would be energetically disfavoured. Therefore, it is interesting to understand this in more details; in particular we consider how the entropic and energetic interplay is affected by the increase of bending rigidity.

Total Bending The bending energy is proportional to κ_b and to the total bending of the ring as defined before (see chapter 3.2.1). The average total bending $\langle B \rangle$, which is shown in Fig. 3.10a decreases monotonically with κ_b. For small value of κ_b, $\langle B \rangle$ reaches very large values and decreases dramatically with the stiffening of the ring. Much more informative is the difference between the average total bending of knotted and unknotted rings: see Fig. 3.10, panel b. For large value of κ_b, knotted rings confirm our intuition: the ΔB shows a ~ 0.1 surplus of total bending of knotted compared to unknotted rings. The asymptotic value of ΔB is also explained by the mathematical result described in paragraph 3.2.1: the minimum value of total curvature that a curve must have to be knotted is 4π, whereas for unknotted curve is only 2π. For large values of κ_b, these bounds are viable approximations for the typical total bending angles of the sampled knots and unknots and the corresponding, for $N = 500$, $\Delta B = \langle B \rangle_{rigid\_knot} - \langle B \rangle_{rigid\_unknots} = N(cos(\frac{2\pi}{N}) - cos(\frac{4\pi}{N})) \sim 0.12$, and is well in line with the observed value.

Decreasing the bending rigidity, $\kappa_b < 8$, the ΔB becomes negative. This unexpected result means that knotted rings have, on average, a smaller total bending than unknotted ones. The largest bending deficit in knotted rings is found in the fully-flexible limit, where it is equal to $\Delta B = -1.3$. Note that P_k attains its maximum when ΔB is approximately equal to 0, and consequently, the lower total curvature of knotted rings must clearly translate into a more favourable energy compared to unknotted ones when $\kappa_b \sim 8$.

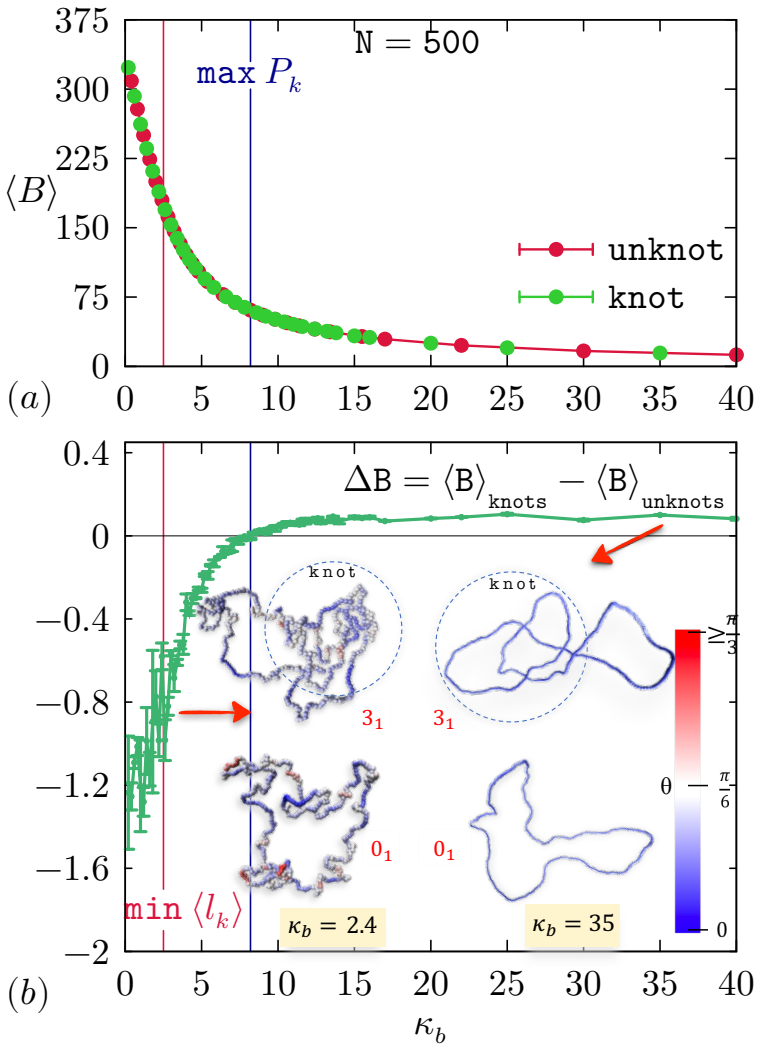


FIGURE 3.10: (a) Average total bending of knotted (green) and unknotted rings (red) of $N = 500$ beads as function of κ_b; (b) Difference of total bending of knots and unknots, ΔB as a function of κ_b. In both panels, the blue (right) and red (left) vertical lines mark the location of the maximum of P_k and the minimum of l_k, respectively. Typical knotted and unknotted configurations for $\kappa_b = 2.4$ and 35 are shown (not to scale) in the inset. The configurations are colored according to their local bending angle, see color bar.

The counterintuitive fact that knots are less bent than unknots is well illustrated by the typical knotted and unknotted conformations at $\kappa_b = 2.4$, as shown in the insert of 3.10. From the color code, which reflects the degree of local bending, one notices two features: the unknotted configuration has more points of high bending (red regions) than the knotted one. Moreover, the knotted region is less locally bent than the rest of the ring. This effect is not observed in the knotted configuration at $\kappa_b = 35$.

A more detailed view of the distribution of bending along the ring is made by looking at the probability of a bending angle θ. More precisely, studying the difference

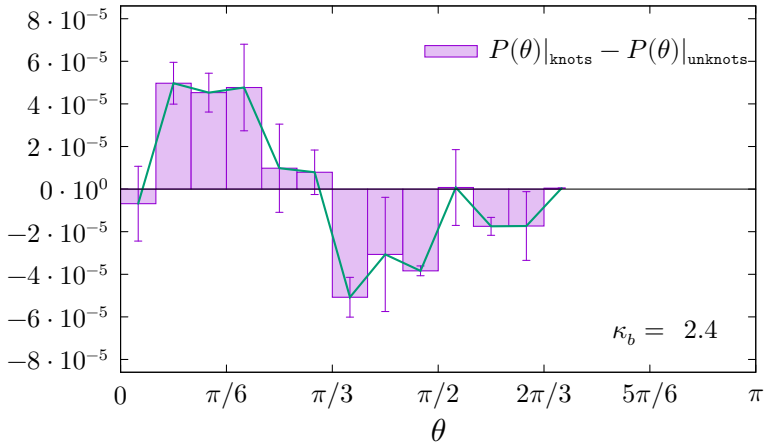


FIGURE 3.11: Panel shows $P(\theta)|_{knots} - P(\theta)|_{unknots}$, that is the difference of the normalized probabilities to observe a certain local bending angle, θ, in knotted and unknotted rings at $\kappa_b = 2.4$.

of the normalized probabilities, $P(\theta)$, between knotted and unknotted rings. This difference, at $\kappa_b = 2.4$, is shown Fig. 3.11. It is clear that small angles are more likely in knotted rings, while the opposite is true for large angles. The balance of the probability difference corresponds to a bending deficit $\Delta B \sim -1.3$ at $\kappa_b = 2.4$.

These observations lead to the conclusion that topological (chain uncrossability) and excluded volume constraints are so effective on flexible knotted rings that their average bending is smaller than for unknotted rings. The effect is analogous to the straightening of knotted strands observed for tightened knots in ref. [77]. For that system, the effect arose from the mechanical tension applied to the knotted chain, which caused a tight contact of the entire knotted region. It is remarkable that the straightening is found in our system too, where no external tension is applied, and the knot size and tightness are controlled only by thermodynamic forces.

Bending Energy Once analysed the total bending, we study the average energy difference, $\Delta U = \langle U \rangle_{knots} - \langle U \rangle_{unknots}$, see Fig. 3.12, red line. The behaviour is non monotonic, but more important is that the difference is zero at the point in which the knotting probability has its maximum.

Indeed, the profile of the average energy difference, ΔU between knots and unknots changes sign at $\kappa_b \sim 8$. Note that ΔU is the product of κ_b and ΔB, which is itself a function of κ_b. As a consequence, ΔU, unlike ΔB, is not monotonic and has a global minimum at $\kappa_b \sim 2.4$. This is approximately the same bending rigidity for which the

relative knotted versus unknotted rings size is maximum, and the average knot length is minimum at all lengths shown in Fig. 3.9, $N = 200$, 500, and $1,000$. It is therefore possible that such weak N-dependence of the $\langle l_k \rangle$ minimum originates from the fact that ΔB and ΔU at low κ_b are controlled by the number of essential crossings, which is effectively equal to three at all the considered chain lengths because trefoil knots are the dominant non-trivial topology, this is still a speculation and an interesting point warranting further investigations.

Free Energy and Entropy If the energetic term, ΔU, in Fig. 3.12 gave the main contribution to the free energy difference, then knotted rings would dominate the equilibrium ensemble for $\kappa_b \lesssim 8$ and would maximally outnumber unknotted rings at $\kappa_b \sim 2.4$.

The actual knotting probability profile of Fig. 3.4 shows instead that unknots are the dominant topology (at the considered N's) and that knots are maximally abundant at $\kappa_b \lesssim 8$, when knots and unknots have about equal bending energies. This implies that the favourable energy bias towards knots must be counteracted by entropic effects.

To address this point quantitatively we used the multiple-histogram technique of section 3.2.3 to compute the relative density of states of knotted and unknotted rings and hence their entropy difference. The results are given in Fig. 3.12 and clarify that as κ_b is lowered, going from the rigid to the flexible limits, the entropy of knotted rings is systematically reduced compared to unknots. The variation of $T\Delta S$ are concentrated in the $2.4 \leq \kappa_b \leq 8$ range. This interval spans approximately from the κ_b associated to the minimum knot length (see Fig. 3.8) to the one associated to the knotting probability maximum (see Fig. 3.4). A relevant point is that at low κ_b the entropic loss eventually outweighs the bending energy gain of introducing knots.

The balance of the entropic and energy terms yields the free energy difference profile, ΔF, of Fig. 3.12, which has a minimum for $\kappa_b \sim 8$. The position of the minimum agrees well with the location of the P_k peak for $N = 500$ in Fig. 3.4. The minimum free energy difference itself, $\Delta F^{min} \sim 4.0$ is also consistent with the direct measurements of the P_k peak, $P_k^{max}(N = 500) \sim 1.9 \cdot 10^{-2}$.

From these results we can put forward the following explanation for the counter-intuitive knotting enhancement at intermediate values of κ_b. Going from rigid to

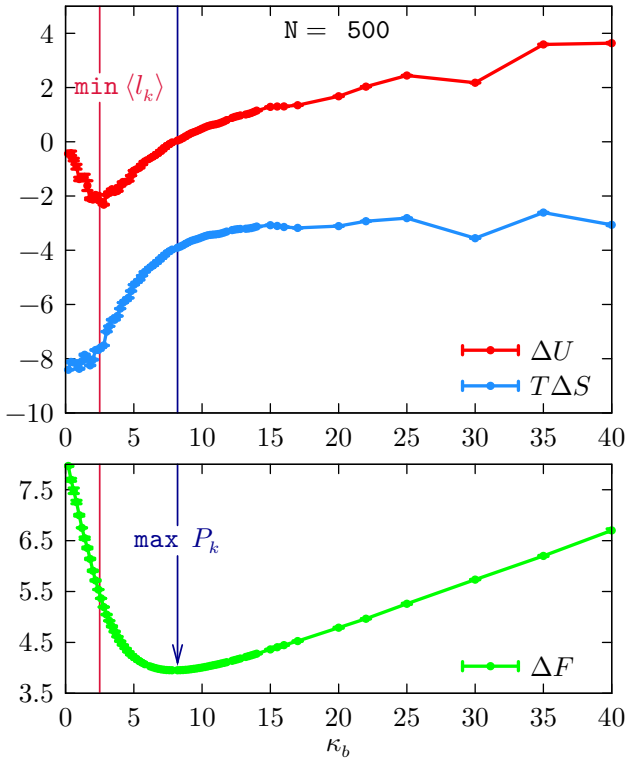


FIGURE 3.12: Top panel: κ_b dependence of the bending energy $\Delta U = \langle U \rangle_{knots} - \langle U \rangle_{unknots}$ and the entropic cost $T \cdot \Delta S = T(S_{knots} - S_{unknots})$ required to tie a knot into a semiflexible ring of contour length $N = 500$. The estimated relative errors are 2% or smaller. Bottom panel: free energy difference $\Delta F = (F_{knots} - F_{unknots})$ between knotted and unknotted rings as a function of κ_b. As expected the minimum of ΔF is located at the value of κ_b at which the knotting probability is a maximum (right vertical blue line). The estimated relative errors are 0.5% or smaller.

fully-flexible rings, these become progressively more compact, and so does the knotted region, which shrinks in contour length too. Non-local contacts inside the knotted region, which are inevitable because of the non-trivial topology, become more probable. The steric hindrance of the contacting strands in the knotted region has two competing effects. On the one hand it straightens the strands (as occurs in tensioned knotted rings [78]) and lowers their bending energy below average. This gives knotted rings an energy gain over unknotted ones. On the other hand, the topological restraints introduce an entropic cost that contrasts and eventually outweighs the energy gain as the bending rigidity is lowered. This competition produces the non-monotonic trend of the knotting probability seen in Fig 3.4.

3.4 Summary and perspectives

To summarize, we studied how the knotting probability of ring polymers varies as a function of the bending rigidity.

The study was motivated by the leak of previous systematic investigation of the effect of bending rigidity on the knotting of polymers rings. We accordingly addressed this problem by using a general model of self-avoiding semiflexible rings, and Monte Carlo sampling simulations.

We found that, when going from the rigid to flexible limit, we observe that the progressive compactification of the rings is neither paralleled by a steady increase of knots abundance, nor of knots size. In fact, the knotting probability, is non-monotonic and has a maximum at intermediate bending rigidities, when the persistence length is a few times larger than the bead size.

We thus profiled various metric and topological properties as well as the energetic and entropic contributions to the free-energy in order to clarify the unexpected result.

We concluded that as the bending rigidity is reduced, and the chain is compactified, the strands of the knotted region have a significant steric interaction due to the topological constraints. Such interaction has two opposite effects: on the one hand, it keeps the knotted region less bent than average, resulting in a bending energy gains. On the other hand the constraints inevitably decrease the entropy of knotted rings compared to unknots. It is the competition of the two effects that produces the striking and counter-intuitive non-monotonicity of the knotting probability.

For sufficiently long chains (of fixed bending rigidity) it is known that the occurrence of knots is inevitable ($P_k \to 1$) because long unknotted rings are entropically disfavoured. In particular, as $N \to \infty$ we expect $P_k \sim 1 - exp((f_{unknotted}(\kappa_b) - f(\kappa_b))N)$ where $f(\kappa_b)$ and $f_{unknotted}(\kappa_b)$ are the limiting free energies per bead of respectively the set of all rings and the subset of the unknotted ones. It would therefore be very appealing to analyse the interplay of energy and entropy for long rings and check whether the non-monotonic dependence of P_k on κ_b still holds in the large N limit [64]. As matter of fact, we will extend our study in this direction, in particular, studying if and how the profile of knotting probability changes for very large N. This study will be described in Chapter 4.

Other interesting issues related to the problem of knotting in semiflexible rings are the effect that the interplay of bending energy and entropy may have on the populations of knot types (knot spectrum) and the possibility of designing semiflexible knotted rings with non-homogeneous stiffness to control the size and position along the chain of the hosted knots, including composite ones.

Knotting probability of semiflexible rings: theoretical results

4.1 Introduction

In chapter 3 we discussed how the knotting properties of semiflexible chains have an unexpected non-monotonic dependence on bending rigidity. The results were based on numerical simulations of rings of beads.

Another commonly used polymer model is the ring of cylinders, also termed self-avoiding polygon. A well-known example of its application in knotting contexts is the work of Rybenkov *et al.* [43] where it was used to interpret experimental data on DNA rings in solutions of different ionic strengths. More recently, the group of T. Deguchi has systematically studied the topological spectrum of SAPs [69, 79], and provided approximate analysic expressions for how the knotting probabilities of the rings of cylinders depend on their thickness and contour length [80].

In this Chapter, I will discuss how a heuristic model mapping can be used to adapt the available analytical expressions for the knotting probabilities of self-avoiding polygons to the case of semiflexible rings-of-beads. This will allow us to extend the numerical results of the previous chapter, that were obtained for chains of up to $1,000$ beads, to much longer chain lengths. From such length extrapolation we obtain a further interesting result, namely that for sufficiently-long contour lengths, the knotting probability becomes bimodal as a function of the bending rigidity.
The content of this chapter is based on the publication of ref. [81].

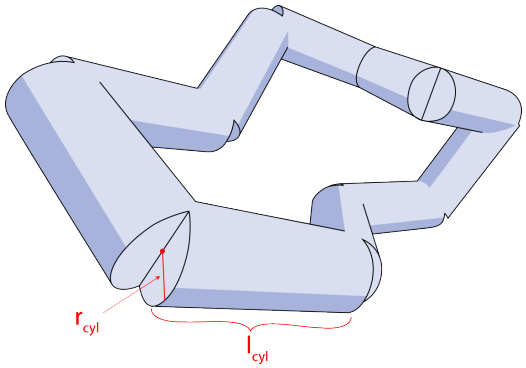


FIGURE 4.1: Representation of a self-avoiding polygon, or ring of cylinders. The shown case consist of $N_{cyl} = 10$, and characterized by its r_{cyl} and l_{cyl}. Image taken from ref. [81]

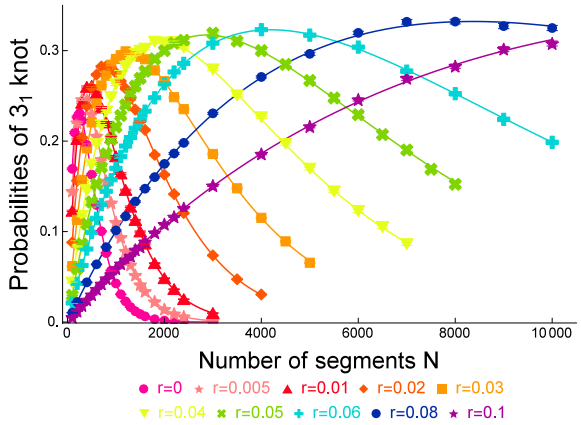


FIGURE 4.2: Knotting probability of trefoil knot, $P_{3_1}(N_{cyl})$, for value of radius equal to $0.0, 0.005, 0.01, 0.02, 0.03, 0.04, 0.05, 0.6, 0.08, 0.10$. The fitted curves are using the eq. 4.3 and the value of table 4.1. Image taken from ref. [80].

4.2 Model

Self-avoiding polygons Self-avoiding polygons (SAPs) are a circular succession of freely-jointed cylindrical segments, with radius r_{cyl} and length l_{cyl}, which is taken as length unit. Excluded volume effects are enforced by preventing any two non-consecutive cylinders from overlapping. An example is given in Fig. 4.1.

The work of ref. [80], presented a systematic study of the incidence of various knot types, k, as function of the number N_{cyl} and radius r_{cyl} of the cylinders. In particular, it was established that the occurrence probability, P_k, of the simplest prime knots: $3_1, 4_1, 5_1, 5_2$ could be well approximated by the following analytic expression:

$$P_k(N_{cyl}, r_{cyl}) = C_k(r_{cyl}) \frac{N_{cyl} - \Delta N_k}{N^0_{cyl}(r_{cyl})} e^{-\frac{N_{cyl} - \Delta N_k}{N^0_{cyl}(r_{cyl})}} \tag{4.1}$$

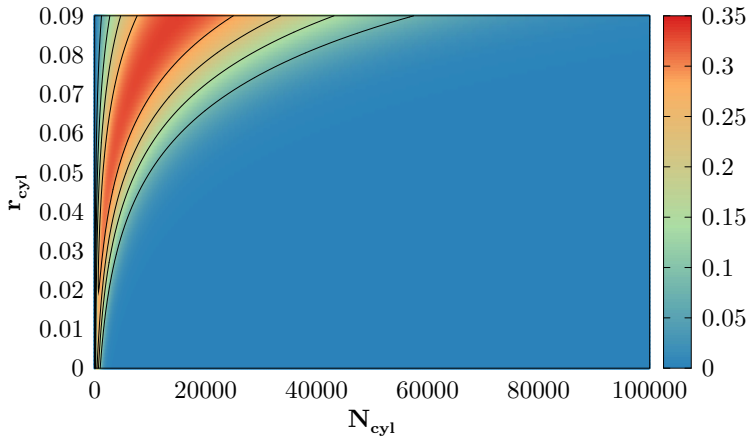


FIGURE 4.3: Observed knotting probability for trefoil knots for various combination of N_{cyl} and r_{cyl}, in two-dimensional density map. Image taken from ref. [81]

where the C_k's are the coefficients listed in table 4.1; ΔN_K's are finite size correction, that is of the order of $20 - 30$, and N^0 the characteristic knotting lengthscale [79]:

$$N^0_{cyl}(r_{cyl}) = 292.0e^{43.5r_{cyl}} \tag{4.2}$$

We recall, as described in ref. [82], that for on-lattice self-avoiding polygons N^0 is very large, typically of the order of 10^5. For sufficiently large N_{cyl} is possible to omit the finite-size correction. Thus, eq: 4.1 simplifies as:

$$P_k(N_{cyl}, r_{cyl}) = C_k(r_{cyl})\frac{N_{cyl}}{N^0_{cyl}(r_{cyl})}e^{-\frac{N_{cyl}}{N^0_{cyl}(r_{cyl})}} \tag{4.3}$$

and their results for the incidence of trefoil knots are represented by the curves of Fig. 4.2. Finally, it is worth to mention another connection with self-avoiding polygons defined on lattices. Specifically, for SAP thickness about equal to 0.125, i.e. $r \sim 1/8$, it was observed in ref. [80] that the ratios of knot coefficients C_{k_1}/C_{k_2} for a given pair of knots k_2 and k_2 become close to the universal ratios of self-avoiding rings on a lattice, which were evaluated in ref. [83].

Rings of beads We introduced the ring of beads model in Chapter 1 were we used it to describe self-avoiding semi-flexible chains. The model parameters were the radius of the impenetrable beads, $b/2$, the bending rigidity of the chain, κ_b and the number of beads N_{beads}. Using this model, as already mentioned, we showed that the knotting probabilities has a non-monotonic dependency on κ_b for chain lengths up

K	$C_K(r_{cyl})$	*parameters*		
		a_0	a_1	b_1
3_1	$a_0 \cdot (1.0 - a_1 \cdot \exp(-b_1 \cdot r_{cyl}))$	0.919 ± 0.003	0.327 ± 0.002	33.1 ± 0.8
4_1			0.1357 ± 0.0013	8.82 ± 0.27
5_1	$a_1 \cdot \exp(-b_1 \cdot r_{cyl})$		0.04387 ± 0.00042	20.81 ± 0.34
5_2			0.07741 ± 0.00046	21.91 ± 0.22

TABLE 4.1: Approximate expressions for the dependence of the C_K coefficient on the cylinders' radius, r_{cyl} for the simplest prime knot. Data are from ref. [80].

to 1000 beads, the longest we could sample efficiently, see Fig. 3.4. In the following section I will introduce an heuristic mapping with the SAP model that will allow to extrapolate our earlier results for much longer chains.

4.2.1 Mapping

An approximate mapping between chains of cylinders and chains of beads can be made by matching the length of the cylinders to the nominal Kuhn length of the chain of beads.

This heuristic mapping lie in the connection between the simple model in which bonds are free to rotate and bend, freely-jointed chain (FJC) and Kratky-Porod (KP) one, the latter is a discrete chains with bending rigidity, κ_b, and no excluded volume interactions. It is important to stress that the latter case represents an approximation of the ring of beads model; the implication of this approximation will be described later.

For long Kratky-Porod chains, metric observables such as the radius of gyration are equivalent to those of freely-jointed (i.e. fully-flexible) chains with the same contour length, but where bonds have the Kuhn statistical length [48], which is about twice the chain persistence length, $l_{Kuhn} \sim 2\kappa_b b$. Based on these considerations, a chain of N_{beads}, unitary diameter and bending rigidity κ_b can be mapped in a ring of N_{cyl} cylinders of radius r_{cyl}, that follows three rules:

$$\boxed{b \to r_{cyl} = b/2 \qquad \kappa_b \to l_{cyl} = 2\kappa_b b \qquad N_{beads} \to N_{cyl} = N_{beads} \cdot b/l_{cyl}} \tag{4.4}$$

which means that, if we conventionally take length of cylinder axis of SAPs as the unit length for such system, and the beads' diameter, b, as the one for rings of beads,

we have the following correspondences: $N_{cyl} = N_{beads}/2\kappa_b$ and $r_{cyl} = 1/4\kappa_b$.

Based on this mapping, the expression 4.3 for the knotting probability of SAPs can be adopted to rings of beads as follows:

$$P_k^{map} = C_k \frac{N_{beads}}{2\kappa_b N_{cyl}^0(1/4\kappa_b)} e^{-\frac{N_{beads}}{2\kappa_b N_{cyl}^0(1/4\kappa_b)}} \tag{4.5}$$

where N_{cyl}^0 is given in eq. 4.2 and C_k in table 4.1.

Range of validity We stress that this mapping is expected to hold only approximately because it relies on the notion of Kuhn segments defined for Kratky-Porod chains, which are free from excluded volume interactions.

As in other contexts, excluded volume effects are too challenging to account for exactly. This is because of their non-local character, as they introduce constraints between regions at arbitrarily large distance along the sequence.

Local excluded volume effects, instead, are more tractable and can be used to establish limits of validity for the above mapping.

For instance, in the chain of beads model, the thickness of the beads limits the angle θ between two bonds $\vec{b}_i, \vec{b}_{i+1}$ in $[0 : 2\pi/3]$ range, otherwise next-neighbouring beads along the sequence would overlap. Consider now, such angle constraint applied to a fully-flexible chain, then the average value of $\cos(\theta)$ is:

$$\langle \cos(\theta) \rangle = \frac{1}{Z} \int_1^{-1/2} d\left(\cos(\theta)\right) \cos(\theta) = \frac{1}{4} \tag{4.6}$$

As we mentioned in Chapter 1, the bending angles are statistically independent and the correlation function of two bonds $\langle \vec{b}_i \cdot \vec{b}_{i+k} \rangle$, at given sequence separation k, can be factorised and yield an exponential decay. We thus have that the persistence length $l_p \sim b/ln(4)$. This fix the bound in which the mapping breaks down: indeed, for $\kappa_b \lesssim 1$ the effect of the excluded volume cannot be neglected.

At the same time, large values of the bending rigidity, which correspond to persistence lengths comparable to the contour length of the rings, should be avoided too.

We recall that the parametrizations underpinning eq. 4.5 are based on extensive numerical data gathered for SAPs of up to 10 000 cylinders and thickness of up to 0.1 [80]. These therefore, might represent practical additional limits for the scope of

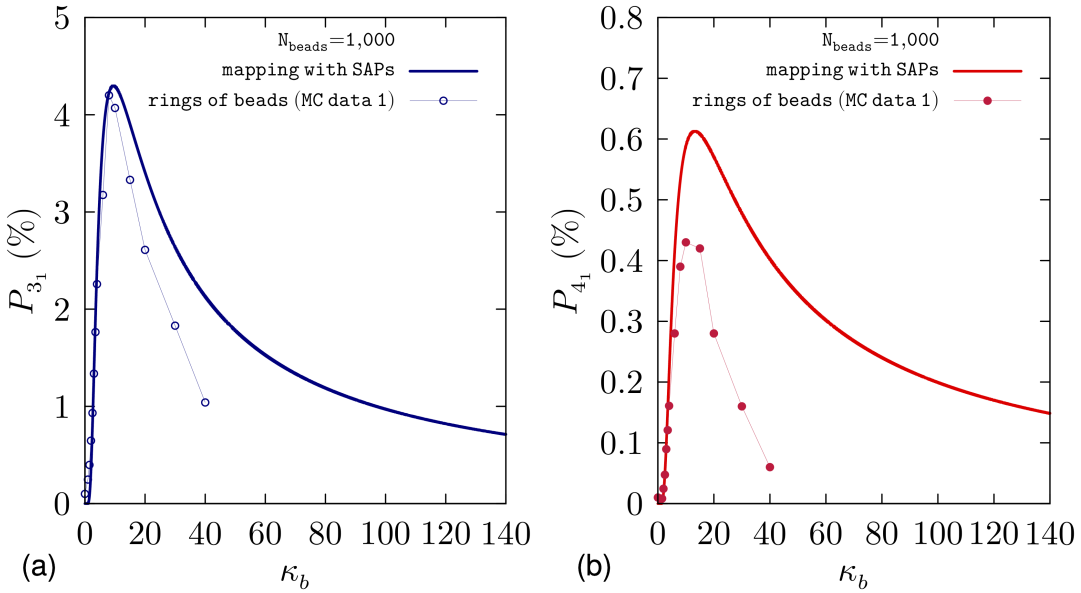


FIGURE 4.4: Observed and predicted probabilities of (a) trefoil knots and (b) figure-of-eight 4_1 knots for rings of 1 000 beads as a function of their bending rigidity, κ_b. The data points of the observed probabilities are from the Monte Carlo simulations of Chapter 3. Image taken from ref. [81]

applicability of the expressions in eq. 4.5.

Finally, we also neglected ΔN_K in 4.3; for this reason, the approximation does not work well below $N_{cyl} \sim 25$, which occurs for $\kappa_b > N_{beads}/50$, since typically $\Delta N_K \sim 20 - 30$.

4.3 Comparison of knotting probabilities

We first apply expression 4.5 to predict the probability of 3_1 and 4_1 knots in rings of 1000 beads, and compare it with the knotting probabilities obseserved in simulations of Chapter 3 and ref. [72].

Predicted and measured knotting probabilities The predicted and measured knotting probabilities of 3_1 and 4_1 in rings of 1,000 beads, are shown in Fig. 4.4. Expressions 4.4 is able to reproduce the non-monotonicity of the P_k: the curves are well superposed for small κ_b that is for equivalent SAPs made of several (and thick) cylinders. For the case of 3_1 knots it also reproduces the height of the peak. For both topologies, there is a good agreement that holds for small values of bending rigidity, or thick cylinders, in a range of $2 < \kappa_b < 20$. Then, discrepancy of the curves increases progressively beyond the peak, due to a systematic overestimate of

the predicted knotting probabilities. As matter of fact, beyond the upper bound value, the effective number of cylinder become comparable to the fine size correction, and, as we discussed in the previous section, this invalidates the use of the asymptotic expression 4.4.

This corroborates the results presented in Chapter 3 and ref. [72]. It also allows us to provide a physical explanation behind the non-monotonicity: the qualitative trend and the mapping show that, when increasing κ_b two distinct effects arise: on one hand, there is reduction of the Kuhn segments which suppresses the knotting probability, but on the other hand, increasing κ_b also reduces the thickness of the cylinders, and this instead promotes the knotting propensity at fixed number of cylinders. These considerations clarify that the non-monotonicity observed in Chapter 3 for rings of beads can be explained in terms of the competition between the number of cylinders and their thickness in the corresponding SAP model.

Predicted and measured knotting probabilities comparison with linear chain results In Fig. 4.5 we provide a further comparison of the knotting probability profiles with the knotting probabilities of long model DNA *linear* chains up to $5 \cdot 10^5$ beads . All the detailed of the latter model are available in the work of Rieger and Virnau [45].

For our model, we took $N_{cyl} = N_{beads}/(2l_p) = N_{beads}/(2\kappa_b)$, with $\kappa_b = 11.673$ and N_{cyl} is given by the stated equivalent DNA size, in base pairs, divided by 294 bp, the number of base pairs in a Kuhn segment of 100nm, as in Riegers model. The model DNA chains considered were linear, however, the bending rigidity had been set equal to $\kappa_b = 11.673$ so to reproduce the experimental knotting probabilities of DNA chains after they spontaneously circularise in solution. Such circularised DNA rings, are customarily taken as representative of knotting in equilibrium [43]. As shown in Fig. 4.5, the agreement is very satisfactory, both for P_{3_1} and P_{4_1} , especially at small numbers of cylinders, namely < 1000. For longer chains the predicted trefoil knotting probability is progressively underestimated, while the good accord of the 4_1 curves persists.

The overall agreement of the trends of predicted and observed knotting probabilities in Fig. 4.4 and Fig. 4.5 validates the parametric expressions of eq. 4.5 which thus

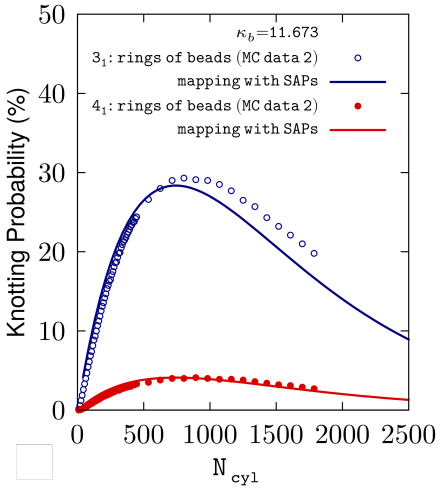


FIGURE 4.5: The data points from the Monte Carlo study of ref. [45] where linear semi-flexible chains of beads where parametrised to reproduce the knotting properties observed in experiments on circular DNA. For such system, the number of nominal Kuhn segments (corresponding to a chain stretch of 100nm or 294 bp) was taken as the number of equivalent cylinders, N_{cyl}, and $\kappa_b = 11.673$. Image taken from ref. [81]

could be of practical use in theoretical and computational contexts. As already mentioned, arguably, the most informative use of the mapping of eq. 4.4 is to predict the knotting properties of semi-flexible rings for previously unexplored combinations of ring lengths and bending rigidities.

4.4 Bimodal knotting probability

We shall now use the analytic expression of eq. 4.5 to explore the dependence of the knotting probability on κ_b for chain lengths much larger than those we could address with the MC simulation of ref. [72].

4.4.1 Extension to long rings

Firstly, in Fig. 4.6, panel a, we present the predicted knotting probability for rings of $N_{beads} = 10000, 20000$ and 50000, these are much longer than the ones of ref. [72]. Indeed, we increase of one order of magnitude the largest explored in the ring of beads model. For $N_{beads} = 10000$ and 20000, the profile probability is unimodal trend, however it becomes more complex with increasing number of segments.

For rings of $N_{beads} = 50000$ the 3_1 knots probability is multimodal, with a pronounced minimum at $\kappa_b \sim 10$. This minimum falls in the region directly covered by the tested

parametrizations of ref. [80] ($\kappa_b > 2.5$) and therefore appears to be a genuine, non-artefactual property of long rings of beads.

Origin of the bimodal knotting probability profile The unexpectedly rich behaviour is conveyed more comprehensively by the density map of Fig. 4.6b. One notes that the color-coded regions, corresponding to various probability intervals, bifurcate and become bi-lobed as N_{beads} is increased above $\sim$ 25000. This feature precisely translates in a bimodal knotting probability profile, such as that for $N_{beads} = 50000$ in Fig. 4.6a, since each lobe hosts a local maximum of the knotting probability, and in the region in between them falls a minimum.

The crossover from unimodal to bimodal profiles is conveniently illustrated by considering how the trefoil knotting probability of self-avoiding polygons of Fig. 4.2 is transformed by the model mapping (eqs: 4.4).

The knotting probability manifold of SAPs, shown in Fig. 4.2b, is simpler than the one of semiflexible rings, shown in Fig. 4.7b, because it is free of bifurcations. In fact, for a given number of edges, N_{cyl}, there is a single maximum of the trefoil knotting probability that occurs for a finite value of the cylinders radii. The locus of points corresponding to these maxima is marked with circles in the contour map of Fig. 4.7a. In the same two-dimensional parameter space of the contact map, the locus of points corresponding to semiflexible rings of beads of given length, N_{beads}, and various bending rigidities, κ_b, has a simple geometry. Specifically it corresponds to a straight line of equation $r_{cyl} = N_{cyl}/(2N_{beads})$.

Two such lines, representative of rings of 5000 and 30000 beads are shown in Fig. 4.7a. The line with the largest slope corresponds to the shortest ring, and it crosses the circled curve (the locus of the maxima at fixed N_{cyl}) in one point only. Consequently, the corresponding probability profile presents a single maximum as a function of κ_b. As the length of the rings of beads is progressively increased, the slope of the line decreases until it eventually crosses the locus of the maxima twice (but no more), as it happens for the 30000 beads case. The double crossings clearly reflect in a bimodal probability distribution as a function of κ_b, as illustrated in 4.7c.

The analysis of eq. 4.5 for asymptotically-long rings of bead clarifies that the two

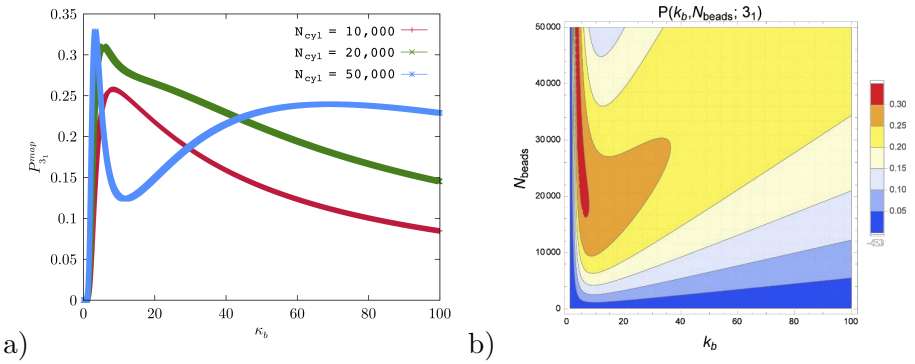


FIGURE 4.6: (a)Predicted probability of trefoil knots for rings of 10,000, 20,000 and 50,000 beads as a function of their bending rigidity, κ_b. (b) Density map of the probability of trefoil In the right figure, the contour plot of the derivative of $P(\kappa_b, N_{beads})$ with respect to κ_b is colored with knots for rings of beads as a function of their length and bending ridigity. Image taken from ref. [81]

maxima of the trefoil knotting probability occur at:

$$\kappa_b^* \sim 10.9/\log(N_{beads}/6351) \tag{4.7}$$

$$\kappa_b^{**} \sim N_{beads}/584 \tag{4.8}$$

These correspond to rings of, respectively, increasing flexibility and increasing rigidity as N_{beads} gets larger. In the plot of Fig. 4.7a, the first maximum corresponds to the crossing point at the largest r_{cyl} values. Note, that because the locus of the maxima (circled line) grows slowly with N_{cyl}, the location of the extremal value of $\kappa^*$of eq. 4.7 will likewise change slowly with N_{beads}.

Physical Interpretation A physical interpretation of the crossover from unimodal to bimodal profiles can be given in terms of the interplay of the equivalent number of cylinders, N_{cyl} and the characteristic topological lengthscale, N^0, which depends on the cylinders thickness, see eq. 4.2.

This is best discussed for trefoil knots for which the C_K coefficient, has only a mild dependence on r, see Table 4.1. From eq. 4.2 and 4.3 one has that the maximum possible incidence of trefoil knots is reached when N_{cyl} is about equal to N^0. In fact, fewer segments prevent from reaching the target knot complexity, while more segments favour more complex knot types. The κ_b-dependence of N_{cyl} and N^0, which is controlled by the mapping of eq. 4.5, produces a non-monotonic, unimodal profile of N_{cyl}/N^0 versus κ_b.

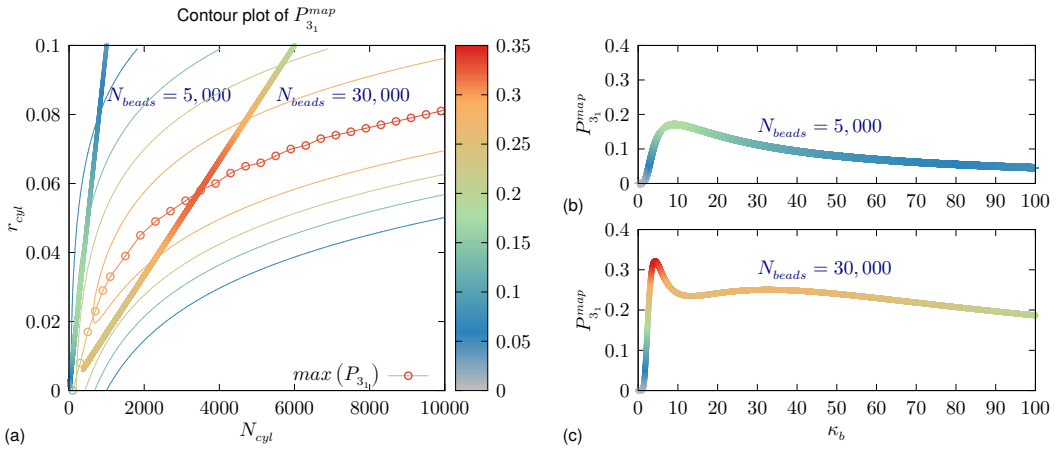


FIGURE 4.7: (a) The curved lines are the contour lines of the probability for trefoil knots in self-avoiding polygons of radius r_{cyl} and with N_{cyl} edges. The straight lines are the locus of points, that in the model mapping of eqs. 4.4, correspond to rings of 5,000 and 30,000 beads of various bending rigidities, κ_b. The actual probability profiles as a function of κ_b are shown in panels (b) and (c)

When $N_{beads} \lesssim 17,260$, the peak value of N_{cyl}/N^0 remains below unity and a single maximum of the knotting probability is observed.

However, when N_{beads} is increased beyond $17,260$, the maximum of N_{cyl}/N^0 becomes larger than 1. There will then be two increasingly separated values of κ_b for which N_{cyl} matches N^0, and each of them will correspond to a peak in knotting probability. The peak at the larger κ_b value corresponds to relatively short and thin SAPs, while the other corresponds to SAPs that are not only longer, but also thicker.

4.5 Summary and perspectives

We used a simple physical mapping to adapt the known asymptotic expressions for the knotting probabilities of self-avoiding polygons to the case of semiflexible rings of beads. The mapping thus yielded analytical, though necessarily approximate, expressions for how the incidence of the simplest knot types, 3_1, 4_1, 5_1 and 5_2 depends on the number of chain beads, N_{beads}, and bending rigidity, κ_b. These two parameters are, in fact, mapped to an effective number of cylinders with an associated effective thickness in an equivalent self-avoiding polygon model.

We validated the approach by comparing the predicted knotting probabilities with those established in previous studies where stochastic simulations were used for rings with 100 - 1,000 beads and for bending rigidities in the $0 \leq \kappa_b \leq 40$ range. The

observed agreement was very satisfactory except when κ_b, becomes large enough that the effective number of cylinders becomes too small for the asymptotic expressions to hold.

The results are useful in several respects. From a practical point of view, the mapping can be used to obtain immediate estimates of knotting probabilities in contexts where the actual determination by stochastic simulations (molecular dynamics or Monte Carlo) is particularly onerous, e.g. for semiflexible chains of thousands of beads. This ought to be particularly helpful when one wishes to precondition a simulations setup and choose the best combination of parameters for a target abundance or complexity of knots.

More in general, the mapping is useful because it offers further insight into the knotting propensity of rings of beads by interpreting them in terms of equivalent polygons of different length and thickness. As an illustration of this, we examined the non-monotonic κ_b-dependence of the knotting probability of rings of hundreds of beads, hat was recently reported by some of us based on Monte Carlo simulations for rings of up to 1,000 beads [72].

Here, by using the mapping with SAPs, we could extend the study to chains of tens of thousands of beads and, by doing so, we observed a previously unsuspected complex phenomenology. Specifically, we found that the non-monotonic profile of the knotting probability switches from unimodal to bimodal for rings longer than about 20,000 beads. It would be most interesting to verify this prediction in the future, when computational resources will allow for a systematic exploration of the κ_b-dependence of the knotting probability profiles for such long rings.

Finally, we remark that, for simplicity, the results presented here have been purposely based on a minimalistic mapping between the two models. This is an aspect that can certainly be extended in future investigations, with more precise estimates for the number of cylinders and their thickness in the equivalent SAP model.

Knotted supercoiled DNA rings

5.1 Introduction

In the previous chapter I reported on a systematic exploration of how the knotting probability of polymer rings is influenced by their bending rigidity.

In this chapter, I move on the second main topic of the thesis. I will broaden the perspective to a more detailed coarse-grained model to study the interplay between knotting and torsion in closed double-stranded DNA rings. The structural organization of bacterial plasmids is profoundly affected by homeostatic catalytic processes involving DNA transactions such as transcription, replication and recombination. These are processes which require strands separation of the DNA double helix, and opening the DNA leads to accumulation of torsional strain ahead and behind the transcriptional machinery. If not relaxed, this torsion can eventually prevent the DNA or RNA polymerases involved in these processes, and leads to detrimental effects.

The best known example is the level of negative supercoiling, $\sim -5\%$, found *in vivo* in bacterial DNA, which is maintained by a family of enzymes called topoisomerases [84–86]. This baseline negative supercoiling creates an interplay between DNA twist (local winding around the double-helix axis) and writhe (winding of the double-helix axis around itself), that affects the structural organization of the DNA. In addition, topoisomerase action can affects DNA entanglement, introducing for example knots [13, 17, 87–89] or removing it [89–91]. The type and the aboundance of knots can vary with experimental conditions, the specific topoisomerase action on well-defined patterns of crossings [87] as well as DNA length [13, 17, 88, 89, 91–93], which is also a key determinant of knot complexity in viral [11, 23, 25, 94] and eukaryotic DNA [24].

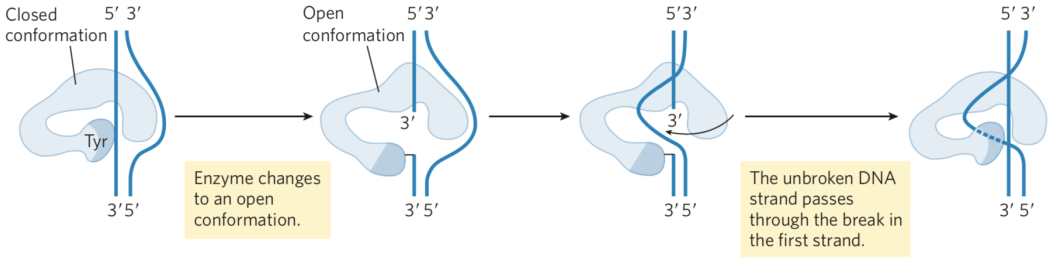


FIGURE 5.1: Bacterial topoisomerase I mechanis: breaking one DNA strand, passing the unbroken strand through the break, then resealing the break. The action of topoisomerase I lead one strand to turn around the other, hence it changes L_k by a unit. Image taken from ref. [111].

The structural constraints associated with supercoiling and knotting also have functional implications. Plectonemes, for instance, control the degree of branchedness of DNA rings [95–98] and this, in turn, affects the contact probability of *loci*, close contact, at large genomic separations [95, 99–102]. In addition, the mechanics of superhelical stress, which can have long-range effects [103, 104], can facilitate the unzipping of AT-rich regions that are upstream of genes, and thus assist the binding of promoters [103, 105–110].

5.1.1 Topoisomerases

The topoisomerases (topo) are enzymes able to increase or decrease the extent of the winding of the two DNA strands around each other, and thus control the torsional stress accumulated during biological processes. The topo activity introduces transient topological changes in the linking number (defined in eq. 2.3 of Chapter 2). Since the ability to regulate the latter is essential for cell survival, these enzymes are ubiquitous in all organisms.

Topoisomerases are classified in two categorises:

- Type I topoisomerases, which introduce transient single-strand breaks and either allow one strand to rotate around the other (swivelling) or let the unbroken DNA strand passes through the broken strand (enzyme-bridge strand passage) and then reseal the latter. The change of linking number by one unit, see Fig. 5.1;

- Type II topoisomerases, which cut both strands of the DNA double helix, pass another unbroken double strand DNA filament through the gap, and then reseal

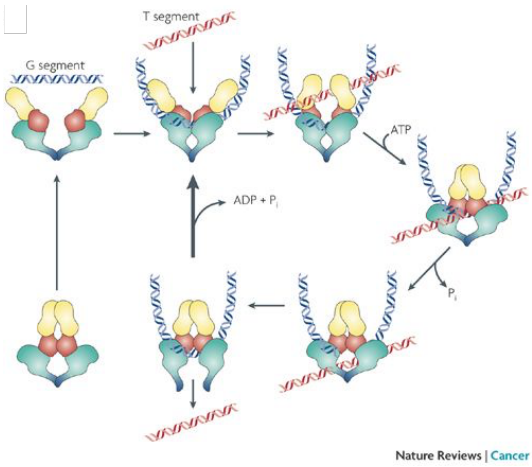


FIGURE 5.2: The hypothesized general mechanism of topoisomerase II features the passage of one intact duplex DNA segment through a transient double-strand break in another segment. Image taken from ref. [112]

it. These topoisomerases rely on ATP hydrolysis and change the linking number by two units, see Fig. 5.2.

Both types of enzymes have been found in all the domains of life. They are further grouped into subclasses according to the specific mechanism employed to process the DNA, and to their structure.

For instance, type I topoisomerases are further divided into three types: Type IA, which removes only negative supercoiling, $(-)S$, and does it by an enzyme bridge strand passage; Type IB and Type IC [113] which remove positive, $(+)$ and $(-)$ supercoiling through swiveling but differ in structure [114].

The Type II enzymes are Type IIA, which are the only enzymes that can unknot or unlink DNA [90], and IIB. The latter is comprised by the enzymes of the archaea domain, and are exclusively able to relax both $(-)$ and $(+)$ supercoiling. Table 5.1 summarizes the salient features of the enzymes.

The combined action of topo I and topo II enzymes sets the level of DNA supercoiling, unlinks the daughter chromosome in the replication of circular bacterial DNA, and removes accidentally formed DNA knots. For instance, Type II topoisomerases are essential for separating the two daughter strands of circular DNA molecules during the very end of DNA replication, when two single strands are catenated and cannot separate into the two daughter cells.

Besides this crucial role, they are also responsible for resolving the potential tangle where two double-stranded DNA cross each other [30, 91, 115].

Several studies have been modelling the conditions that lead to the simplification of entanglements: they have suggested that Topo II bind to regions characterized by high

Type I			
	enzymes	domains	
Type IA	topoisomerase I, III Reverse Gyrase	bacteria archaea eukarya	Relax only (−) supercoiling
Type IB	topoisomerase I	bacteria eukarya	Relax both (+) and (−) supercoiling
Type IC	topoisomerase V	archaea	Relax both (+) and (−) supercoiling DNA repair

Type II			
	enzymes	domains	
Type IIA	DNA Gyrase	bacteria archaea	Introduce (−) supercoiling and removes knots
	topoisomerase II	eukarya	Relax only (−) supercoiling
	topoisomerase IV	archaea	Relax only (−) supercoiling
Type IIB	topoisomerase VI	archaea	Relax both (+) and (−) supercoiling

TABLE 5.1: Classification of topoisomerases. There are two main categories: Type I and Type II. In Type I, there is further classification which divide the topoisomerases that release the torsion via enzyme-bridge stand passage, form the ones that work through swivelling. In Type II, the subclasses are mainly due to the recent discovered of topoVI, which are limitated to archaea domain.

level of torsional stress, where the different portions of DNA filaments are interlocked, without being chemically bonded. These strands are hooked-juxtapositions and generally, when the action of Topo II is applied in these regions it leads to elimination of knots. Therefore, those regions are indentified as fingerprints of knots [34, 116–118]. Under these assumptions, the presence of supercoiling ideally increases the occurence of portions with high level of stress, such as apical loops or juxtapositions, and therefore they become ideal targets of topoisomerases [29, 34, 63, 117, 119, 120].

5.1.2 Knotting in supercoiling DNA

Knots have been mostly associated with detrimental functional effects [13, 88, 121], such as stalling DNA replication and transcription and yet, their emergence is statistically inevitable.

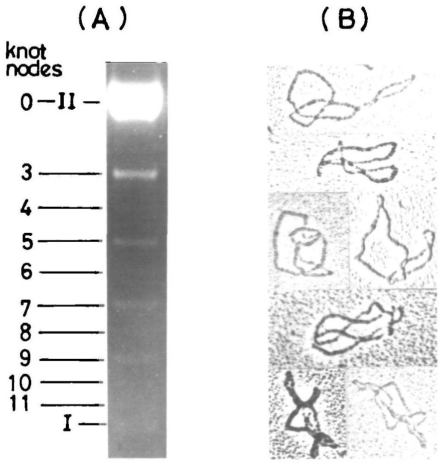


FIGURE 5.3: A) Agarose gel electrophorese. It shows the separation of the knotted nicked DNA species with various numbers of nodes from the unknotted nicked circular form. To indicate the number of knot nodes, they analyse the bands by electron microscopy. B) Typical knotted DNA molecules by electron micrographs. Image taken from ref. [122]

The experimental works of Shishido *et al.* [11, 122], found that, *in vitro* $1-10\%$ of plasmid DNA is a knotted species. This occurs in presence of negative supercoiling with defective gyrases. Besides, they have also shown that only the knots with an odd number of crossings are produced, see Fig. 5.3.

In addition, it has been proved in 1997 by Rybenkov *et al.* [90], that topo II enzymes can reduce the fractions of knotted circular DNA molecules compared with the corresponding equilibrium values, by more than 80 times.

The general mechanism of DNA entanglement simplification by topoisomerases is still far from being completely understood, and one open question is how are the topological constraints affecting the structure of the DNA.

More precisely, the constraints due to the supercoiling and the ones due to knotting, affect DNA in similar ways: they modify the linking number and generate deformations in twist and writhe. In living bacteria, the topoisomerases IV (see table 5.1), decrease the $(-)$ supercoiling by performing strands passages and introducing knots. Concurrently, the topoisomerases type I work to adjust the level of supercoiling that has been changed by the accidental presence of a knot. The key aspect is how topoisomerases, that can only sense DNA local structural information, may get hints whether performing a strand passage in a given place, will succeed at bringing the DNA knotting much below the equilibrium value.

Intriguingly, recent modelling studies by Dietler's and Stasiak lab, have pointed to a primary role of supercoiling in the removal of DNA knots, too [46, 123, 124]. The Stasiak's lab explored the interplay between supercoilng and knotting in DNA molecules with stochastic simulations of coarse-grained DNA filaments tied in trefoil

knots, and found also that the systematic accumulation of twist leads to a tightening, or localization, of the knot. They suggested that the local DNA simplifying action (recognition and strand passage) of topoisomerases is possible thanks to such knot localization. In particular, they suggested that strongly localized knotted portions are characterized by higher curvature than unknotted ones, hence it will be easly recognized by topo enzymes. The size reduction of the knot is the result of the competition bewteen the energy cost of increasing the bending and the entropy lost of accessible configurations [46, 123, 125].

These results add a novel appealing layer to the functional role of supercoiling and suggest that further and more detailed investigations are needed to explain the successful simplification of entanglements. Specifically the interplay of supercoiling and knots, especially the more complex ones reported in plasmids [122], is largely unexplored and key questions are still unanswered.

For instance: how do knots more complex than trefoils affect the branchedness of supercoiled rings? Would the latter be increased by the large writhe of complex knots, or would it be suppressed by the topological constraints? Also, what is the effect of an intricate topology on the internal dynamics of plasmids, and how does it differ from the one of supercoiling? Do knots trap the system in long-lived states and, if so, what are their characteristics?

To address these questions, we consider a mesoscopic model of a DNA duplex and studied, through molecular dynamic simulation how the interplay entanglement and supercoiling affects the geometry and the dynamics of the duplex.

The material of this chapter is pubblished in the article of ref. [126].

5.2 Materials and methods

5.2.1 Model

We consider 2kbp-long DNA rings modelled with oxDNA [47, 50, 51], an accurate model for DNA double filaments based on mesoscopic representation of nucleotides and their interactions that has been presented in Chapter 1. The oxDNA model makes it possible to gather multi-ms trajectories for kbp-long DNAs while retaining the key structural details responsible for the frictional [127] and cholesteric effects [27]

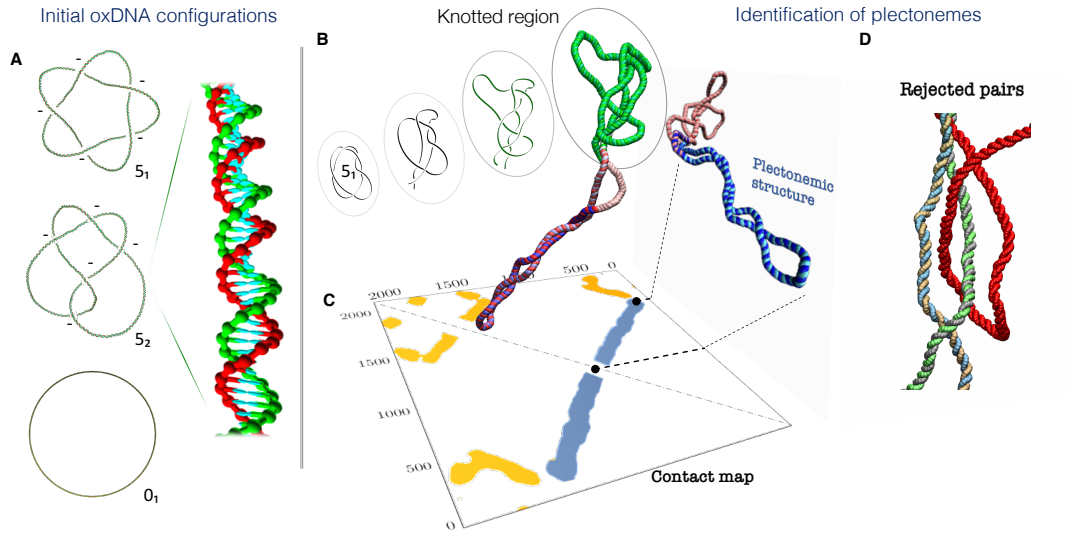


FIGURE 5.4: (a) Initial configurations of the supercoiled double-stranded DNA rings for the three considered topologies, 0_1, 5_1 and 5_2. The latter two are left-handed, i.e. the topological sign of their projected crossings is negative, as indicated. The 5_1 and 5_2 snapshots have been edited to highlight the over- and underpasses. The mesoscopic structural representation of the oxDNA model is illustrated in the inset, which shows a magnified portion of one of the rings. The twist was uniformly adjusted for each of the three cases to yield the same level of negative supercoiling (-5%). (b-d) Identification of the knotted and the plectonemically-wound regions for a typical 5_1-knotted supercoiled conformation, shown in the foreground. (b) The knotted region (green) is the shortest portion that, after suitable bridging of the termini, has the same (5_1) topology of the entire ring. (c) The plectonemically-wound region (blue) is found by using the contact map to identify long superhelical regions ending in a short apical loop and that are free of *cis* or *trans* entanglement, as in the case of panel (d).

arising from self-contacts in the knotted or superhelical regions.

Specifically, we focus on DNA rings with 5% negative supercoiling, which is the typical homeostatic level of supercoiling in plasmids as shown in ref [128], and tied in 5-crossings left-handed knots (5_1 and 5_2 topologies), a complex form of entanglement previously reported in 4kbp-long pBR322 plasmids [11,88]. In addiction, we consider also ring torsionally relaxed and unknotted to use as comparison.

5.2.2 Initial setup

The initial conformations are generated with the following two-tier scheme. First we produce the centerline of the double-stranded DNA rings by using the KnotPlot software [56] to create smooth, symmetric forms of 5_1 and 5_2 knots. They have the same complexity but they belong to the torus and twist group, which typically show different physical behaviour, from mechanical resistance to sliding friction to pore-translocation compliance [129, 130]. With these more complex (respect to earlier studies on 3_1-knotted DNAs) knots we can explore the effects of a larger writhe

and more numerous minimal crossings on the branchedness and dynamics of DNA rings. We also consider the unknotted ring, 0_1, as reference see Fig. 5.4A. To be consistent with experimental observations on pBR322 plasmids [11, 88] the chirality of the five-crossing knots is set to be left-handed, i.e. projected crossings have negative sign. The circular centerlines obtain from KnotPlot, is discretised in 2000 segments, and are next turned into the oxDNA double-helical representation by a fine-graining procedure, where each segment is mapped into the six interaction centers of the two paired nucleotides [127], see inset in Fig. 5.4A. In this fine-graining procedure, the average twist between consecutive bases is adjusted differently for each topology to yield the sought level of supercoiling, i.e. 5% , as discussed in detail in ref. [123, 131] and described below.

5.2.3 Linking number

To set the baseline value of supercoiling, first we recall that in torsionally-relaxed duplex, the double helix makes one complete turn about its axis every ~ 10.5 base pairs. For a planar circular (and hence unknotted) DNA ring, the linking number, L_k is therefore given by an integer close to $L_C/10.5$ were L_C is the ring contour length expressed in base pairs. Because the writhe of a planar circle is zero, by using the Calugareanu-Fuller relationshhip, $L_k = Tw + Wr$ (eq. 2.4) we have that twist and linking number coincide, $Tw = L_k$. The latter expression has to be aptly generalized in order to be applied to knotted chains. Indeed, a knot cannot be arranged as a planar circle. As matter of fact, knots are characterized by a non-zero writhe, hence the number of twists of torsionally relaxed knotted rings is not simply given by $L_C/10.5$ but it is corrected by subtracting a factor equal to $\langle Wr_0 \rangle_K$. This value of writhe only depends on the knot type [132]. For each considered topological state, we accordingly adjusted the twist uniformly to include the two different cases, torsionally-relaxed DNA rings and negatively supercoiled ones. We set the twist from the formula:

$$Tw = L_k - Wr + \langle Wr_0 \rangle_K \tag{5.1}$$

where $\langle Wr_0 \rangle_K$ is average writhe of torsionally-relaxed rings of knot type K. For 5_1 and 5_2 knots, these values are:

$$\langle Wr_0 \rangle_{5_1} = \frac{50}{7} - \frac{6}{7} \sim 6.2857$$
$$\langle Wr_0 \rangle_{5_2} = \frac{20}{7} + \frac{12}{7} \sim 5.285$$

where the rational expression are based on the conjecture of ref [132]. Since the values of the writhe of a knot and its mirror image differ by a sing flip, thus $\langle Wr \rangle_{5_1}^{left} \sim -6.2857$, and $\langle Wr \rangle_{5_2}^{left} \sim -5.285$.

Finally, we set the the relative amount of supercoiling, ΔL_k equal to -0.05, the typical homeostatic level in bacterial plasmids, by uniformly spreading the global excess of twist overall the bases.

5.2.4 Molecular dynamics simulations

For each of the six combinations of knot types (0_1, 5_1 and 5_2) and torsional states (relaxed and negatively-supercoiled) we collected ten different Langevin dynamics trajectories at $T = 300$K. The dynamical evolution was integrated with the LAMMPS package [54], using the implementation of Henrich *et al.* [53] and default values for the mass of the interaction centers, m, solvent viscosity, η and of the time step in the Langevin-type rigid-body integrator [53], $0.01\tau_{LJ}$, where $\tau_{LJ} = \sigma\sqrt{m/\epsilon}$ is the Lennard-Jones characteristic time, σ is the length unit, equal to $10^{-10}m = 0.8518nm$, and $\epsilon = \kappa_B T$ is the energy unit. Each trajectory had a typical duration of $\sim 2 \times 10^7\, \tau_{LJ}$. For the analysis, we omitted the initial relaxation phase of duration $10^6\tau_{LJ}$. The cumulative time span covered by all simulations was $1.2 \times 10^9\tau_{LJ}$ and required about 1.4×10^6 equivalent CPU hours on the Intel-based high-performance computing cluster (Ulysses) based in SISSA, Trieste.

5.2.5 Diffusion coefficient and time mapping

In general, because of the concurrent presence of various dynamical regimes at different spatial scales, a correspondence between real time units and simulation time in coarse-grained models can be set only approximately. Here, we established the mapping *a posteriori* by matching the diffusion coefficient of the simulated DNA rings

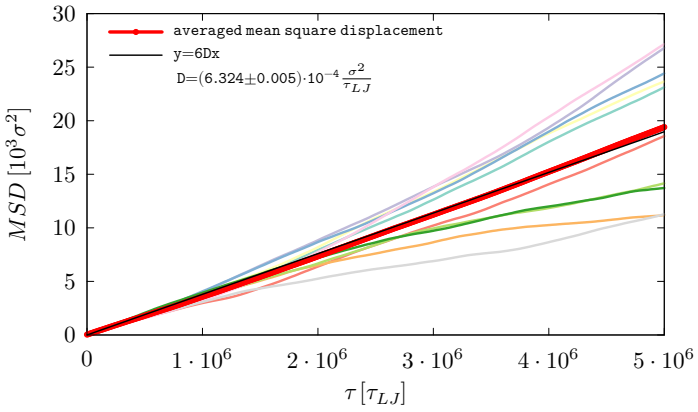


FIGURE 5.5: The thick red line is the mean square displacement (MSD) of the center of mass of 2kb-long supercoiled unknotted rings, at different time lags. This MSD was calculated by averaging the MSD curves of 10 different trajectories (colored thin lines). The diffusion coefficient, computed from the linear best fit (black line), is $D_{theory} \sim (6.324 \pm 0.005) \times 10^{-4}\sigma^2/\tau_{LJ}$.

with the analogous experimental quantity. This conservative approach is expected to be more apt for the spontaneous dynamics of large systems than mappings based on the diffusivity of oligonucleotides [127]. Within the oxDNA setup, the diffusion coefficients of supercoiled or torsionally-relaxed 2kbp-long unknotted rings at 1M monovalent salt, D_{theory}, is obtained by computing the mean square displacement of center of mass. Hence, at different lag time and for each trajectory:

$$MSD = \langle \left(\vec{X}_{CM}(\tau + t) - \vec{X}_{CM}(t) \right)^2 \rangle \tag{5.2}$$

where τ is the lag and the average is done on the number of pairs at time lags. Then, averaging over all the trajectories, it is possible to compute the $\langle MSD \rangle$, see Fig. 5.5 (red line). Finally, as result of a linear interpolation of the latter, we compute the diffusion coefficient: $D_{\text{theory}} \sim 6.324 \times 10^{-4}\sigma^2/\tau_{LJ}$, where $\sigma = 0.8518$nm [47, 50, 51]. Experimental measurements for DNA rings of similar length yield $D_{\text{exp}} \sim 7 \times 10^{-12}$ m^2/s [133]. By equating D_{theory} and D_{exp}:

$$D_{\text{exp}} \sim 7 \times 10^{-12} m^2/s = 6.324 \times 10^{-4}\sigma^2/\tau_{LJ} \tag{5.3}$$

one therefore has $\tau_{LJ} \sim 7 \times 10^{-11}$s.

5.2.6 Metric observables

As an overall metric observable we used the root-mean-square gyration radius, $R_g = \sqrt{\langle \frac{1}{2n^2} \sum_{i,j=1}^{n} (\boldsymbol{r}_i - \boldsymbol{r}_j)^2 \rangle}$ where i and j run over the $n = 4000$ nucleotides, $\boldsymbol{r}_i$ is the position of the center of mass of the ith nucleotide, and the $\langle\rangle$ brackets denote the average over the configurations visited in the trajectories.

For the metric relaxation dynamics we computed the time-lagged autocorrelation function of R_g

$$C(\tau) = \frac{\langle (R_g(t+\tau) - R_g) \times (R_g(t) - R_g) \rangle_t}{\langle (R_g(t) - R_g)^2 \rangle_t} \ , \tag{5.4}$$

where $R_g(t)$ is the instantaneous gyration radius at time t, τ is the time lag, and $\langle\rangle_t$ denotes the average taken over simulation time for the various trajectories. The characteristic timescale was computed as the integral of $C(\tau)$. To limit the effects of the noisy tail of $C(\tau)$ the integral was computed from $\tau = 0$ up to when $C(\tau)$ drops below 10^{-2} for the first time.

5.2.7 Detection of plectonemes

The torsional stress induced by negative supercoiling can leads into folding of the DNA in the form of a supercoil, called plectonemes. To identify these plectonemically-wound region, if any, we designed an algorithm based on persistency of subsequent patterns in the contact maps. We generalized previous approaches [98, 134, 135] and used the multi-step strategy sketched in Fig. 5.4C-D and Fig.5.6.

We first constructed a contact map for the DNA centreline using a tolerant cut-off distance of $40\sigma \sim 32$nm, about three times larger than the typical superhelical diameter [93], see Fig 5.6A.

Next, to identify superhelical regions, we searched for clusters of contacts forming bands perpendicular to the contact map diagonal. These bands correspond to dsDNA stretches in spatial proximity and with antiparallel orientation, i.e. opposite directionality in the an oriented ring.

For each band we then identified its apex, which is the site, i, associated to the longest uninterrupted ladder of contacts $\{(i+\Delta, i-\Delta), (i+\Delta+1, i-\Delta-1), (i+\Delta+2, i-\Delta-2), ...\}$ starting at a sequence separation, Δ, not larger than 400bp. This value is an upper bound that we empirically chose to avoid excessively large apical loops.

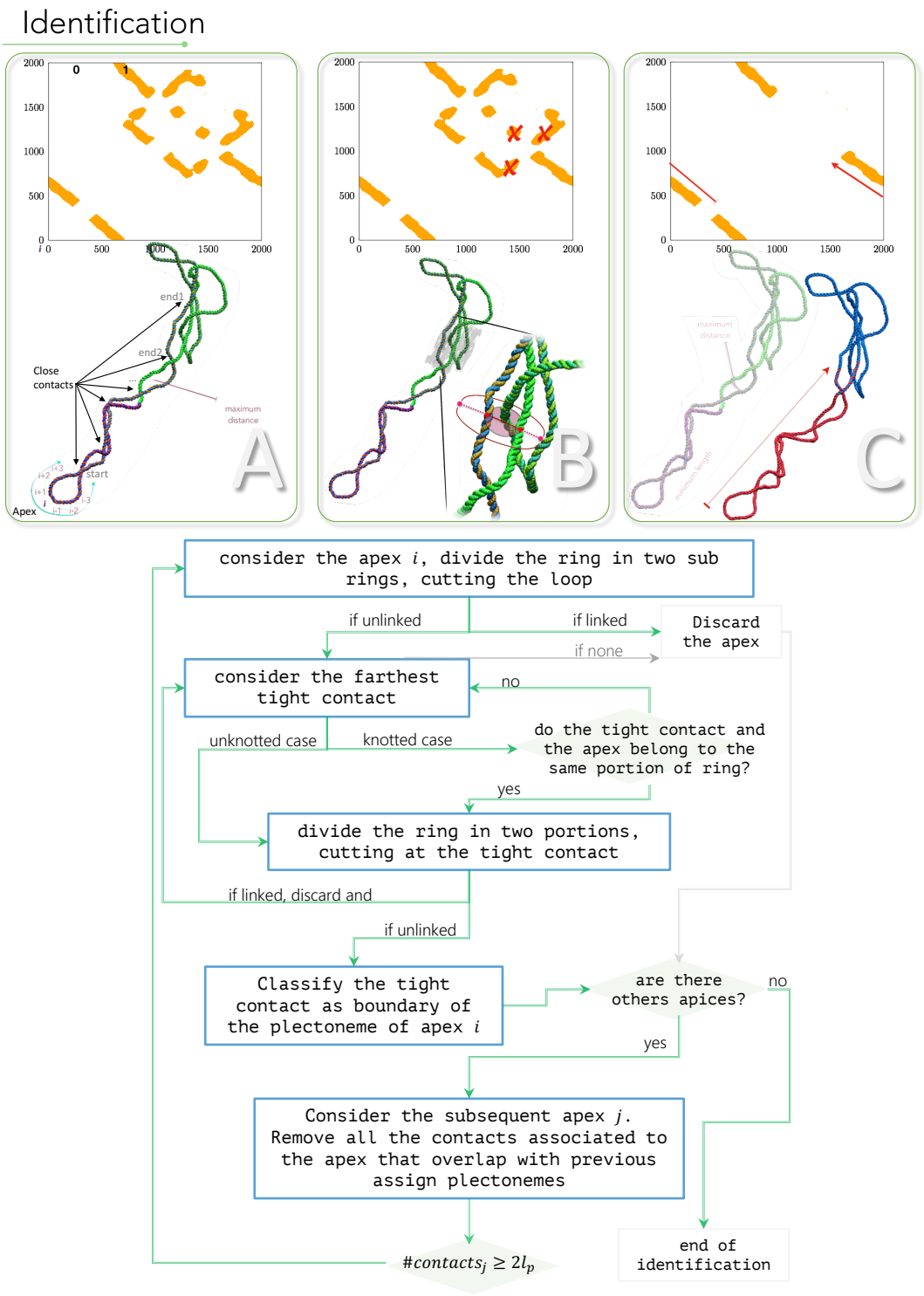


FIGURE 5.6: The identification of plectonemes is based on the distance between the points of the centreline. Firstly we compute a contact map with a tolerant cutoff of 40σ. Then, we look for apices, which are identify as the bead from which develops a continuous bands perpendicular to the diagonal. We classify them according to their length: the persistence of consecutive contacts. Later, we identify and classify the tight contacts, which are the one closer in space than 10σ, and that do not have any other strand in the middle. (B) Once identify the putative apices and all the relative tight contacts we proceed iteratively as summarize in the panel flowchart.

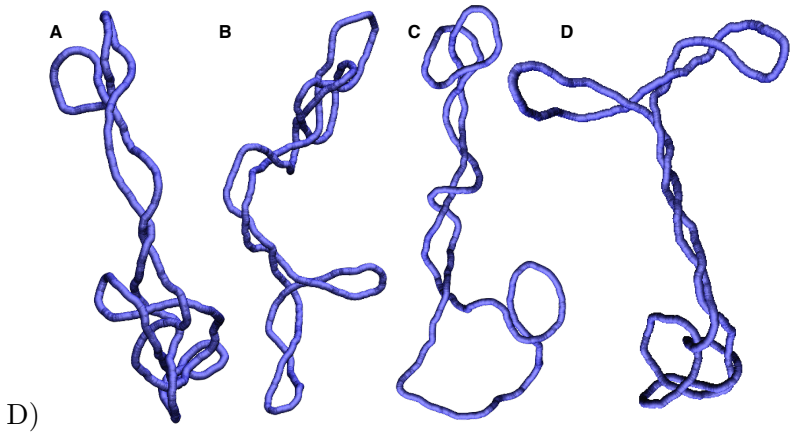
Likewise we do not count short range contacts, i.e. contacts at sequence separation $\Delta < 100$bp since they are spatially and sequencly close and do not provide information on the 3D structure. In addition, omitting them speed up the computation procedure that requires the computation of the distance between all the pairs. Within each band with $\Delta < 400$bp, we then searched for the presence of pairs even closer in space, we refer to them as close contacts, a signature feature of supercoiling, see Fig. 5.6. Specifically, we searched for the contacts at distance smaller than $7.5\sigma(\sim 6$nm), and took the contacting pairs with the smallest and largest sequence distance from the apex as the endpoints of the putative superhelical region, Fig 5.6A. The putative plectoneme, formed by this region and the bridging apical loop was then checked to be free of entanglement, Fig 5.6B, by testing that progressively longer portions of the plectoneme had no physical linking [136] with the remainder of the ring. If this was not the case, the distal endpoints (those farthest from the apex) were progressively backtracked until the region became disentangled. Notice that, this requirement will exclude all the plectonemes that have their loop entangled in the ring.

The plectoneme assignment was then carried out in an iterative non-overlapping manner by ranking the putative plectonemes by the contour length of the superhelical region and disregarding instances where the latter was smaller than 300bp,i.e. twice the persistence length, Fig 5.6C. The region with the longest superhelix was then assigned as the first plectoneme. Next, the distal endpoints of the remainder putative plectonemes (if any) were then backtracked to eliminate eventual overlaps with the sites assigned to the first plectoneme. The length ranking and selection was repeated and the second plectoneme was assigned so on, until exhaustion of the putative plectoneme set.

The iterative scheme allowed for the unsupervised detection of one or more plectonemes in practically all supercoiled configurations except for a small subset (0.5% of 5_1-knotted instances and even smaller for 5_2-knotted and unknotted ones) with non well formed superhelical regions or atypically large apical loops, see Fig. 5.7.

5.2.8 Topological observables

To locate the knotted region along the ring we used the bottom-up search scheme described in Chapter 2.



D)

FIGURE 5.7: Examples of supercoiled conformations for which no plectonemes are identified by the algorithm described due to excessively large apical loops or non well-formed superhelical regions. The topologies of the shown conformations are as follows: 5_1 for panels A and B, 0_1 for panel C and 5_2 for panel D.

The procedure has the goal to find the smallest portion of the ring that, after closure, has the same topology of the entire ring. Firstly, starting from portions of only few nucleotides and systematically expanding to longer ones.

For closing the considered portion we used the minimally interfering closure scheme [58], where the termini of the portion are bridged either with a straight segment or via a path involving the convex hull, depending on their proximity.

Once we obtain a closed curve we establish the topology using Alexander polynomials evaluated at $t = -1$ and $t = -2$.

5.3 RESULTS

5.3.1 Conformational variability

We provide in Fig. 5.8A typical snapshots, representative of the various degrees of branching found in the 5_1- and 5_2-knotted DNA rings, along with instances without knots (0_1 case). Our simulations reveal that the conformational variability of knotted supercoiled DNA rings is significant, it explore various degrees of branching, up to 3 branches, for all the topologies. This large conformational variability is consistent with ref. [106], indeed the latter study shows how wide is conformational variability of minicircles of unknotted DNA containing 336bp, it is remarkable that the DNA was able to adopt such a wide variety of conformations despite the short-length and the accumulated torsion of due to supercoiling, even varying the latter in a range of

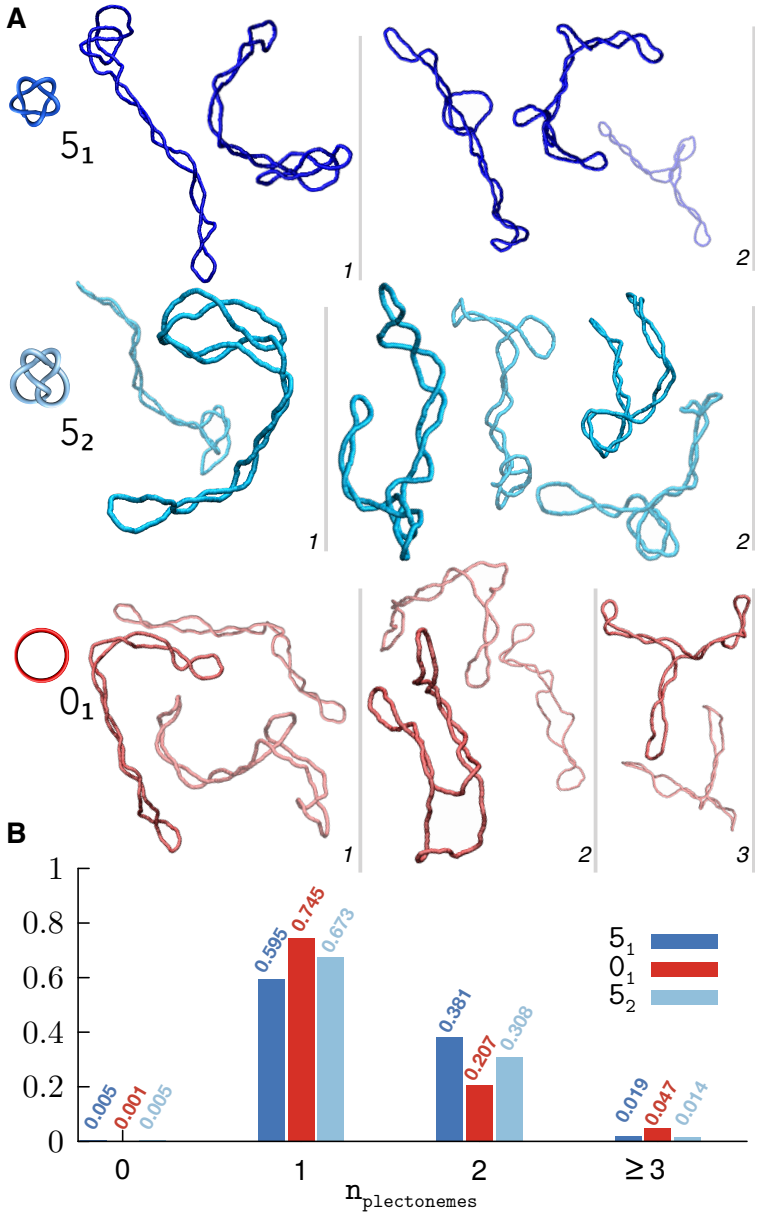


FIGURE 5.8: a) Typical snapshots of supercoiled DNA rings for the three considered topologies. The conformers are grouped by the number of plectonemes (in italics), which increases from left to right, and are shown in colors of different saturation for visual clarity. b) Normalised histogram of the number of plectonemes observed for each topology.

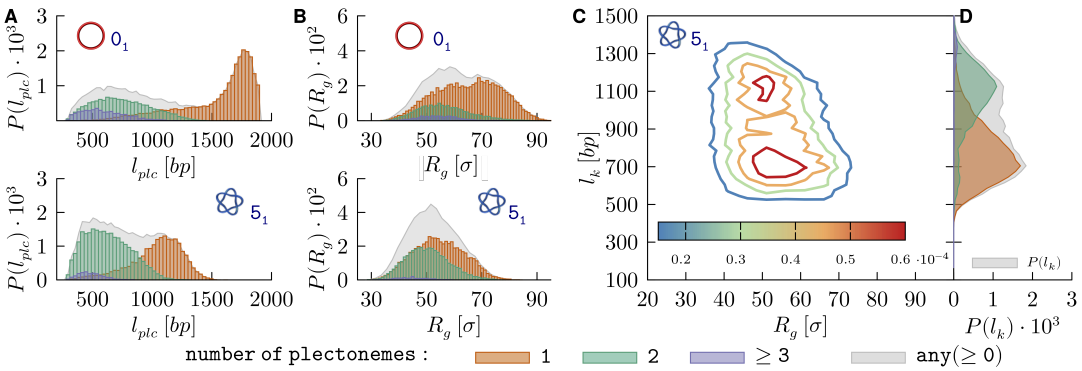


FIGURE 5.9: Probability distributions of the plectonemes' length, l_{plc} and gyration radius, R_g, for supercoiled rings with (A) unknotted and (B) 5_1 topologies. The conditional probabilities for 1 to 3 plectonemes are shown with coloured histograms, see legend, while the normalised combined distribution is shown in grey. (C) Normalised joint probability distribution of R_g and knot length, l_k. (D) Marginal probability distribution of l_k.

$-6 < \Delta L_k < +3$.

In ref. [95] it has been shown via Monte Carlo Methods, that even relatively short DNA molecules, < 3 kilobases, increasing the supercoiling results in the formation of unbranched, rodlike plectonemic conformations in order to accommodate torsional stress due to DNA underwinding. Therefore, it is interesting to examine the relationship between knottedness and the number of branches, or plectonemically-wound regions, because of the competing elements that govern it.

The lobes, apices and clasps inherent to complex knots can favour plectonemes by serving as nucleation points, while the conformational restrictions of the topological constraints can inhibit plectonemes formation. The histograms in Fig. 5.8B clarify that, at this contour length, unknotted DNA rings are actually somewhat richer in plectonemes than knotted ones; in particular, instances with 3 or more plectonemes are practically found in unknotted rings only.

More conspicuous differences related to topology are found in the distributions of plectoneme lengths, l_{plc}, and gyration radius, R_g, see Fig. 5.9A,B, for the unknotted rings and 5_1 ones and Fig. 5.10 for the 5_2 topology. In particular, the conditional distributions of l_{plc} for the common single- and double-plectoneme states are little superposed for unknotted rings, but overlap substantially for knotted ones.

The length of the plectonemes is also different across the 2kbp-long knotted and unknotted rings. For examples, plectonemes longer than 1500bp are common in uknotted rings but rare in 5_1-knotted ones (50.7% and 0.05% of the populations,

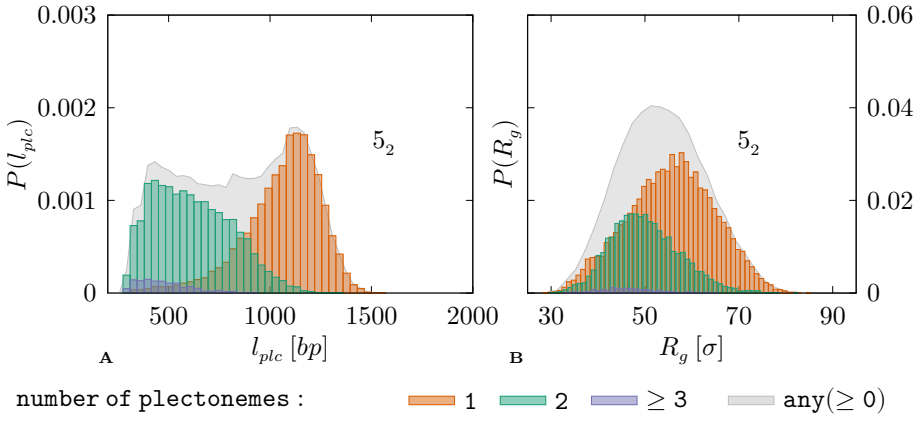


FIGURE 5.10: Normalised probability distributions of (**A**) the plectoneme's lengths, l_{plc}, and of (**B**) the gyration radius, R_g for 2kbp-long supercoiled rings with 5_2-knots.

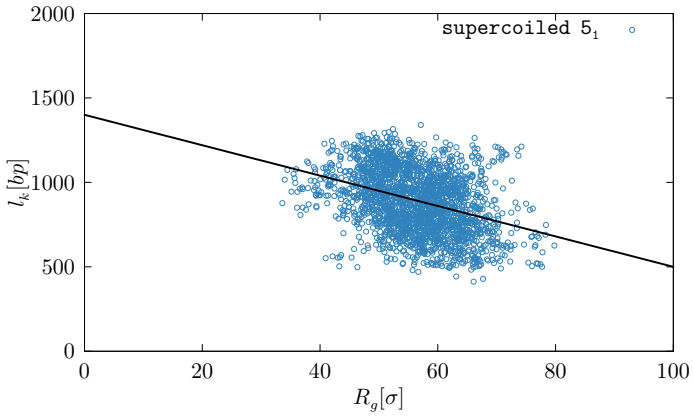


FIGURE 5.11: Scatter plot of knot length, l_k, versus gyration radius, R_g, for 2kbp-long supercoiled rings tied in 5_1 knots. The best linear fit (black line) has correlation coefficient $r = -0.36$.

respectively). Conversely, conformers with only one plectoneme, and shorter than 1000bp, are uncommon in unknotted rings but abundant in 5_1-knotted ones (6.3%and 22% of the populations, respectively).

These differences could be of practical relevance, since they could be exploited in imaging, e.g. cryo-em, experiments to tell apart knotted from unknotted plasmids when supercoiling is present. Such discrimination is generally beyond the scope of gel electrophoresis, the method of choice for DNA topological profiling (but for a notable exception see [137]).

Knot length,l_k, plectoneme length and gyration radius are metric quantities that are correlated due to the fact that the size of the ring is fixed. The interplay is presented in Fig. 5.9C-D. The joint probability distribution in panel C shows that l_k and R_g are anticorrelated, a property observed in other polymer systems too [72, 76].

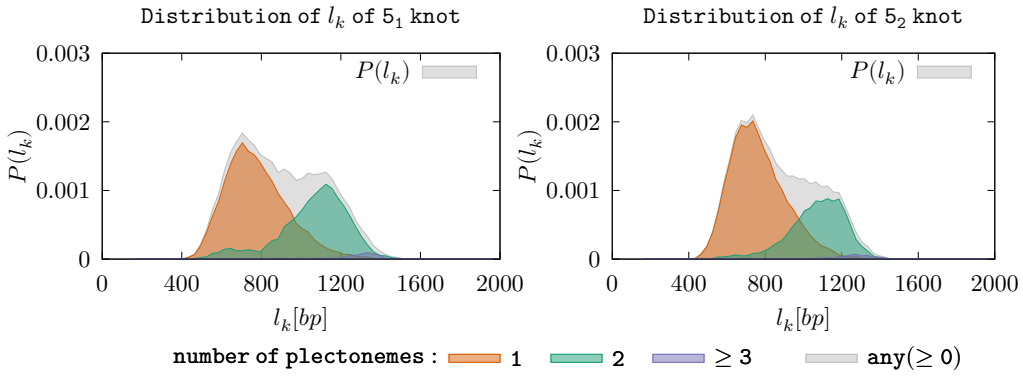


FIGURE 5.12: Normalised probability distribution of the knot length, l_k for 2kbp-long supercoiled 5_2-knotted DNA rings.

In Fig. 5.11 the anticorrelation is shown in a scatter plot of l_k and R_g, for the 5_1 topology. The argument presented in Chapter 3 can be applied also here: shrinking the knot size leads to an increase of the unknotted portion of the ring, which will be characterize by large value of R_g. What is specific of supercoiled knotted DNA rings is, instead, the presence of two peaks in the joint $l_k - R_g$ distribution. The peaks' origin is clarified by their marginal (projected) l_k distributions subdivided for number of plectonemes, see Fig. 5.9D. Specifically, the dominant peak, for $l_k \sim 700$bp, is mostly associated to single-plectoneme states, while the peak at larger knot lengths ($l_k \sim 1100$bp), corresponds to states with two or more plectonemes. Analogous results for the 5_2 topology are shown in Fig. 5.12.

The inverse correlation of l_k and the number of plectonemes is understood by noting that the knotted region, which is the shortest *uninterrupted* portion of the ring accommodating the essential crossings (or essential tangles), must also include all intervening loops between the crossings except for the longest one. In supercoiled rings this remainder loop typically coincides with a plectoneme. Because the average plectoneme length decreases as their become more numerous, one has that l_k is shorter for states with a single plectoneme.

Overall, we can consider the equilibrium ensemble dominated by two different states: conformers with single-large plectoneme, and a localized knot, and conformers with two short plectonemes. These two are well characterized by the distributions of the metric observables, that allows for discerning between the two conformes types.

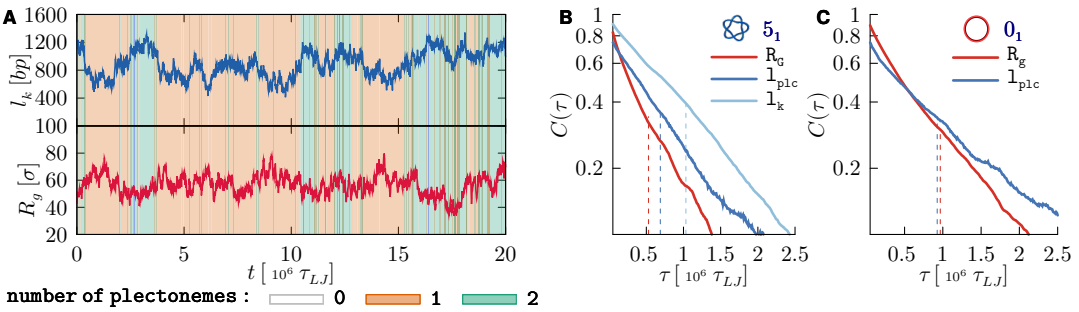


FIGURE 5.13: (A) Typical temporal traces of the length of the knotted region, l_k and the gyration radius, R_g, from a trajectory of supercoiled 5_1-knotted ring. The background is coloured according to the instantaneous number of plectonemes, see legend. (B) Semi-log plot of the autocorrelation functions, based on data from all trajectories, of R_g, l_{plc} and l_k of supercoiled 5_1-knotted rings (B) and of R_g, l_{plc} for unknotted ones (C).

5.3.2 Time evolution of metric and knot-related properties

The oxDNA model allows us to perform molecular dynamics simulation, which gives as the opportunity to look at the time of evolution of metric and topological properties. The data shown in Fig. 5.13A are a kinetic counterpart to the static, or ensemble, view given above of the interplay of the knotted region, the number of plectonemes and the gyration radius.

One notes that over the typical duration of a trajectory ($2.2 \times 10^7 \tau_{LJ}$ corresponding to about 1.5ms) both l_k and R_g have significant fluctuations, and clearly of opposite sign. These are accompanied by several changes in the number of plectonemes, as conveyed by the colored background.

A more quantitative analysis of the characteristic timescales of these variations is given in Fig. 5.13B,C. These panels present the autocorrelation curves of R_g, l_k and of l_{plc}. The latter, was used in place of the number of plectonemes (on which it clearly depends) for its broader range of values, which makes it more amenable to the autocorrelation analysis. By integrating the autocorrelation curves, one has that the characteristic times of l_k and R_g are respectively equal to $1.03 \times 10^6 \tau_{LJ}$ and $0.53 \times 10^6 \tau_{LJ}$, while for plectonemes' length it is $0.69 \times 10^6 \tau_{LJ}$. Consistent with visual inspection, these timescales are all of the same order, about $10^6 \tau_{LJ}$, which is much shorter than the duration of each simulated trajectory.

It is interesting that unknotted rings have a somewhat slower internal kinetics than knotted rings, cf. panels B and C in Fig. 5.13. In fact, the characteristic times of R_g and l_{plc} are, respectively 80% and 35% longer for unknotted rings (the same holds for

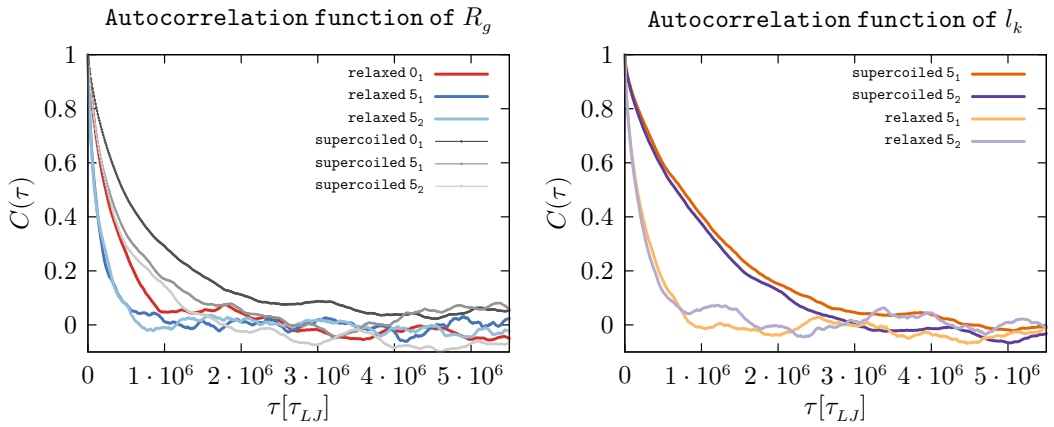


FIGURE 5.14: The autocorrelation functions, based on data from all trajectories, of R_g and l_k of supercoiled knotted rings.

R_g in the torsionally-relaxed case, see Fig. 5.14).

The result is not obvious, as one might expect a slower internal dynamics for knotted rings due to the friction of their self-contacts. It can be explained by considering that a finite portion of topological constraints necessarily uses up a finite portion of the chain, and therefore knotted rings have a shorter effective contour length than unknotted ones and a smaller gyration radius too (see Fig. 5.9). This, in turn, reflects in a reduced breadth of the relevant conformational space and hence a faster relaxation kinetics.

From the above analysis of overall metric and topological properties we conclude that, at physiological supercoiling, 2kbp-long knotted rings have enough conformational freedom to fluctuate spontaneously between two main states, related to the peaks in Fig. 5.9C, differing by knot size as well as the length and number of plectonemes. The characteristic timescale of these variations is $\sim 3 \times 10^6 \tau_{LJ}$, corresponding to about 0.02ms, see Fig. 5.15.

5.3.3 Slowly-moving boundaries of the knotted region

To understand the kinetics of the concerted variations of knot size and the number of plectonemes, we decouple entanglement and supercoiling by studying the time evolution of DNA rings where either the knot or supercoiling were present, but not both of them, as in the kymographs of Fig. 5.16 panel A and B, for the torsional relaxed knotted ring and the supercoiled unknotted one, respectively.

Considering that the trajectory covers a timespans of $2 \times 10^7 \tau_{LJ} \sim 1.5$ms, it is

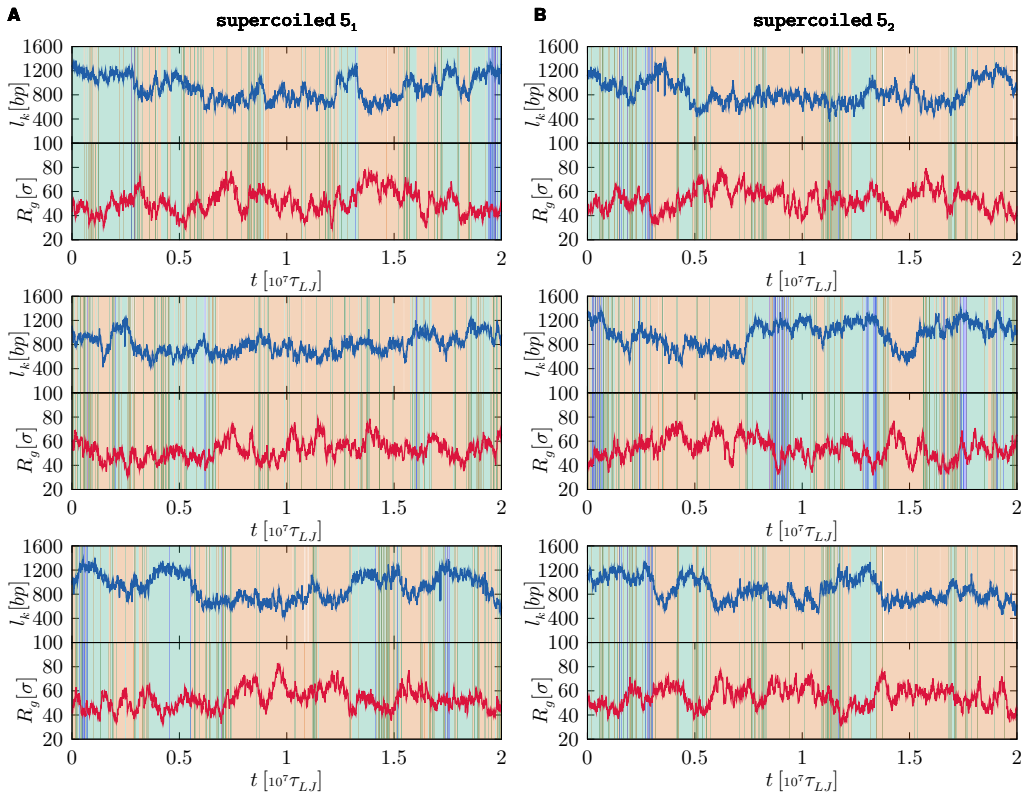


FIGURE 5.15: Temporal traces of the length of the knotted region, l_k, and the gyration radius, R_g, for 2kbp-long supercoiled rings with (**A**) 5_1 and (**B**) 5_2 topologies. The color-coded background reflects the number of plectonemes present: orange for 1 plectoneme, green for 2 plectonemes, as in the legend of Fig.5.13A.

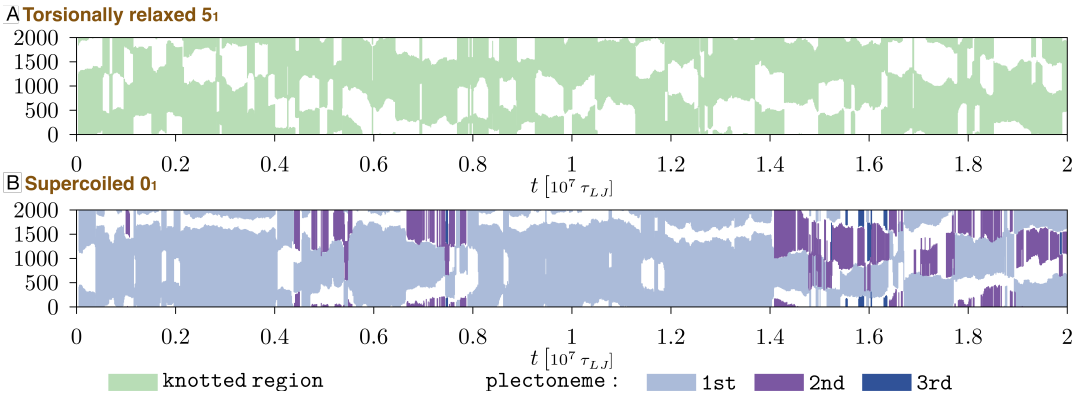


FIGURE 5.16: Kymographs showing the typical time evolution along the ring contour of knotted and plectonemically-wound regions, see legend for color code. The two kymographs are for: (A) torsionally-relaxed 5_1-knotted rings and (B) supercoiled unknotted rings.

clear that the contour motion of the knotted region occurs over timescales that are comparable to those discussed previously. These, we recall, were the relaxation time of the gyration radius and knot length, of about $10^6\tau_{LJ}$, and the changes between the tight and delocalised knotted states, which occur at intervals of about $3\times10^6\tau_{LJ}$, see Figs. 5.13B. Likewise, the diffusion of supercoiled regions along the ring has the same characteristics: the diffusion occurs on timescale comparable with the one of Figs. 5.13C.

However, the interplay of the knots' constraints and supercoiling in the ring changes dramatically the time evolution of these regions: the kymograph clearly indicates that there exists an additional relevant kinetic process besides those discussed before, namely a surprisingly slow stochastic motion of the knot along the ring contour. Note, in fact, that the region covered by the knot at the beginning of the simulation in Fig. 5.17A (bp600-bp1800) still has not moved appreciably by the end of the trajectory (bp700-bp2000).

We thus conclude that in supercoiled DNA rings, the contour motion of the knotted region is slower than these other processes by an order of magnitude (and likely more since the practical bound is the duration of the simulated trajectories).

Besides the knot, Fig. 5.17A shows the concurrent evolution of plectonemically-wound regions, too. Note that most of the time, there is a single long plectoneme that spans the ring portion complementary to the knotted region; but the latter can occasionally nest a plectoneme too. The kymograph clarifies that such nested instances have the following properties: (i) they occur in addition, and not in substitution, of

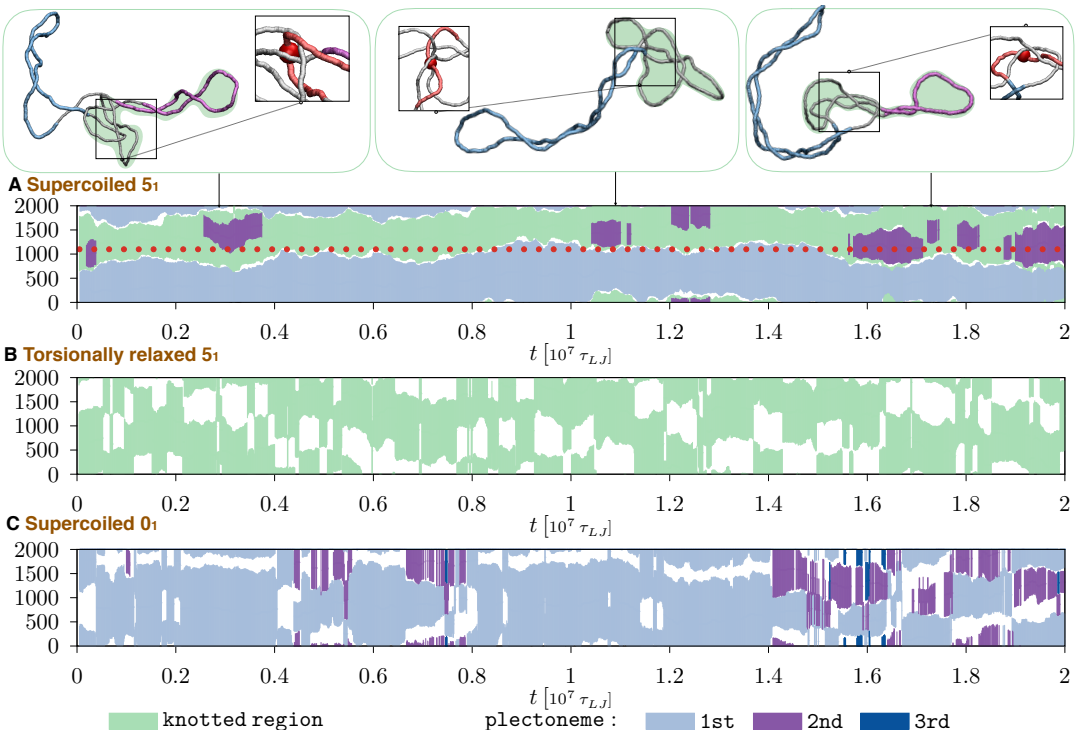


FIGURE 5.17: Kymographs showing the typical time evolution along the ring contour of knotted and plectonemically-wound regions, see legend for color code. The three kymographs are for: (A) supercoiled 5_1-knotted rings, (B) torsionally-relaxed 5_1-knotted rings and (C) supercoiled unknotted rings. The boundaries of the knotted and the main plectonemically-wound regions of case (A) are noticeable stabler than for case (B) and (C) due to persistent interlocking of multiple strands. This is illustrated in the snapshots above panel (A), where the same region at the knot-superhelix boundary (bp1000-bp1200, highlighted in red in the insets) remains entangled with other ring portions throughout the trajectory. The midpoint of this region is marked with a red bead in the insets and with a dotted red line in panel A.

the typically longer "dominant" plectoneme that complements the knotted region; (ii) their characteristic lifetime is $7 \times 10^5 \tau_{LJ}$ and (iii) consecutive appearances are separated by intervals of highly variable duration.

Overall, the several observed fluctuations of plectonemes' number and length during the entire trajectory are in line with the relatively fast relaxation dynamics of R_g and l_{plc} of Fig. 5.13. Strikingly, these variations are accompanied by a noticeable persistence of the plectoneme boundaries, which mirrors the one of the knotted region.

These features are ubiquitous across the collected trajectories for both 5_1 and 5_2 topologies, as shown in other examples of Fig.5.18. One concludes that, irrespective of their torus (5_1) or twist (5_2) character, these supercoiled knotted rings have persistent boundaries between knotted and plectonemically-wound regions. This, in turn, poses the question of which of these two components, knots or plectonemes, is the primary

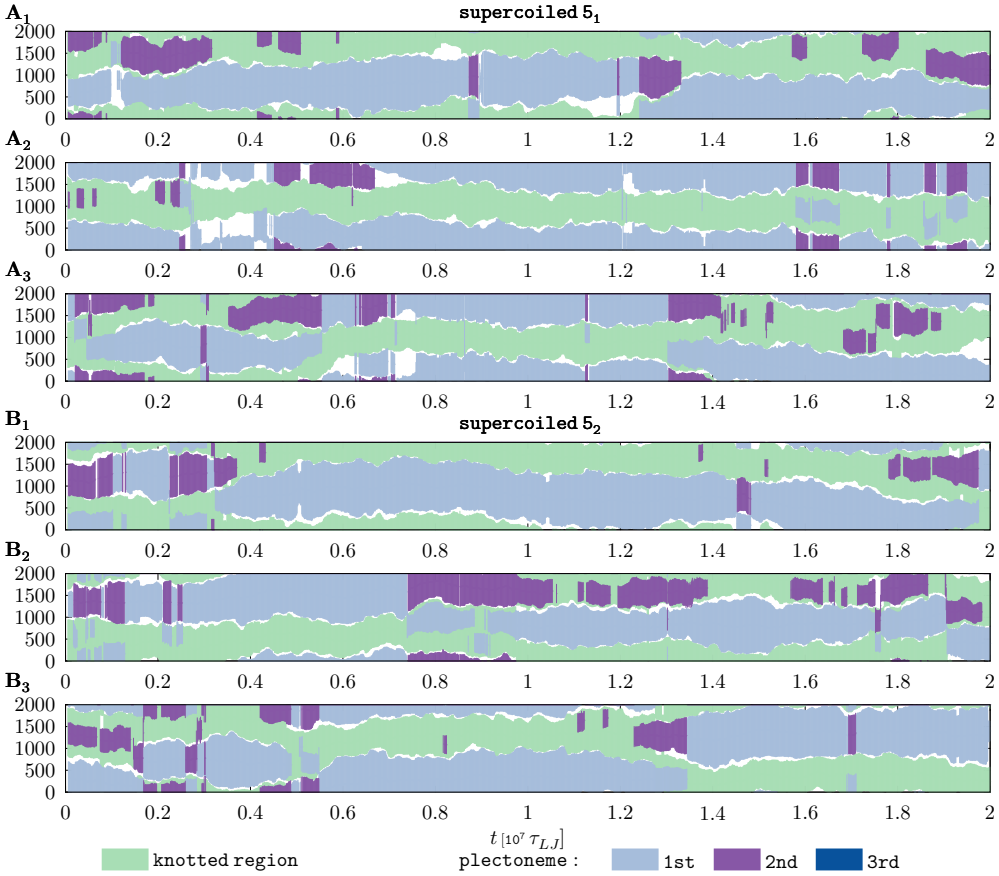


FIGURE 5.18: Kymographs showing the contour motion of the knotted and plectonemically-wound regions. Data are for different trajectories of 2kbp-long supercoiled rings with 5_1 or 5_2-knots, as indicated. Notice the persistence of the boundaries between the knotted and the main plectonemically-wound region.

cause for the slow evolution of these boundaries.

5.3.4 Persistent interlockings

As mentioned in the previous paragraph the properties that emerges when we consider torsionally-relaxed knotted rings and supercoiled unknotted one, are in stark contrast with those of panel A, where knotting and supercoiling are both present. In fact one observes that neither the boundaries of the knotted region (without supercoiling) nor those of plectonemes (without knotting) are inherently persistent. In fact, in both cases the boundaries vary on the same relatively fast timescales of the metric relaxation, i.e. $\sim 10^6 \tau_{LJ}$.

We thus conclude that it is precisely the synergistic action of complex topology and supercoiling that is responsible for the locked boundaries, and the latter disappears when either of them is missing.

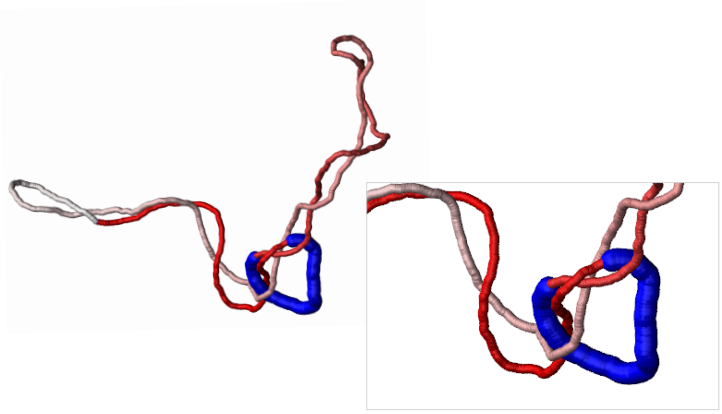


FIGURE 5.19: Example of interlocking. The ring is knotted and supercoiled. The insert shows how the strands are arranged in the so called persistent interlockings.

To clarify the mechanism underpinning this effect we inspected in detail the dynamical evolution of the rings. We thus established that the persistent boundaries correspond to specific points of tight and complex self-contacts of the knotted region, a typical example is given in Fig. 5.19.

A large number of self-contacting *loci* are clearly introduced by supercoiling in any DNA ring, regardless of its topological state. In knotted rings, due to the tightening of the intrinsic essential crossings, these *loci* typically involve several clasped or hooked double strands, as highlighted in the snapshots of Fig. 5.17A. We observed that, similarly to what happens during the pore translocation of knotted filaments [38, 129, 130, 138], the topological friction at these points is so high that the DNA strands are locally pinned while other parts of the chain can still reconfigure.

This effect accounts for the observed separation of timescales between the metric relaxation time and the contour motion of the clasped points, at the boundary of the knotted region.

Incidentally, we note that the relevance of these persistent regions of self-contacts reinforces *a posteriori* the necessity to use models, such as oxDNA, where the spatial description is sufficiently fine to capture the internal friction that develops when two tightly interacting DNA strands slide against each other. For this reason, we surmise that the observed sliding hindrance of contacting strands would be even higher in atomistic or more fine-grained models at this same high salt conditions. At the same time, the friction between DNA strands could be relieved by increasing their electrostatic repulsion with a lower salt concentration. We believe these would be

worthwhile points to address in future modelling studies and possibly also experimentally, e.g. using setups akin to those of ref. [139].

Importantly, not all intrinsic (or essential) crossings of the knots create persistent interlockings, but only a subset of them. This is visible in the snapshots of Fig. 5.16A where one notes that the two boundaries of the five-crossing knot are pinned by their high local physical entanglement and yet, the DNA strands can slide, despite the several points of pairwise contacts. It is precisely this internal sliding that creates the possibility for plectonemes to form transiently, but repeatedly, within the knotted region. As a matter of fact, the persistent interlocking appears in either of the two qualitatively different conformers populated by supercoiled knotted rings (i.e. those with local or non-local knots in Fig. 5.9.

In this regard, it is relevant to recall the seminal work of Liu *et al.* [34, 117], who pointed out that hooked DNA juxtapositions are an ideal target substrate of topoII enzymes because local strand passages at these point generally produces a simpler topology. The kinetic persistence of these multi-strand interlocking adds a novel temporal dimension to other, more thermodynamical effects of supercoiling, such as knot localization, that are credited to favour the local, yet globally-disentangling action, of topo II. The present results, in fact, complement the insight from earlier thermodynamic sampling [124, 125], by showing that once hooked or multiply-clasped juxtapositions are formed, they are long-lived. This, we speculate, is key for making such forms of local entanglement persistent enough to be recognised and processed by topo II enzymes.

5.4 Conclusions

To summarize, we used molecular dynamics simulations and the oxDNA mesoscopic model to study the effect of complex, five-crossing knots, on the conformational and kinetic properties of 2kbp-long plasmids with the typical 5% negative supercoiling found in bacterial plasmids. We particularly focussed on whether and how complex topologies, with their numerous points of high curvature and self-contact, can alter the branchedness of supercoiled plasmids and the dynamical evolution at long timescales.

On both accounts, we found that the interplay of knotting and supercoiling has major consequences that would have been difficult to anticipate *a priori*.

For the structural properties, we found that the conformational ensemble explored during the spontaneous dynamical evolution is largely dominated by two qualitatively-distinct states. They differ both by knot size and degree of branching. This fact, noteworthy *per se*, is accompanied by two intriguing kinetic effects. The first is that spontaneous fluctuations between these two states occur on timescales that are comparable to metric relaxation times of unknotted rings. The second is that certain boundaries separating the knotted and plectonemically-wound regions are very long-lived: they remain persistent throughout the 1.5ms-long simulated trajectories, and hence vary over much slower timescales. The relaxation times are much longer - arguably by at least an order of magnitude - than the metric relaxation times which is about 0.3ms. This complex phenomenology is shown to arise exclusively from the cooperative action of supercoiling and topological constraints; removing either of the two suffices to remove the persistent boundaries. The latter are shown to occur in correspondence of *loci* where multiple strands become interlocked. The interlockings have a local geometry that is analogous to the so-called hooked juxtapositions [34,117,118], argued to be ideal local targets for the topoisomerases' knot simplifying action. We accordingly surmise that their long-lived character, besides their structural features, could also be instrumental to favour their recognition by topoisomerases. We believe this would be a noteworthy problem to address in future studies, for instance using mesoscopic models incorporating the interaction of DNA and proteins, which would be essential for a realistic description of DNA organization and processing *in vivo*. In addition, we expect that the incidence and long-lived character of multi-strand interlockings could also be probed with advanced single-molecule manipulation techniques such as pore translocation that, having been successfully applied DNA rings with either knots or supercoiling [38, 60, 127, 140–143], is ideally suited to address their concurrent effects. In addition to previously established conformational features such as hooked juxtapositions or tight knots [34, 117, 118, 123], the long-lived character of these regions could aid the recognition, and hence knot simplification, by topoisomerases.

Knotting probability of intracellular DNA

6.1 Introduction

In Chapter 5 we discussed how negative supercoiling, combined with complex topologies, affects static and dynamical properties of DNA rings. Knotting and supercoiling arise from the incessant action of various DNA-processing enzymes, particularly RNA polymerases plus Topo I and Topo II enzymes.

As noted in the previous chapter, the DNA strands separation operated by the RNA polymerase causes the accumulation of positive supercoiling, $(+)S$, ahead of the polymerase and negative supercoiling, $(-)S$, behind, see Fig. 6.1, which are next relaxed by topoisomerase enzymes to prevent the stalling of the transcription process. Note that the two oppositely-directed types of supercoiling are produced at the same rate, which is typically of 10 turns/s, but are relaxed with different efficiency by Topo I and Topo II enzymes. In bacterial DNA, it is this interplay of creation and differential relaxation of $(+)S$ and $(-)S$ that leads to the overall negative level of supercoiling at steady state. In eukayotic DNA, the chromatin architecture itself can strongly affects the way that supercoiling is introduced, relaxed or propagated. As matter of fact the DNA wrapping around histones can act as a buffer for the torsional stress, and adds a further level of complexity in supercoiling homeostasis.

Besides, topoisomerases act differently in different chromatin regions: in particular, Topo II preferentially relaxes the denser DNA regions organised by histones, whereas Topo I acts preferentially on the proteins-free ones.

It is important to stress that both types of topoisomerases are involved in the transcription process but the latter can be stalled only by inhibiting Topo II, not Topo I [144]. Therefore, the entanglement features capable of hindering transcription must

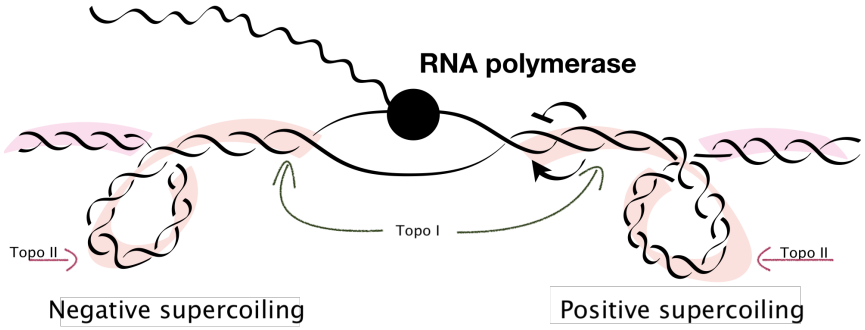


FIGURE 6.1: Illustration of how the strands-separation accumulates positive and negative supercoiling.

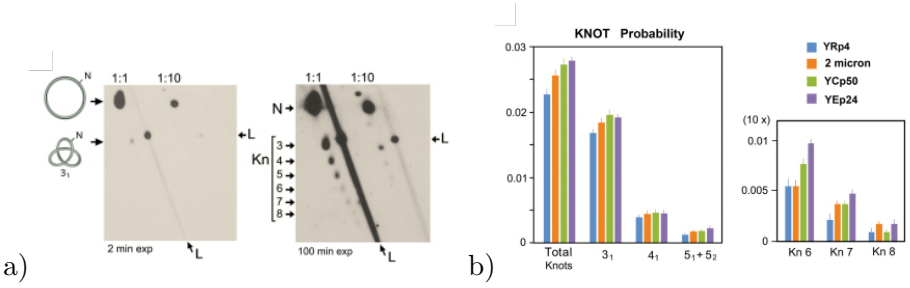


FIGURE 6.2: Yeast minichromosomes DNA. Image and caption are from ref. [24]. a) *in vivo DNA sample of YRp4 was enzymatically nicked and loaded in a 2D gel for analysis of DNA knotting. Gel lanes show 1:1 and 1:10 dilutions of the sample. Left and right panels show, respectively, short (2 min) and long exposures (100 min) of the gel-blot.* b) *Total DNA knot probability and of knots* 3_1, 4_1, $5_1 + 5_2$, *and knot species of 6 to 8 crossings (Kn6, Kn7, Kn8) in the indicated yeast minichromosomes. The plots show the mean and SD of three experiments.* [24]

be strongly related to the different properties of the substrates processed by Topo I and Topo II, such as the number of strand passages, see chapter 5.1.1.

This necessary role of the Topo II leads to the question of assess the extent to which DNA knots, which cannot be relaxed by Topo I, are associated to the positive supercoiling that accumulated ahead of the RNA polymerase.

This question is reinforced by the recent seminal study by Valdés *et al.* [24] that has shown, for the first time, that eukaryotic DNA can be knotted *in vivo*. They used minichromosomes YRp4 of 4.4 kb, which were nicked and analysed via 2D gel electrophoresis, see Fig. 6.2a. The gel bands indicate that 3_1 knots are the most probable, but not the sole, topology and that the overall knotting probability of these minichromosomes is $\sim 0.02 - 0.03$. The later value clearly sets the baseline for DNA knotting in minichromosomes *in vivo*.

They also hypothesized that, *in vivo*, the observed knotting probabilities are explicable by a possible hierarchical structural organization of chromatin. The interpretation

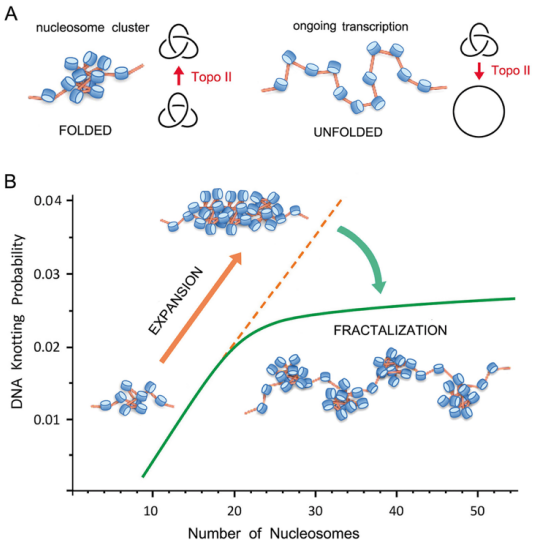


FIGURE 6.3: Image and caption adapted from ref. [24]. *Model of chromatin architecture inferred from DNA knotting probability. (A) Intricate folding of nucleosome arrays favours topo II-mediated knotting of intracellular DNA. Knotted fractions are reduced when nucleosomal clusters unfold during DNA transcription. (B) Uninterrupted expansion of nucleosomal fibers would produce proportional scaling of DNA knot formation (orange dashed line). Fractalization of the chromatin architecture minimizes instead the potentially harmful scaling of DNA entanglements (green line). The kPCHR data supports a fractal model, in which the beads on a string architecture of the 10 nm nucleosomal fiber reiterates in its next level of organization by forming clusters of about 20 nucleosomes.*

that Valdés *et al.* provide for these results is that intramolecular DNA segments are much more juxtaposed due to the high flexibility of nucleosomal fibers compared to naked DNA. Notice that the flexibility arises thanks to the free segments between nucleosomes, which are called linkers. They behave like rigid sticks completely free to bend. The DNA knotting probability shows to be proportional to the number of nucleosomes, at least up to ~ 20 nucleosomes, see Fig. 6.3. Valdés *et al.* show that for high flexibility of nucleosomal fibers the DNA knotting grows with length however some mechanism minimizes the scaling of DNA knot formation throughout intracellular chromatin. They also postulate that the controlling mechanism is related to architecture of chromatin, and in particular to its compactness.

As matter of fact, *in vitro* studies have shown that topo II produces abundant and complex knots when there are several DNA strands that are brought close together due to DNA condensation [145, 146]; and computer simulations [21, 147, 148] of polymer chains have largely investigated the effect of DNA compaction on knot abundance and complexity. Moreover studies [149, 150]have also shown that introducing $(+)S$ can cause the rapid compatification of nucleosomal fibers.

Overall, these results point at a connection between the knotting observed in eukaryotic DNA *in vivo* and its compactness and supercoiling properties.

Here I will report on the use of a minimalistic model of minichromosome as a first step to clarify the interplay of these three elements. The material presented here is mostly based on the study published in ref. [151], which resulted from a collaboration

with Prof. Roca's experimental lab at the University of Barcelona.

6.2 Model and Methods

Model We represent YRp4 minichromosomes with a strings and beads model. The YRp4 DNA is $\sim$ 4.4kb-long and is wrapped around 25 nucleosomes. The latter are modelled by spherical beads of diameter D, and the DNA duplex linkers between them are described by infinitely-thin straight segments of length $L + D$ connecting the centers of consecutive beads, see Fig. 6.4. Note that free portion of a segment, describing the DNA linker has length L. For YRp4 minichromosome, D is about 10nm, and the length of the linkers can vary between 10 to 20nm. Here, for simplicity, we assume all linkers to be of the exactly the same length.

We tune the ratio D/L in order to reproduce the baseline knotting probability of $\sim$ 0.02. The beads are impenetrable and the excluded volume effect is introduced by assigning infinite energy to configurations with overlapping beads and zero energy otherwise.

The model has the following limitations: is not possible to properly define the torsion and perform a systemic study on it. It is possible, instead, to use the confinement as observable, which enable to investigate the effect on knotting and writhe. We recall that the latter, is a measurement of the amount of winding of the axis of a double-strands and it is described in Chapter 2.4.

Another important point to stress is that, for simplicity, we will assume for simplicity to be in equilibrium, even thought the processes results from the concurrent dynamic (out-of-equilibrium) action of RNA polymerase and topoisomerases (and hence do not account for effects related to their concentration in solution, DNA binding affinities, efficiency etc.)

Monte Carlo A Metropolis Monte Carlo scheme based on unrestricted crankshaft moves is used to sample the conformational space, see Chapter 1.3. The moves allow the linkers to cross so that the sampled space corresponded to torsionally relaxed and topology unrestricted minichromosomes. For different combinations of the D/L ratio in the $[0 : 1]$ range, we collect no less than 10^5 uncorrelated conformations: they are

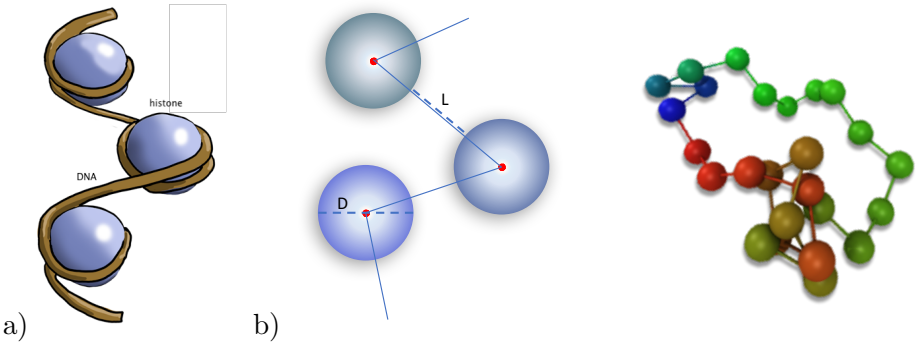


FIGURE 6.4: a) Representation DNA wrapped in 1.67 left-handed superhelical turn ($\sim$ 146 bp) around an histone. Image adapted from Shmoop Editorial Team. (2008, November 11). b) Model of minichromosomes of 4.4kbp represented by ring of beads and linkers.

picked at time intervals larger than the autocorrelation time of the radius of gyration radius, R_g.

Methods We consider different combinations of the D/L ratio in the $[0 \div 1]$ range to tune the optimal ratio that reproduce the value of basal knotting probability observed in ref. [24]. For each of them we collected over 10^6 independent conformations (i.e. picked at time intervals larger than the autocorrelation time of the radius of gyration radius, R_g) which were topologically-profiled. This has been done by comparing the Dowker code of their 2D projections against tabulated values using the Knotscape software [152]. The $\sim$ 2% knotting probability observed experimentally for torsionally relaxed YRp4 rings was recovered for $D/L \sim 0.47$, see Fig. 6.5, which is plausible based on the aforementioned nominal histone diameter and linker lengths in YRp4 (nucleosomes diameter of 10 and linkers of about $10 - 20$).

6.3 Results

Computational results Firstly, we address the role of compactification on the knotting probability at fixed value of $D/L = 0.47$. We collect an extensive set of $\sim 4 * 10^6$ conformations. The packing is introduced by resampling *a posteriori* the conformations to retain only those with gyration radius, normalised to the unconstrained average value R_g^0, smaller than a maximum threshold value, max R_g/R_g^0. By varying the latter in the $[0.55 \div 1.8]$ we obtain equilibrated ensembles of conformers with different upper bound for their diameter.

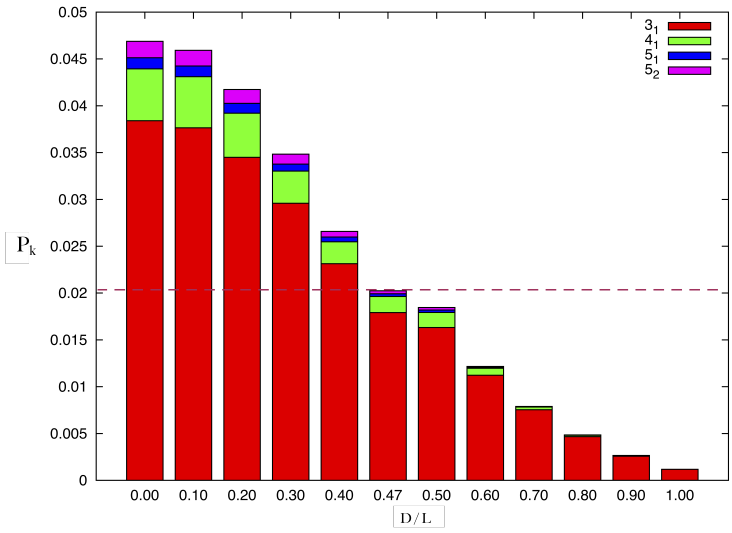


FIGURE 6.5: Knotting probability as function of the ratio D/L. The dotted line is the target experimental knotting probability.

Fig. 6.6B illustrates the effect of increasing compactification, the knotting probability increases, reaching $\sim 50\%$ for strong packing, i.e. when the R_g is $< 60\%$ of the average unconstrained gyration radius R^0, which correspond to a 5-fold volume compaction. Notice that knotting, in the strongest compactification achieved, is enhanced ~ 20 times, compared with the baseline of 0.02 for unconstrained rings.

This boost varies with the complexity of the knot, as clarified by the profiles for the enhancement of $3_1, 4_1$ and $5_1 + 5_2$ knot types. These are enhanced by a factor of 10 ,25 and 60 times, respectively.

Secondly, we tackle the question of the interplay between the compaction level and the DNA supercoiling. Because torsion cannot be inserted in this minimalistic minichromosome model, it is not possible to study directly the effects of supercoiling. As a proxy, we instead consider the average unsigned writhe, $|Wr|$ of the rings. A heuristic rationale for this is that DNA is characterized by large torsional rigidity persistence length $\sim$ 80nm, and hence the introduction of torsion is mostly converted in writhe rather than DNA twist [153].

Experimental results Here, I briefly recall the main experimental findings obtained in Roca's Lab using minichromosomes YRp4 and *E. Coli TopA*; they accumulate positive supercoiling by inactivation of topo II while *TopA* is kept active, thus can relax $(-)S$. Instead, the accumulation of $(-)S$ is obtained by of thermal inactivation

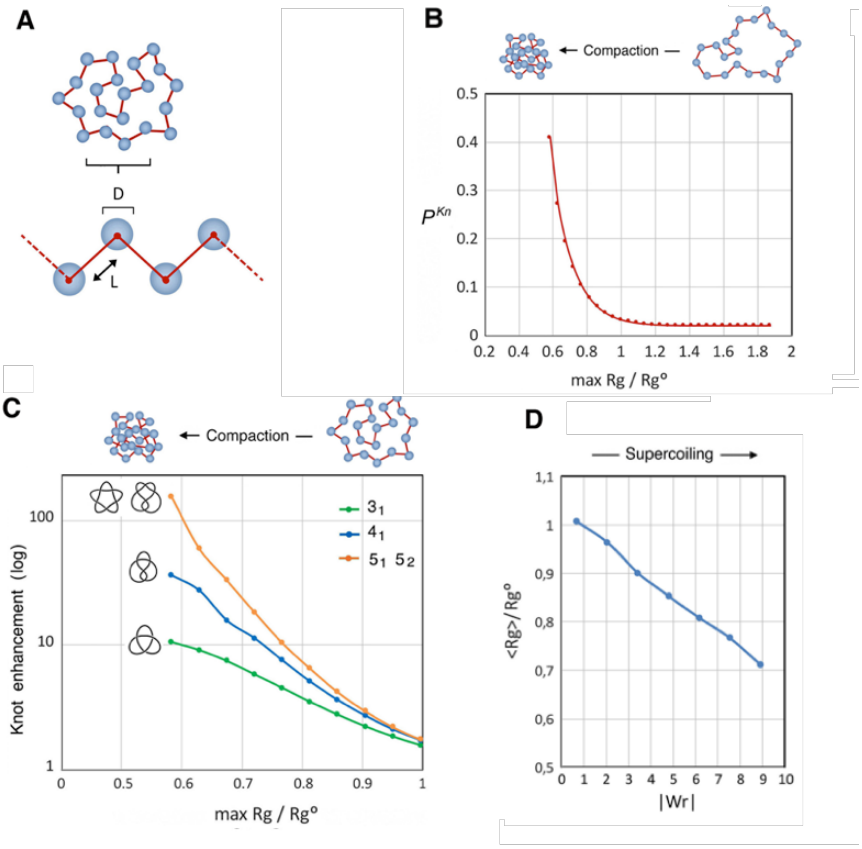


FIGURE 6.6: (A) Model that simulates nucleosomal fibers. (B) Effect of compaction of knot probability. P_{Kn} values ($N = 25, D/L = 0.47$) are plotted as a function of the gyration radius (R_g). Each point computes the P_{Kn} of those configurations of R_g below a cutoff value ($\max R_g$) relative to the average gyration radius of the entire distribution of conformers (R_g^0). (C) Enhancement of individual knot populations by the effect of compaction. (D) Reduction of $\langle R_g \rangle / R_g^0$ as a function of the absolute writhe ($|Wr|$). Image adapted from ref. [151].

of topo II.

2D-gel electrophoresis is then used to quantify L_k and hence the level of supercoiling. Finally, after properly nicking the YRp4 minichromosomes DNA and using a different 2D-gel electrophoresis setup, they examine the incidence of knots and their complexity.

The results are shown in Fig. 6.7 panels C and D. The 2D-gel electrophoresis clearly shows that is DNA knotting increases strongly with the accumulation of $(+)S$ (Fig. 6.7D), instead, in the accumulation of $(-)S$ (Fig. 6.7C) there are no appreciable differences in the gel runned before and after the accumulation of $(-)S$. As a result of the increase of $(+)S$, the enhancement of P^{kn} is $\sim 10 - fold$, increasing from ~ 0.02 to ~ 0.2, see Fig. 6.7F.

The emergence of knots is directly due to the residual activity of topo II when there is an unbalance between the two directions of supercoiling. The experimental data

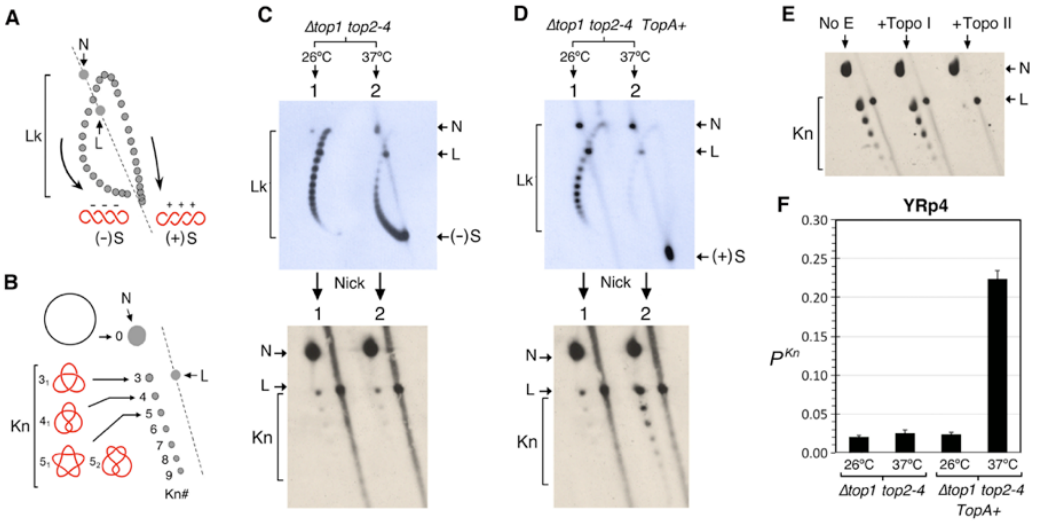


FIGURE 6.7: DNA knotting probability during (+) and () supercoiling of chromatin. (A) *Lk* distribution of DNA topoisomers in a 2D-gel electrophoresis; N, nicked DNA circles; L, linear DNA. (B) Relative position of unknotted (N) and knotted nicked DNA circles (Kn) in a 2D-gel electrophoresis. (C and D) DNA topology of YRp4 in top1 top24 (C) and in top1 top24 TopA+ (D) cells. (E) Incubation of the nicked DNA sample of (+)S YRp4 (no E) with topo I and topo II activities in vitro. (F) DNA knotting probability (PKn) of YRp4 in the four conditions analyzed in panels (C) and (D). Image from ref. [151].

collected by Roca's Lab indicate that transient DNA knots, which are present in intracellular chromatin, are strongly enhanced by positive supercoiling.

In Fig. 6.8B, is possible to see how the knotting probability starts increasing when the $(+)S$ is accumulating, and it affects half of the supercoilied population after $20 - 40min$ (green bars), see ref. [151] for more details.

The boost of knotting also affects the knot spectrum. Indeed, the enhancement of individual knot population increases accordingly to the complexity of the knot see Fig. 6.8D: the trefoil knot which is the most abundant, increases its incidence by ~ 12 times compared to the torsionally-relaxed case, the enhancement is 25-fold for 4_1 knots and 60-fold for the combined population of 5_1 and 5_2 knots .

6.4 Discussion

The YRp4 minichromosome model provide a first valuable insight into the structural properties that are likely responsible for the enhancement of knots observed experimentally during transcription.

First, we found that the 20-fold knot enhancement can be accounted for by a substantial compactification of the minichromosome, namely a 5-fold reduction of the

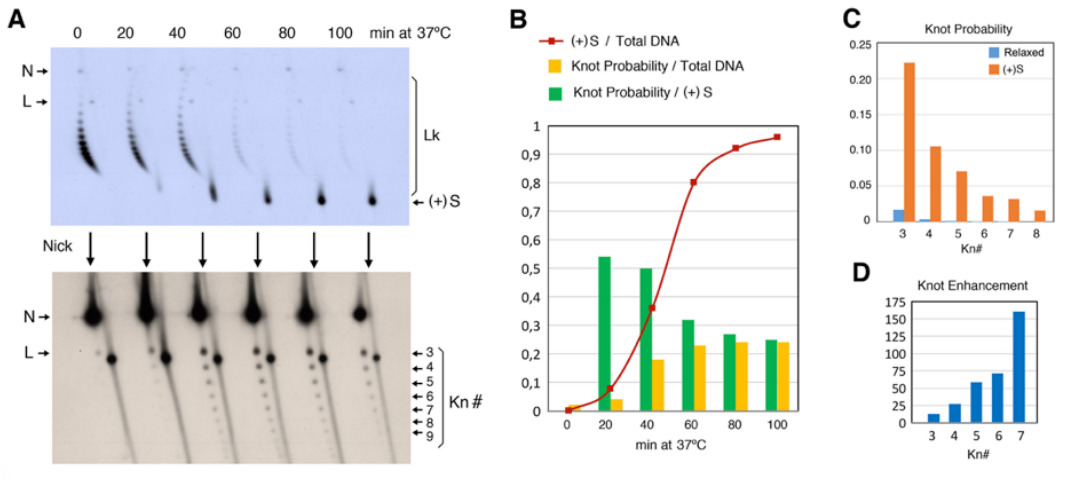


FIGURE 6.8: (A) DNA supercoiling and knotting of YRp4 in top1 top24 TopA+ cells sampled at different time points (min) after shifting the cultures to 37C. (B) Comparison of the accumulation rate of (+)S (red), PKn values relative to total DNA (yellow) and PKn values relative to the fraction of (+)S DNA (green). (C) Probability of individual knot populations in cells sampled at 0 min (relaxed chromatin, blue) and after shifting them to 37C for 100 min ((+)S chromatin, orange). (D) Enhancement of individual knot populations shown in panel (C) upon accumulation of (+)S. Image and caption from ref. [151].

occupied volume.

Second, the packing effect is compatible with the accumulation of supercoiling. In fact, we observe that a progressive compactification of the model minichromosome reflects in a systematic increase of the unsigned writhe, see Fig. 6.6D.

Finally, the properties of the model aptly complement the experimental ones for having insight on the simplifying action of topoisomerases. More precisely, Roca's Lab was able to detect the transient knotting probability that arises during the accumulation of supercoiling and the action of topoisomerases.

As already mentioned in Chapter 5, the DNA disentangling action of topoisomerase has long been investigated because it poses the conundrum of how the necessarily local action of these enzymes can be reconciled with the capability to remove global forms of entanglement. Over the years several appealing explanations have been suggested, including active biological mechanisms that promote the tightening and hence localization of knots, [19,123]the temporal persistence of essential crossings [126], and the fact that knots are accompanied by characteristic local geometric features, such as hooked DNA juxtapositions [63].

The transient enhancement of knots observed in Roca's lab points, instead, to an apparently contradicting and opposite effect, namely that topoisomerases introduce

rather than remove knots. This opposite effect can, however, be rationalised too with the same framework based on hooked juxtapositions that can account for topoisomerase disentangling action. The apparent contradiction is, in fact, solved by noting that the compactification caused by the transient accumulation of supercoiling will inevitably cause the juxtaposition of several DNA strands.

The strand-passage action of topoisomerases on these supercoiling induced juxtapositions on an initially *unknotted* DNA, will lead to introducing knots. Eventually, the torsional relaxation that accompanies these strand-passages will lower the incidence of supercoiling-induced juxtapositions in favour of the juxtapositions that were introduced by knotting. At this stage, the action of topoisomerases will be directed on topologically-induced juxtapositions, thereby leading to an overall simplification of DNA topology. This mechanism can account for the experimentally transient boost of knotting that accompanies transcriptional supercoiling.

One interesting aspect to study in the future would be how the efficiency rates of topoisomerases, controlled by the chromatin architecture, affect the resulting topo-mediated knotting. At the same time, the presence of the entanglement itself carries question on possible functional aspect of the knotting that can be studied too.

Concluding remarks

The common theme through the chapters of this thesis was how intrinsic physical properties of polymers, such as their bending or torsional ones, affect and constrain the emergence of knots and their dynamics too.

The first part of the thesis was focussed on the interplay between knotting propensity and bending rigidity of equiibrated rings polymers. We started by using a general model of semi-flexible self-avoiding rings of up to 1000 beads and systematically profiled the knotting probability versus bending rigidity.
We found a surprising result: the equilibrium incidence of knots has a strongly non-monotonic dependence on bending, with a maximum at intermediate flexural rigidities. We next provided a quantitative framework, based on the balance of bending energy and configurational entropy, that allowed for rationalizing this counter-intuitive effect.
We next extended the investigation to rings of much larger number of beads, via an heuristic model mapping between our semiflexible rings of beads and self-avoiding rings of cylinder. The latter model was chosen because earlier studies had provide for it approximate analytical expressions for their knotting probability as a function of the number and thickness of cylinders. By these mapping, we not only confirmed the unimodal knotting profile for chains of 1,000 beads, but further found that chains of $> 20,000$ beads are expected to feature a bi-modal profile. We believe it would be most interesting to direct future efforts to confirm this transition from uni- to bi-modality using advanced sampling techniques for very long polymer rings.

The second part of the thesis focused on the interplay of DNA knots and supercoiling which are typically simultaneously present *in vivo*. We first studied this interplay by using oxDNA, an accurate mesoscopic DNA model and using it to study

rings of thousands of base pairs tied in complex knots and with or without negative supercoiling (as appropriate for bacterial plasmids). By monitoring the dynamics of the DNA rings we found that the simultaneous presence of knots and supercoiling, and only their simultaneous presence, leads to a dramatic slowing down of the system reconfiguration dynamics. In particular, the essential tangles in the knotted region acquire a very long-lived character that, we speculate, could aid their recognition and simplification by topoisomerase.

Finally, motivated by the recent experimental breakthrough that detected knots in eukaryotic DNA, in the last chapter, we investigated the relationship between the compactness, writhe and knotting probability. The model was tuned to capture some of the salient properties of yeast minichromosomes, which were shown experimentally to become transiently highly knotted during transcription (work done by our collaborator group of Prof. Roca in Barcelona). The model, with its designed simplicity helped clarify the connection between chromatin compactification, knotting and supercoiling, and further suggested possible mechanisms for the transient increase of knots promoted by topoisomerases. Again, we believe this would be an worthy topic to study further with more accurate models, possibly incorporating the action of topoisomerases too.

Each project presented has provided new information about the interplay of knotting probability and geometrical properties of polymers, helping in the clarification of the spontaneous emergence of knots as well as the behaviour of the proprieties related. The results presented in this manuscript underscore the fact that suitable coarse-grained models, such as rings of beads or oxDNA, and simulation techniques can be used to elucidate the complex interplay of structure, function and entanglement of DNA filaments *in vivo*. They give an unique insight on the mechanisms whereby the knots form or are suppressed, inaccessible by experimental techniques.

Finally, as we already mentioned, all the obtained results offer a valuable starting point for future extensions towards theoretical and computational characterizations of DNA with even more detail and realistic properties.

Acknowledgements

This thesis, in its form and contents, owes much to the dedicated and competent effort of my supervisor, Cristian. I am extremely grateful to him for his mentorship and teaching with which he guided me during these four years.
All the professors that I met have contributed to my growth, I want to thank Enzo, for his valuable collaboration, and the professors of the SBP group for their teaching.
In four years the members of SBP group have changed a lot but I've been lucky enough to know many of them and share a piece of my journey with their ones. I want to thank my colleagues Stefano and Giulia, who have been extremely supportive and caring in the every day life in Sissa; Mattia and Antonio, for all the help they gave and give me even across the seas. I also want to thank Andrea Giulia and Francesca with whom I started and shared this journey, and all the ex-members that are spread around the world, among them I thank Andrea for all he did for me.
In these years in Trieste I encountered so many friends that have to be thanked, we shared fun moments, conversations, dances, scooter trips, walks and barbecues, but in particular in this last year some of them have been a fundamental presence, they have supported me, enriched and yes, also feed me: Martina, Irene, Juraj, Daniele and Serafina, the ones who do not live in Trieste, and yet on which I can always count, Anna, Davide, Andrea N., Raffaele G., Guido, Christhian, Riccardo, Pamela, Silvia and Valentina. Thank you all.
To my wonderful family, I am lucky to have all of you in my life, you are my heroes.
Finally, my gratitude goes to people like Alexandra Elbakyan, idealistic, who stands.

Bibliography

[1] Micheletti, C., Di Stefano, M. & Orland, H. Absence of knots in known rna structures. *Proceedings of the National Academy of Sciences* **112**, 2052–2057 (2015).

[2] Mallam, A. L., Rogers, J. M. & Jackson, S. E. Experimental detection of knotted conformations in denatured proteins. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 8189–8194 (2010).

[3] Mallam, A. L. & Jackson, S. E. Knot formation in newly translated proteins is spon- taneous and accelerated by chaperonins. *Nat Chem Biol* **8**, 147–153 (2012).

[4] Virnau, P., Mallam, A. & Jackson, S. Structures and folding pathways of topologically knotted proteins. *J. Phys. Condens. Matter.* **23**, 033101 (2011).

[5] Jackson, S. E., Suma, A. & Micheletti, C. How to fold intricately: using theory and experiments to unravel the properties of knotted proteins. *Current Opinion in Structural Biology* **42**, 6 – 14 (2017).

[6] Mansfield, M. L. Are there knots in proteins? *Nature Structural Biology* **1**, 213–214 (1994).

[7] King, N. P., Yeates, E. O. & Yeates, T. O. Identification of rare slipknots in proteins and their implications for stability and folding. *J Mol Biol* **373**, 153–166 (2007).

[8] Faísca, P. F. N., Travasso, R. D. M., Charters, T., Nunes, A. & Cieplak, M. The folding of knotted proteins: insights from lattice simulations. *Physical Biology* **7**, 016009 (2010).

[9] Faísca, P. F. N. Knotted proteins: A tangled tale of structural biology. *Computational and structural biotechnology journal* **13**, 459–468 (2015).

[10] Soler, M. A. & Faísca, P. F. N. Effects of knots on protein folding properties. *PLOS ONE* **8**, 1–10 (2013).

[11] Shishido, K., Komiyama, N. & Ikawa, S. Increased production of a knotted form of plasmid pBR322 DNA in Escherichia coli DNA topoisomerase mutants. *J. Mol. Biol.* **195**, 215–218 (1987).

[12] Dean, F. B., Stasiak, A., Coller, T. & Cozzarelli, N. Duplex dna knots produced by escherichia coli topoisomerase i. structure and requirements for formation. *J Biol Chem.* **260**, 4975–4983 (1985).

[13] Deibler, R. W., Mann, J. K., Sumners, D. W. L. & Zechiedrich, L. Hin-mediated DNA knotting and recombining promote replicon dysfunction and mutation. *BMC Mol Biol* **8**, 44–14 (2007).

[14] Liu, L. F., Depew, R. E. & Wang, J. C. Knotted single-stranded DNA rings: A novel topological isomer of circular single-stranded DNA formed by treatment with Escherichia coli ω protein. *J. of Mol. Biol.* **106**, 439–452 (1976).

[15] Liu, L. F., Perkocha, L., Calendar, R. & Wang, J. C. Knotted dna from bacteriophage capsids. *Proceedings of the National Academy of Sciences* **78**, 5498–5502 (1981).

[16] Schvartzman, J. B., Martinez-Robles, M. L., Hernandez, P. & Krimer, D. B. Plasmid dna replication and topology as visualized by two-dimensional agarose gel electrophoresis. *Plasmid* **63**, 1–10 (2010).

[17] Sogo, J. M. *et al.* Formation of knots in partially replicated dna molecules. *J. Mol. Biol.* **286**, 637–643 (1999).

[18] Valle, F., Favre, M., De Los Rios, A., P. Rosa & Dietler, G. Scaling exponents and probability distribution of dna end-to-end distance. *Phys. Rev. Lett.* **95**, 158105 (2005).

[19] Ercolini, E. *et al.* Fractal dimension and localization of dna knots. *Phys. Rev. Lett.* **98**, 058102 (2007).

[20] Meluzzi, D., Smith, D. E. & Arya, G. Biophysics of knotting. *Annu Rev Biophys* **39**, 349–366 (2010).

[21] Marenduzzo, D., Micheletti, C. & Orlandini, E. Biopolymer organization upon confinement. *J. Phys. Condens. Matter.* **22**, 283102 (2010).

[22] Shaw, S. Y. & Wang, J. C. Knotting of a dna chain during ring closure. *Science* **260**, 533–536 (1993).

[23] Arsuaga, J., Vázquez, M., Trigueros, S., Sumners, D. W. & Roca, J. Knotting probability of dna molecules confined in restricted volumes. *Proc. Natl. Acad. Sci. USA* **99**, 5373–5377 (2002).

[24] Valds, A., Segura, J., Dyson, S., Martnez-Garca, B. & Roca, J. DNA knots occur in intracellular chromatin. *Nucleic Acids Research* **46**, 650–660 (2017).

[25] Arsuaga, J. *et al.* Dna knots reveal a chiral organization of dna in phage capsids. *Proc. Natl. Acad. Sci. USA* **102**, 9165–9169 (2005).

[26] Matthews, R., Louis, A. A. & Yeomans, J. M. Knot-controlled ejection of a polymer from a virus capsid. *Phys. Rev. Lett.* **102**, 088101 (2009).

[27] Marenduzzo, D., Micheletti, C., Orlandini, E. & Sumners, D. W. Topological friction strongly affects viral dna ejection. *Proc. Natl. Acad. Sci. USA* **110**, 20081–20086 (2013).

[28] S. A. Wassermann, N. R. C. Biochemical topology: applications to dna recombination and replication. *Science* **232**, 951–960 (1986).

[29] Liu, Z., Deibler, R. W., Chan, H. S. & Zechiedrich, L. The why and how of DNA unlinking. *Nucleic Acids Research* **37**, 661–671 (2009).

[30] Liu, L. F., Liu, C.-C. & Alberts, B. M. Type ii dna topoisomerases: Enzymes that can unknot a topologically knotted dna molecule via a reversible double-strand break. *Cell* **19**, 697–707 (1980).

[31] Wang, J. C. Dna topoisomerases. *Annual Review of Biochemistry* **65**, 635–692 (1996). PMID: 8811192.

[32] Sumners, D. W. Untangling dna. *Math. Intelligencer* **12**, 71–80 (1990).

[33] Sumners, D. W. Knot theory and dna. *New Scientific Applications of Geometry and Topology: Proceedings of Symposia in Applied Mathematics* **45**, 39–72 (1992).

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

[35] Viovy, J.-L. Electrophoresis of dna and other polyelectrolytes: Physical mechanisms. *Rev. Mod. Phys.* **72**, 813–872 (2000).

[36] Trigueros, S., Arsuaga, J., Vázquez, M., Sumners, D. W. & Roca, J. Novel display of knotted dna molecules by two-dimensional gel electrophoresis. *Nucleic Acids Res* **29**, e67 (2001).

[37] Olavarrieta, L. *et al.* Supercoiling, knotting and replication fork reversal in partially replicated plasmids. *Nucleic Acids Research* **30**, 656–666 (2002).

[38] Plesa, C. *et al.* Direct observation of dna knots using a solid-state nanopore. *Nat. nanotechnol.* **11**, 1093–1097 (2016).

[39] Delbruck, M. Knotting problems in biology. *Mathematical Problems in the Biological Sciences, edited by R. E. Bellman, Proceedings of Symposia in Applied Mathematics,American Mathematical Society, Providence, Rhode Island)* **14**, 55–63 (1962).

[40] Micheletti, C. & Orlandini, E. Numerical study of linear and circular model dna chains confined in a slit: Metric and topological properties. *Macromolecules* **45**, 2113–2121 (2012).

[41] Drube, F., Alim, K., Witz, G., Dietler, G. & Frey, E. Excluded volume effects on semiflexible ring polymers. *Nano Letters* **10**, 1445–1449 (2010). PMID: 20235507.

[42] des Cloizeaux, J. Ring polymers in solution : topological effects. *Journal de Physique Letters* **42**, L433–L436 (1981).

[43] Rybenkov, V. V., Cozzarelli, N. R. & Vologodskii, A. V. Probability of dna knotting and the effective diameter of the dna double helix. *Proc. Natl. Acad. Sci. USA* **90**, 5307–5311 (1993).

[44] Katritch, V. *et al.* Geometry and physics of knots. *Nature* **384**, 142–145 (1996).

[45] Rieger, F. C. & Virnau, P. A monte carlo study of knots in long double-stranded dna chains. *PLOS Computational Biology* **12**, 1–11 (2016).

[46] Witz, G. & Stasiak, A. Dna supercoiling and its role in dna decatenation and unknotting. *Nucleic Acids Res.* **38**, 2119–2133 (2009).

[47] Šulc, P. *et al.* Sequence-dependent thermodynamics of a coarse-grained dna model. *J. Chem. Phys* **137**, 135101 (2012).

[48] Flory, P. J. *Statistical Mechanics of Chain Molecules* (Hanser, 1969).

[49] Caracciolo, S. & Sokal, A. D. Dynamic critical exponent of some monte carlo algorithms for the self-avoiding walk. *J. Phys. A: Math. Gen.* **19**, L797–L801 (1986).

[50] Ouldridge, T. E., Louis, A. A. & Doye, J. P. K. Structural, mechanical, and thermodynamic properties of a coarse-grained dna model. *J. Chem. Phys* **134**, 02B627 (2011).

[51] Snodin, B. E. K. *et al.* Introducing improved structural properties and salt dependence into a coarse-grained model of dna. *J. Chem. Phys* **142**, 234901 (2015).

[52] Ouldridge, T. E., Louis, A. A. & Doye, J. P. K. Dna nanotweezers studied with a coarse-grained model of dna. *Phys. Rev. Lett.* **104**, 178101 (2010).

[53] Henrich, O., Gutierrez-Fosado, Y., Curk, T. & Ouldridge, T. Coarse-grained simulation of dna using lammps. *Eur. Phys. J. E* **41**, 57 (2018).

[54] Plimpton, S. J. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).

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

[56] Scharein, R. Knotplot. URL http://www.knotplot.com/.

[57] Klenin, K. & Langowski, J. Computation of writhe in modeling of supercoiled dna. *Biopolymers* **54**, 307–317 (2000).

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

[59] Tubiana, L., Polles, G., Orlandini, E. & Micheletti, C. Kymoknot: A web server and software package to identify and locate knots in trajectories of linear or circular polymers. *The European Physical Journal E* **41**, 72 (2018).

[60] Adams, C. C. *The Knot Book* (Freeman, 1994).

[61] J. H. White, J. H. Self-linking and the gauss integral in higher dimensions. *Am. J. Math.* **91**, 683–728 (1969).

[62] Frank-Kamenetskii, M. D., Lukashin, A. V. & Vologodskii, A. V. Statistical mechanics and topology of polymer chains. *Nature* **258**, 398–402 (1975).

[63] Buck, G. R. & Zechiedrich, E. L. Dna disentangling by type-2 topoisomerases. *Journal of Molecular Biology* **340**, 933 – 939 (2004).

[64] Sumners, D. W. & Whittington, S. G. Knots in self-avoiding walks. *Journal of Physics A: Mathematical and General* **21**, 1689–1694 (1988).

[65] Deguchi, T. & Tsurusaki, K. A statistical study of random knotting using the vassiliev invariants. *Journal of Knot Theory and Its Ramifications* **03**, 321–353 (1994).

[66] K. Koniaris, M. M. Knottedness in ring polymers. *Phys. Rev. Lett.* **66**, 2211–2214 (1991).

[67] Virnau, P., Kantor, Y. & Kardar, M. Knots in globule and coil phases of a model polyethylene. *J. Am. Chem. Soc.* **127**, 15102–15106 (2005).

[68] Bao, X. R., Lee, H. J. & Quake, S. R. Behavior of complex knots in single dna molecules. *Phys. Rev. Lett.* **91**, 265506 (2003).

[69] M. K. Shimamura, T. D. Characteristic length of random knotting for cylindrical self-avoiding polygons. *Phys. Lett. A* **274**, 184–191 (2000).

[70] Orlandini, E., Tesi, M. C. & Whittington, S. G. Entanglement complexity of semiflexible lattice polygons. *J. Phys. A: Math. Gen.* **38**, L795–L800 (2005).

[71] Virnau, P., Rieger, F. C. & Reith, D. Influence of chain stiffness on knottedness in single polymers. *Biochemical Society Transactions* **41**, 528–532 (2013).

[72] Coronel, L., Orlandini, E. & Micheletti, C. Non-monotonic knotting probability and knot length of semiflexible rings: the competing roles of entropy and bending energy. *Soft matter* **13**, 4260–4267 (2017).

[73] Fary, I. Sur la courbure totale d'une courbe gauche faisant un nœud. *Bulletin de la Société Mathématique de France* **77**, 128–138 (1949).

[74] Milnor, J. W. On the total curvature of knots. *Annals of Mathematics* **52** (1950).

[75] A. M. Ferrenberg, R. H. S. Optimized monte carlo data analysis. *Phys. Rev. Lett.* **63**, 1195–1198 (1989).

[76] D'Adamo, G., Dietler, G. & Micheletti, C. Tuning knot abundance in semiflexible chains with crowders of different sizes: a monte carlo study of dna chains. *Soft matter* **12**, 6708–6715 (2016).

[77] Matthews, R., Louis, A. A. & Likos, C. N. Effect of bending rigidity on the knotting of a polymer under tension. *ACS Macro Letters* **1**, 1352–1356 (2012).

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

[79] Uehara, E. & Deguchi, T. Characteristic length of the knotting probability revisited. *Journal of Physics: Condensed Matter* **27**, 354104 (2015).

[80] Uehara, E. & Deguchi, T. Knotting probability of self-avoiding polygons under a topological constraint. *The Journal of Chemical Physics* **147**, 094901 (2017).

[81] Uehara, E., Coronel, L., Micheletti, C. & Deguchi, T. Bimodality in the knotting probability of semiflexible rings suggested by mapping with self-avoiding polygons. *Reactive and Functional Polymers* **134**, 141 – 149 (2019).

[82] E. J. Janse van Rensburg, S. G. W. The knot probability in lattice polygons. *J. Phys. A: Math. Theor.* **23**, 3573–3590 (1990).

[83] Janse van Rensburg, E. J. & A Rechnitzer, A. On the universality of knot probability ratios. *Journal of Physics A: Mathematical and Theoretical* **44**, 162002 (2011).

[84] Bates, A. D., O'Dea, M. H. & Gellert, M. Energy coupling in Escherichia coli DNA gyrase. *Biochemistry* **35**, 1408–1416 (1996).

[85] Drlica, K. Control of bacterial DNA supercoiling. *Mol. Microbiol.* **6**, 425–433 (1992).

[86] Zechiedrich, E. L. *et al.* Roles of Topoisomerases in Maintaining Steady-state DNA Supercoiling in Escherichia coli. *J. Biol. Chem.* **275**, 8103–8113 (2000).

[87] Wasserman, S. A., Dungan, J. M. & Cozzarelli, N. R. Discovery of a predicted DNA knot substantiates a model for site-specific recombination. *Science* **229**, 171–174 (1985).

[88] Olavarrieta, L., Hernández, P., Krimer, D. B. & Schvartzman, J. B. DNA Knotting Caused by Head-on Collision of Transcription and Replication. *J. Mol. Biol.* **322**, 1–6 (2002).

[89] López, V., Martínez-Robles, M.-L., Hernández, P., Krimer, D. B. & Schvartzman, J. B. Topo IV is the topoisomerase that knots and unknots sister duplexes during DNA replication. *Nucleic Acids Res.* **40**, 3563–3573 (2011).

[90] Rybenkov, V., Ullsperger, C., Vologodskii, A. V. & Cozzarelli, N. R. Simplification of dna topology below equilibrium values by type ii topoisomerases. *Science* **277**, 690–693 (1997).

[91] Deibler, R. W., Rahmati, S. & Zechiedrich, E. L. Topoisomerase IV, alone, unknots DNA in E. coli. *Genes Dev* **15**, 748–761 (2001).

[92] Krasnow, M. A. *et al.* Determination of the absolute handedness of knots and catenanes of DNA. *Nature* **304**, 559–560 (1983).

[93] Postow, L., Crisona, N. J., Peter, B. J., Hardy, C. D. & Cozzarelli, N. R. Topological challenges to DNA replication: conformations at the fork. *Proc. Natl. Acad. Sci. USA* **98**, 8219–8226 (2001).

[94] Marenduzzo, D. *et al.* DNA-DNA interactions in bacteriophage capsids are responsible for the observed DNA knotting. *Proc. Natl. Acad. Sci. USA* **106**, 22269–22274 (2009).

[95] Krajina, B. A. & Spakowitz, A. J. Large-Scale Conformational Transitions in Supercoiled DNA Revealed by Coarse-Grained Simulation. *Biophys. J.* **111**, 1339–1349 (2016).

[96] Boles, T. C., White, J. H. & Cozzarelli, N. R. Structure of plectonemically supercoiled dna. *J. Mol. Biol.* **213**, 931–951 (1990).

[97] Adrian, M. *et al.* Direct visualization of supercoiled dna molecules in solution. *The EMBO journal* **9**, 4551–4554 (1990).

[98] Vologodskii, A. V., Levene, S. D., Klenin, K. V., Frank-Kamenetskii, M. & Cozzarelli, N. R. Conformational and thermodynamic properties of supercoiled dna. *J. Mol. Biol.* **227**, 1224–1243 (1992).

[99] Marko, J. F. The internal slitheringdynamics of supercoiled dna. *Physica A Stat. Mech. Appl.* **244**, 263–277 (1997).

[100] Jian, H., Schlick, T. & Vologodskii, A. Internal motion of supercoiled DNA: brownian dynamics simulations of site juxtaposition. *J. Mol. Biol.* **284**, 287–296 (1998).

[101] Huang, J., Schlick, T. & Vologodskii, A. Dynamics of site juxtaposition in supercoiled DNA. *Proc. Natl. Acad. Sci. USA* **98**, 968–973 (2001).

[102] van Loenhout, M. T., de Grunt, M. & Dekker, C. Dynamics of dna supercoils. *Science* **338**, 94–97 (2012).

[103] Hatfield, G. W. & Benham, C. J. Dna topology-mediated control of global gene expression in escherichia coli. *Annu. Rev. Genet.* **36**, 175–203 (2002).

[104] Sutthibutpong, T. *et al.* Long-range correlations in the mechanics of small dna circles under topological stress revealed by multi-scale simulation. *Nucleic Acids Res.* **44**, 9121–9130 (2016).

[105] Jeon, J.-H., Adamcik, J., Dietler, G. & Metzler, R. Supercoiling induces denaturation bubbles in circular dna. *Phys. Rev. Lett.* **105**, 208101 (2010).

[106] Irobalieva, R. N. *et al.* Structural diversity of supercoiled DNA. *Nat. Commun.* **6**, 1–10 (1).

[107] Mitchell, J. S., Laughton, C. A. & Harris, S. A. Atomistic simulations reveal bubbles, kinks and wrinkles in supercoiled dna. *Nucleic Acids Res.* **39**, 3928–3938 (2011).

[108] Hwa, T., Marinari, E., Sneppen, K. & Tang, L.-h. Localization of denaturation bubbles in random dna sequences. *Proc. Natl. Acad. Sci. USA* **100**, 4411–4416 (2003).

[109] Jost, D., Zubair, A. & Everaers, R. Bubble statistics and positioning in superhelically stressed dna. *Phys. Rev. E* **84**, 031912 (2011).

[110] Matek, C., Ouldridge, T. E., Doye, J. P. K. & Louis, A. A. Plectoneme tip bubbles: coupled denaturation and writhing in supercoiled dna. *Sci. Rep.* **5**, 7655 (2015).

[111] Nelson, D. L. & Cox, M. M. *Lehninger Principles of Biochemistry, Fourth Edition* (2004), fourth edition edn.

[112] Nitiss, J. L. Dna topoisomerase ii and its growing repertoire of biological functions. *Nature Reviews Cancer* **9**, 327337 (2009).

[113] Forterre, P. Dna topoisomerase v: a new fold of mysterious origin. *Trends in Biotechnology* **24**, 245 – 247 (2006).

[114] Baker, N. M., Rajan, R. & Mondragn, A. Structural studies of type I topoisomerases. *Nucleic Acids Research* **37**, 693–701 (2008).

[115] Goto, T. & Wang, J. C. Yeast dna topoisomerase ii. an atp-dependent type ii topoisomerase that catalyzes the catenation, decatenation, unknotting, and relaxation of double-stranded dna rings. *J. Biol. Chem.* **257**, 5866–5872 (1982).

[116] Randall, G. L., Pettitt, B. M., Buck, G. R. & Zechiedrich, E. L. Electrostatics of dna–dna juxtapositions: consequences for type ii topoisomerase function. *J. Phys. Condens. Matter.* **18**, S173 (2006).

[117] Liu, Z., Mann, J. K., Zechiedrich, E. L. & Chan, H. S. Topological information embodied in local juxtaposition geometry provides a statistical mechanical basis for unknotting by type-2 dna topoisomerases. *J. Mol. Biol.* **361**, 268–285 (2006).

[118] Liu, Z. & Chan, H. S. Consistent rationalization of type-2 topoisomerases' unknotting, decatenating, supercoil-relaxing actions and their scaling relation. *J. Phys. Condens. Matter.* **27**, 354103 (2015).

[119] Dong, K. G. & Berger, J. M. Structural basis for gate-dna recognition and bending by type iia topoisomerases. *Nature* **450**, 1201–1205 (2007).

[120] Vologodskii, A. V. *et al.* Mechanism of topology simplification by type ii dna topoisomerases. *Proceedings of the National Academy of Sciences* **98**, 3045–3049 (2001).

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

[122] Shishido, K., Ishii, S. & Komiyama, N. The presence of the region on pBR322 that encodes resistance to tetracycline is responsible for high levels of plasmid DNA knotting in Echerichia coli DNA topoisomerase I deletion mutant. *Nucleic Acids Res.* **17**, 9749–9759 (1989).

[123] Witz, G., Dietler, G. & Stasiak, A. Tightening of DNA knots by supercoiling facilitates their unknotting by type II DNA topoisomerases. *Proc. Natl. Acad. Sci. USA* **108**, 3608–3611 (2011).

[124] Racko, D., Benedetti, F., Dorier, J., Burnier, Y. & Stasiak, A. Generation of supercoils in nicked and gapped dna drives dna unknotting and postreplicative decatenation. *Nucleic Acids Res.* **43**, 7229–7236 (2015).

[125] Rawdon, E. J., Dorier, J., Racko, D., Millett, K. C. & Stasiak, A. How topoisomerase iv can efficiently unknot and decatenate negatively supercoiled dna molecules without causing their torsional relaxation. *Nucleic Acids Res.* **44**, 4528–4538 (2016).

[126] Coronel, L., Suma, A. & Micheletti, C. Dynamics of supercoiled DNA with complex knots: large-scale rearrangements and persistent multi-strand interlocking. *Nucleic Acids Research* **46**, 7533–7541 (2018).

[127] Suma, A. & Micheletti, C. Pore translocation of knotted DNA rings. *Proc. Natl. Acad. Sci. USA* **114**, E2991–E2997 (2017).

[128] Bliska, J. B. & Cozzarelli, N. R. Use of site-specific recombination as a probe of DNA structure and metabolism in vivo. *Journal of Molecular Biology* **197**, 205–218 (1987).

[129] Rosa, A., Di Ventra, M. & Micheletti, C. Topological jamming of spontaneously knotted polyelectrolyte chains driven through a nanopore. *Phys. Rev. Lett.* **109**, 118301 (2012).

[130] Suma, A., Rosa, A. & Micheletti, C. Pore translocation of knotted polymer chains: How friction depends on knot complexity. *ACS Macro Lett.* **4**, 1420–1424 (2015).

[131] Burnier, Y., Dorier, J. & Stasiak, A. Dna supercoiling inhibits dna knotting. *Nucleic Acids Res.* **36**, 4956–4963 (2008).

[132] Huang, J.-Y. & Lai, P.-Y. Crossings and writhe of flexible and ideal knots. *Phys. Rev. E* **63**, 021506 (2001).

[133] Hammermann, M. *et al.* Salt effects on the structure and internal dynamics of superhelical dnas studied by light scattering and brownian dynamics. *Biophysical journal* **73**, 2674–2687 (1997).

[134] Liu, Z. & Chan, H. S. Efficient chain moves for monte carlo simulations of a wormlike dna model: Excluded volume, supercoils, site juxtapositions, knots, and comparisons with random-flight and lattice models. *J. Chem. Phys.* **128**, 145104 (2008).

[135] Witz, G. *PhD Thesis* (École Polytechnique Fédérale de Lausanne, 2010).

[136] Caraglio, M., Micheletti, C. & Orlandini, E. Physical links: Defining and detecting inter-chain entanglement. *Sci. Rep.* **7**, 1156 (2017).

[137] Shaw, S. Y. & Wang, J. C. Chirality of dna trefoils: implications in intramolecular synapsis of distant dna segments. *Proc. Natl. Acad. Sci. USA* **94**, 1692–1697 (1997).

[138] Szymczak, P. Translocation of knotted proteins through a pore. *Eur Phys J Spec* **223**, 1805–1812 (2014).

[139] Noom, M. C., Van Den Broek, B., Van Mameren, J. & Wuite, G. J. Visualizing single dna-bound proteins using dna as a scanning probe. *Nat. methods* **4**, 1031 (2007).

[140] Dekker, C. Solid-state nanopores. *Nat. nanotechnol.* **2**, 209 (2007).

[141] Steinbock, L. J., Bulushev, R. D., Krishnan, S., Raillon, C. & Radenovic, A. Dna translocation through low-noise glass nanopores. *ACS Nano* **7**, 11255–11262 (2013).

[142] Trepagnier, E. H., Radenovic, A., Sivak, D., Geissler, P. & Liphardt, J. Controlling dna capture and propagation through artificial nanopores. *Nano Letters* **7**, 2824–2830 (2007). PMID: 17705552.

[143] Traversi, F. *et al.* Detecting the translocation of dna through a nanopore using graphene nanoribbons. *Nat. nanotechnol.* **8**, 939 (2013).

[144] Joshi, R. S., Pia, B. & Roca, J. Topoisomerase II is required for the production of long Pol II gene transcripts in yeast. *Nucleic Acids Research* **40**, 7907–7915 (2012).

[145] Hsieh, T.-S. Knotting of the circular duplex dna by type 11 dna topoisomerase from drosophila melanogaster. *J. Biol. Chem.* **258**, 8413–8420 (1983).

[146] Roca, J., Berger, J. M. & Wang, J. C. On the simultaneous binding of eukaryotic dna topoisomerase ii to a pair of double-stranded dna helices. *The Journal of Biological Chemistry* **268**, 14250–14255 (1993).

[147] Micheletti, C., Marenduzzo, D., Orlandini, E. & Sumners, D. W. Simulations of knotting in confined circular dna. *Biophys. J.* **95**, 3591–3599 (2008).

[148] Arsuaga, J., Vázquez, M., Trigueros, S., Sumners, D. W. & Roca, J. Knotting probability of dna molecules confined in restricted volumes: Dna knotting in phage capsids. *Proceedings of the National Academy of Sciences* **99**, 5373–5377 (2002).

[149] Naughton, C. *et al.* Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. *Nature Structural &Amp; Molecular Biology* **20** (2013).

[150] Lavelle, C., Victor, J.-M. & Zlatanova, J. Chromatin fiber dynamics under tension and torsion. *International Journal of Molecular Sciences* **11**, 15571579 (2010).

[151] Valdés, A. *et al.* Transcriptional supercoiling boosts topoisomerase II-mediated knotting of intracellular DNA. *Nucleic Acids Research* **47**, 6946–6955 (2019).

[152] Hoste, J. & Thistlethwaite, M. Knotscape 1.0. URL http://pzacad.pitzer.edu/~jhoste/HosteWebPages/kntscp.html.

[153] Salceda, J., Fernández, X. & Roca, J. Topoisomerase ii, not topoisomerase i, is the proficient relaxase of nucleosomal dna. *The EMBO Journal* **25**, 2575–2583 (2006).