UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO di FISICA e ASTRONOMIA "GALILEO GALILEI"
Corso di Laurea Magistrale in Fisica

TESI DI LAUREA MAGISTRALE

Dynamic properties of coarse-grained models of linear and circular DNA in nanochannels

Internal supervisor:
Prof. Enzo Orlandini
External supervisor:
Prof. Cristian Micheletti

Student:
Francesca Rizzato
Matricola: 1014337

Academic Year 2011/2012

Contents

Riassunto

Questa tesi ha lo scopo di analizzare, tramite simulazioni di dinamica molecolare, le proprietà di equilibrio e quelle cinetiche di catene di DNA lineari e circolari (125-250 nm) confinate in un nanocanale di larghezza paragonabile alle loro dimensioni.

Questo studio è motivato dai risultati di recenti analisi sperimentali e computazionali [1, 2, 3] che hanno dimostrato che un forte confinamento in un canale può avere pesanti conseguenze sulle proprietà dei polimeri che contiene, ad esempio un diverso coefficiente di diffusione per catene di diversa lunghezza [4]. L'eventuale scoperta di caratteristiche dinamiche differenti dipendenti soltanto dalla topologia della catena (lineare, circolare o annodata), potrebbe portare alla progettazione di dispositivi in grado di separare, tramite dei nanocanali, catene con vincoli topologici diversi. L'importanza pratica di risultati di questo tipo è evidente se si considera che la presenza di nodi in catene di DNA influenza pesantemente l'efficienza di alcuni eventi genetici fondamentali, come la replicazione e la trascrizione.

Abbiamo modellizzato i filamenti di DNA tramite catene coarse-grained semiflessibili di sfere autoevitantisi e le abbiamo inserite in dei nanocanali quadrati di varia dimensione (32.7-65.4 nm) caratterizzati da pareti ruvide. Abbiamo quindi fatto evolvere il sistema tramite delle simulazioni di Dinamica Molecolare basate su un approccio di Langevin sovrasmorzato. Alla luce di precedenti studi computazionali [5] che, pur trascurando l'idrodinamica, hanno evidenzato delle differenze dipendenti dalla sola topologia nel comportamento cinetico di polimeri sottoposti ad elettroforesi, abbiamo deciso di non includere nel nostro modello gli effetti idrodinamici, in modo tale da rendere il procedimento utilizzato il più trasparente possibile.

Per quanto riguarda le proprietà di equilibrio, abbiamo constatato che l'andamento delle dimensioni medie per le catene confinate è significativamente diverso tra catene aperte e chiuse a tutti i gradi di confinamento, confermando quanto già trovato da studi precedenti [3, 6] e abbiamo osservato come la riduzione della dimensione del canale provochi un'orientazione preferenziale della catena sempre più marcata, come previsto da considerazioni teoriche [7]. Abbiamo completato lo studio delle proprietà di equilibrio analizzando la distribuzione dei monomeri nel piano trasverso del canale e verificando che essi si dispongono con massima probabilità nelle vicinanze delle pareti, a causa dell'attrito che ne rallenta il moto.

Contrariamente a quanto atteso, lo studio delle proprietà dinamiche non ha fatto emergere, a parità di lunghezza di catena, delle differenze significative nel coefficiente di diffusione per catene confinate di diversa topologia, mentre variazioni sostanziali sono state osservate in funzione della lunghezza della catena. Questo ci porta a pensare che,

per vedere una dipendenza dalla topologia, attesao a causa delle diverse dimensioni medie di catena libera, sia necessario includere nel modello anche gli effetti idrodinamici. Abbiamo concluso l'analisi della dinamica delle catene confinate studiando l'andamento del tempo di Rouse in funzione del confinamento ed abbiamo trovato che, in accordo con alcuni studi sperimentali, esso è più grande per le catene di DNA circolari che per le corrispondenti catene lineari.

Chapter 1

Introduction

The aim of the present work is to analyse, by Molecular Dynamics (MD) simulations, the equilibrium and kinetic properties of linear and circular double stranded DNA chains confined inside rough nanochannels, by focusing, in particular, on the interplay between topological and spatial constraints.

This study is motivated by the great advancements in the fabrication of nanofluidic devices that, in recent years, have enabled accurate experimental investigations of the effects that a severe confinement may have on the conformational and dynamical properties of DNA chains [1]. In particular recent experimental and computational studies have shown that, in narrow nanochannels, chains with different contour lengths have different diffusion coefficients [4] and relaxation times[2]. This mechanism is nowadays exploited to design nanofluidic devices capable of separating, by lengths, a polydispersed solution of linear DNA filaments. The possible emergence of different dynamical features for chains of constant length but different topology (linear, circular or knotted chains) may therefore allow for future applications, such as the design of devices, based on nanochannels, capable to sort chains out and separate the different topologies. The practical importance of these studies becomes evident if we consider that knots and other entanglements in DNA severely affect the efficiency of fundamental genetic events, such as replication and transcription [3].

In this thesis the DNA filaments are modelled as coarse-grained self-avoiding semi-flexible chains of beads, whose parameters are set to match those of DNA in a physiological solution of monovalent counterions. The MD simulations are based on the overdamped Langevin equation and the confining square channels have a size comparable with the chains dimension in bulk. Since previous computational studies [5] have established topology-dependent differences in the kinetic properties of chains in electrophoretic conditions even without considering hydrodynamic effects, we choose not to include them in our model, in order to make the simulation setup as trasparent as possible.

We will first analyse and discuss the equilibrium properties for confined open and closed chains, by focusing, in particular, on the scaling of the chain average size, on the spatial correlation of monomers and on the distribution of the chain monomers in the channel transverse plane, as functions of the degree of confinement.

As far as the kinetic properties are concerned, we will investigate the impact of chain topology, contour length and confinement on both the overall motion of the molecule and the dynamics of single monomers. In particular, we will study the variation of the

diffusion coefficient of the chains centre of mass and compare our findings with what found in previous studies. We will also analyse the evolution of the Rouse time for different degrees of confinement and topologies, in order to find out possible interesting differences between linear and circular chains.

The material presented in this dissertation is organised as follows:

- in chapter 2 we will discuss the coarse-grained model used to describe DNA chains, focusing in particular on the intra-chain potential energy; we will then introduce the observables that characterize polymer configurations, such as size and spatial correlation, which will be necessary in the prosecution of the work;
- chapter 3 describes the computational method of Molecular Dynamics and the Langevin approach used in the simulations and gives details about the choice of parameters;
- in chapter 4 we will report on the model for the rough nanochannel embedding the confined polymers and we will discuss some theoretical properties of confined chains;
- chapter 5 contains our results concerning equilibrium and dynamical properties of the simulated chains: the first part enlightens our findings on the equilibrium metric properties of the chains, while the second part focuses on the diffusion of the centre of mass and on the motion of single monomers in the chain;
- finally, chapter 6 sums up the obtained results and gives an outlook on possible future improvements and on other perspectives for the analysed problems.

Chapter 2

DNA chain model

Linear polymers are molecules obtained by the repetitive addition of monomeric units at one or both ends of the chain and, in particular, DNA is a biopolymer composed of four kinds of nucleotides containing a backbone of sugar (desoxyribose), a nucleobase (adenine, guanine, cytosine or thymine) and a negatively charged phosphate group each. In living organisms, DNA is generally found in a double-helix structure with bare diameter of 2 nm (double-stranded DNA or dsDNA), where the nucleobases of two opposite strands bind through hydrogen bonds (adenine-thymine and cytosine-guanine) forming a base pair as in *Figure* 2.1.

The dsDNA is a semi-flexible polymer, as we can see from the image of a DNA molecule in *Figure* 2.2, taken through an atomic force microscope (AFM)[8]. A similar method (photography through an electron microscope) provided in 1979 [9] the first direct measure of the DNA persistence length, that is the length over which the molecule tends to stay stiff: while its bare diameter measures 2 nm, the DNA persistence length is about 50 nm. It means that, if we observe a molecule of DNA at very small scale, for example at the scale of its diameter, we will see it as a stiff rod, while, if we analyse its behaviour on distances of the order of some micrometers, it will look like a flexible chain.

From the chemical point of view, DNA is a very complex molecule, so describing its motion with all details, would need a huge amount of calculations and time: typically a full-atom simulation of a 30-nm DNA chain can treat time-scales from ps to maximum μs, while it is totally unthinkable for longer periods. As we are not interested in the motion of each single atom, but in the macroscopic behaviour of the entire molecule, a natural solution, called *coarse-graining*, is to keep the number of variables to a minimum, for example by considering one or a few base pairs as a single entity and leaving behind all the atomic details. The simulation of a coarse-grained system requires, in fact, less resources and time than the full-atom one, increasing of orders of magnitude the simulation length and time scales that can be achieved. In this way, instead of considering the DNA filament as a set of atoms bound to each other, we will treat it as a coarse-grained chain of beads held together by a potential chosen in such a way to reproduce as well as possible DNA macroscopic features.

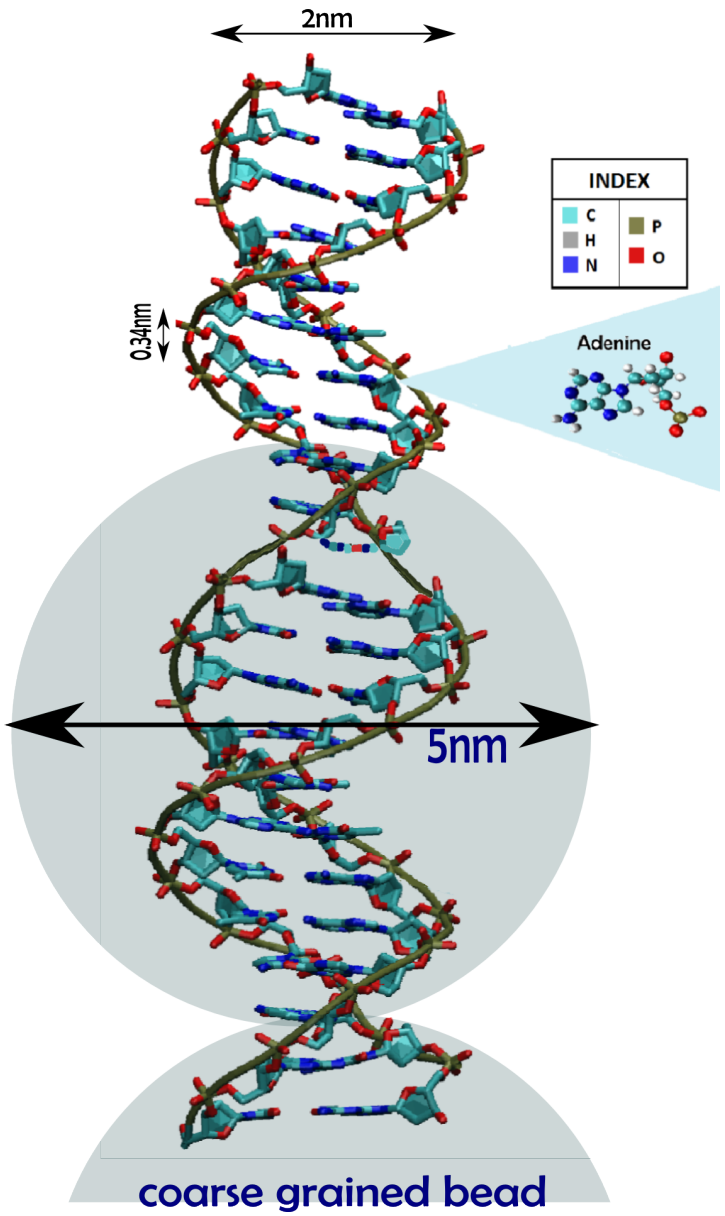


Figure 2.1: *dsDNA in the double-helix structure, with the superposition of some coarse-grained beads*

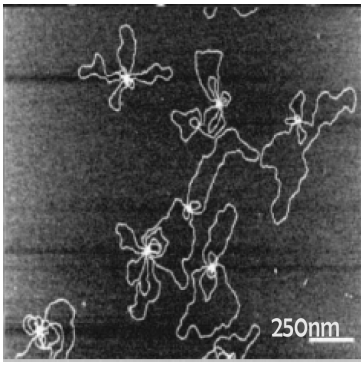


Figure 2.2: *Image of knotted dsDNA from atomic force microscopy (AFM) adapted from ref. [8]*

2.1 Model for linear and circular DNA chains

We will describe a DNA filament as a chain of N beads. With reference to a definite Cartesian frame, the spatial position of the beads centers will be indicated as $\{\vec{R}_i\} = \{\vec{R}_0, \cdots, \vec{R}_{N-1}\}$ and the set of the n bond vectors connecting two neighbouring beads as $\{\vec{r}_i\} = \{\vec{R}_i - \vec{R}_{i-1}\} = \{\vec{r}_1, \cdots, \vec{r}_n\}$, as shown in *Figure* 2.3. Let us remark that, for a fixed number of beads N the number of bonds is $n = N - 1$ for a linear chain and $n = N$ for a circular one.

The simplified structural representation of DNA requires, in turn, to describe the molecule self-interaction with a simplified effective potential energy which accounts for:

- the chain connectivity, given by the chemical bonds between consecutive beads;
- the chain excluded volume interaction;
- the chain bending rigidity;
- the electrostatic self-repulsion, being the DNA a negatively charged polyelectrolite.

The simulated chain is plunged into a solution rich in monovalent counterions, so that the screened electrostatic self-repulsion of the molecule is captured through an enhancement of the molecule thickness, r_0, as we will see in *Section* 3.5.

Then, the total intra-chain energy of the polymer is the sum of the first three components mentioned above:

$$U_{in} = U_{bond} + U_{ev} + U_{bend} \tag{2.1}$$

The total force applied to the centre of mass, $\vec{F}_{CM}$, is zero because there are no external interactions, while the total internal force acting on the i-th bead due to the intra-chain potential energy, is:

$$\vec{F}_{i,in} = -\vec{\nabla}_i U_{in} \tag{2.2}$$

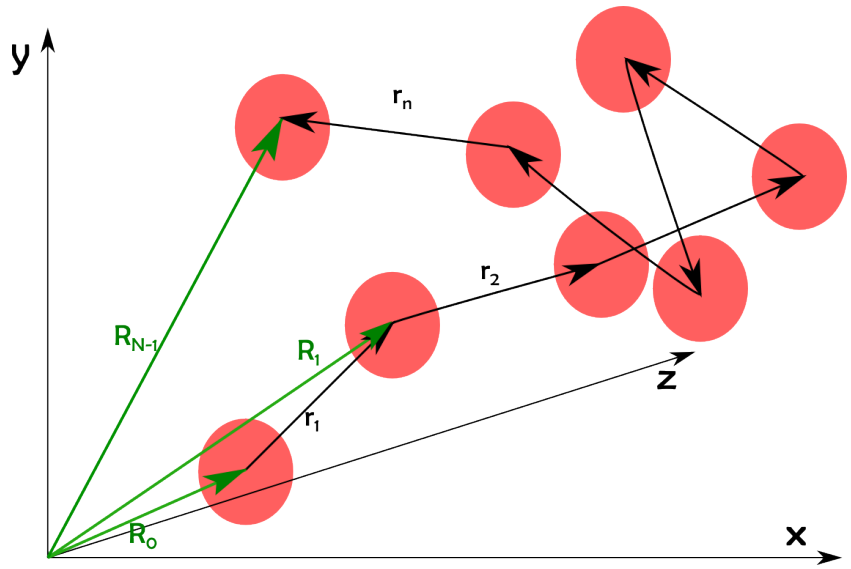


Figure 2.3: *Parametrizations of a linear chain by the set of the N position vectors $\{\vec{R}_i\} = \{\vec{R}_0, \cdots, \vec{R}_{N-1}\}$ of the beads (green), or by the set of the $n = N-1$ bond vectors $\{\vec{r}_i\} = \{\vec{r}_1, \cdots, \vec{r}_n\}$ (black)*

where $\vec{\nabla}_i$ is the gradient calculated along the coordinates of the i-th bead.

We will now analyse in detail each component of the intra-chain energy for linear and circular DNA filaments.

2.1.1 Lennard-Jones potential

As used in most of MD simulations, we modelled the excluded volume effect between two beads by the repulsive part of a Lennard-Jones (LJ) potential, which is characterized by a strong repulsive core ($\propto 1/r^{12}$) and a weak attractive tail ($\propto 1/r^6$) (blue curve in *Fig.* 2.4). The LJ potential acts on all the possible pairs of beads, so that two of them can not occupy positions too close to each other, as in the real world two balls of a certain volume can not overlap.

This introduces a total excluded volume energy of:

$$U_{ev} = 4\varepsilon \sum_{i=0}^{N-2} \sum_{j>i}^{N-1} \left[\left(\frac{r_0}{r_{ij}} \right)^{12} - \left(\frac{r_0}{r_{ij}} \right)^{6} + \frac{1}{4} \right] \Theta(2^{\frac{1}{6}} r_0 - r_{ij}) \tag{2.3}$$

where r_0 is the effective[1] diameter of a bead, N is the number of beads, ε is a factor controlling the strength of the potential, r_{ij} is the distance between bead i and bead j and $\Theta(x)$ is the Heaviside theta function. The constant term, $1/4$, was added to have zero

[1] the meaning of *effective* will be clarified in Section 3.5.1

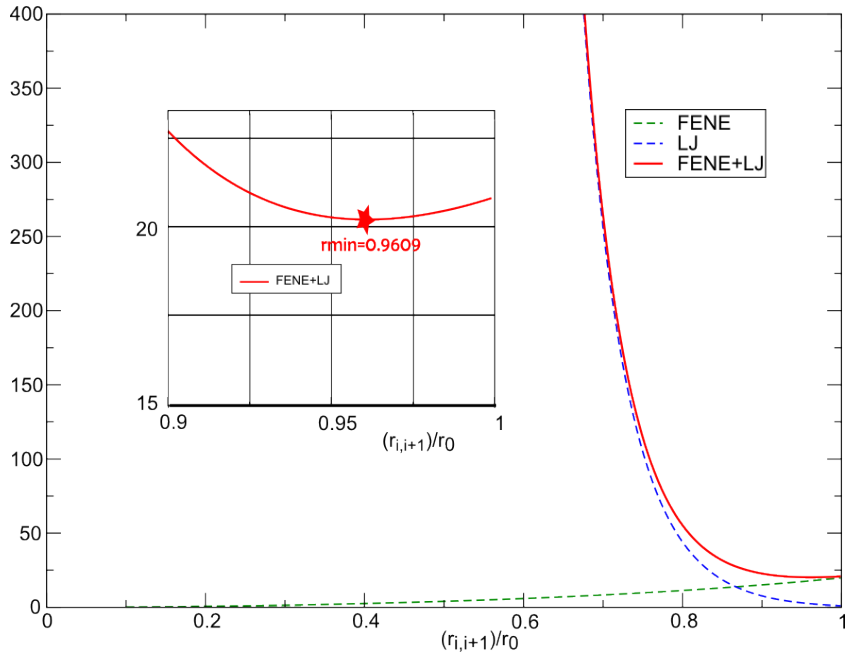


Figure 2.4: *U_{bond}, U_{ev} and $U_{bond}+U_{ev}$ for consecutive beads as function of their relative distance $r_{i,i+1}$ in units of bead's diameter*

when $r_{ij} = 2^{\frac{1}{6}} r_0$ and assures in this way the continuity of the potential: the LJ potential modified in this way is known as the Weeks-Chandler-Andsersen (WCA) potential [10].

It is important to notice that, with the growing of the chain size N, the LJ contribution, acting on every possible pair, increases the computational time by a factor N^2, while we will see that binding and bending, which act only between subsequent beads, increase it by just a factor N. We can conclude that the constraint of self-avoidance is the main upper limit determining the maximal length of the polymer in function of the available computational time.

2.1.2 Finite Extensible Non-linear Elastic potential

In order to mimic the finite extensibility of DNA chains subject to stretching, we make two consecutive beads i and $i+1$ interact by the Finite Extensible Non-linear Elastic(***FENE***) potential [11] (shown in green in *Figure* 2.4):

$$u_{bond,(i\leftrightarrow i+1)} = \begin{cases} -\frac{k}{2} \cdot R_{max}^2 \ln\left[1 - \left(\frac{r_{i,i+1}}{R_{max}}\right)^2\right] & \text{if } r_{i,i+1} \leq R_{max}, \\ +\infty & \text{if } r_{i,i+1} > R_{max} \end{cases} \tag{2.4}$$

where R_{max} is the maximal bond length and $r_{i,i+1}$ is the Euclidian distance between bead i and bead $(i+1)$. The parameter k is strongly dependent on the ε appearing

in equation 2.3 and is chosen in such a way that the sum of LJ and FENE potentials between subsequent beads (in red in Fig. 2.4) have a minimum energy point assuring the presence of an equilibrium bond length between consecutive beads, as shown by the inset in *Figure* 2.4. On the other hand, from Figure 2.4, one can also deduce that the bond length distribution induced by the superposition of LJ and FENE potentials is asymmetric, so that the average bond length is expectedly greater than the bond length for which the energy minimum occurs.

By summing on all bonds, we obtain the total bond energy:

$$U_{bond} = \sum_{i=0}^{\tilde{N}} u_{bond,(i\leftrightarrow i+1)} \tag{2.5}$$

where

$$\tilde{N} = \begin{cases} N-2 & \text{if linear chain} \\ N-1 & \text{if circular chain} \end{cases} \tag{2.6}$$

being N the total number of beads, with the implied identification $R_N \equiv R_0$ for the circular chain.

2.1.3 Bending rigidity

We shall now insert an elastic energy which hinders the bending between consecutive bonds (*Kratky-Porod model* [12]) and compute the intra-chain energy due to bending as:

$$U_{bend} = k_{bend} \sum_{i=0}^{\tilde{N}} \left(1 - \frac{\vec{r}_{i,i+1} \cdot \vec{r}_{i+1,i+2}}{r_{i,i+1} r_{i+1,i+2}}\right) \tag{2.7}$$

where k_{bend} is a prefactor determining the strength of the potential, $\vec{r}_{i,i+1}$ is the distance between two consecutive beads and $r_{i,i+1}$ is its modulus. In this case:

$$\tilde{N} = \begin{cases} N-3 & \text{if linear chain} \\ N-1 & \text{if circular chain} \end{cases} \tag{2.8}$$

with the usual identification of $R_{N+i} \equiv R_i$.

Let us notice that the expression $\frac{\vec{r}_{i,i+1} \cdot \vec{r}_{i+1,i+2}}{r_{i,i+1} r_{i+1,i+2}}$ is the cosine of the angle θ formed between the two subsequent bonds, as shown in Figure 2.5. In this way, consecutive bonds with small deviations from colinearity have a bending energy of:

$$u_{i,bend} = k_{bend} \cdot \theta^2 \tag{2.9}$$

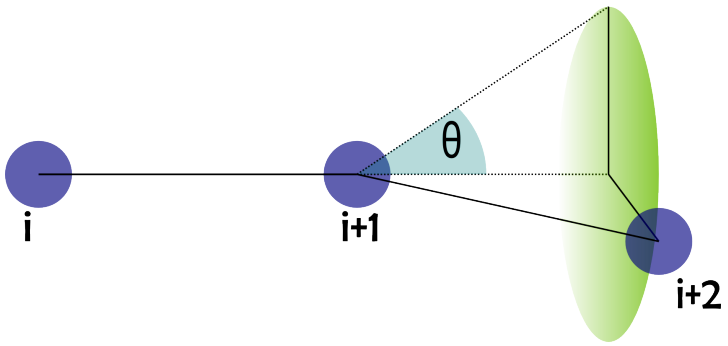


Figure 2.5: *θ angle for consecutive bonds*

2.2 Observables

In this section we will review the most important quantities related to polymer physics, that we will need later when dealing with DNA chains properties.

2.2.1 Effects of the local bending

We will now analyse the effects that the local bending energy has on the correlation of the bonds directionality in the chain by considering the **tangent-tangent correlation function**, $C_{tt}(j)$, that is the correlation between bonds separated by an index j:

$$C_{tt}(j) = \left\langle \frac{\vec{r}_i \cdot \vec{r}_{i+j}}{|\vec{r}_i| \cdot |\vec{r}_{i+j}|} \right\rangle \tag{2.10}$$

where $\vec{r}_i$ is the bond with index i (as in Figure 2.3), $\cdot$ is the scalar product, $|\ |$ is the modulus and $\langle\rangle$ denotes the average over different realizations.

We expect that the correlation function decreases for increasing j: the farther the monomers are, the less the bonds will be correlated. For example, for chains with fixed bond length b and without excluded volume interactions (called ***Kratky-Porod chains***), the decay of the tangent-tangent correlation function is exponential [12]:

$$C_{tt}(j) = e^{-j \cdot b/\tilde{l}_p} \tag{2.11}$$

where the quantity $\tilde{l}_p$ is the ***effective persistence length*** for a Kratky-Porod chain, taking into account its discrete and finite nature. Its value can be directly deduced from equation 2.7[6]:

$$\tilde{l}_p = -\frac{b}{\ln\left[\coth \frac{k_{bend}}{k_B T} - \frac{k_B T}{k_{bend}}\right]} \tag{2.12}$$

In particular, from this equation one can obtain more information on the prefactor in the bending energy equation (eq. 2.7): being the DNA a semi-flexible polymer, we are in

the limit case $k_{bend} >> k_B T$ and by a Taylor expansion of equation 2.12 we get

$$\tilde{l}_p \approx \frac{b k_{bend}}{k_B T}$$

Then, one can rewrite equations 2.7 and 2.12 as

$$U_{bend} = \frac{k_B T l_p}{r_0} \sum_{i=0}^{\tilde{N}} \left(1 - \frac{\vec{r}_{i,i+1} \cdot \vec{r}_{i+1,i+2}}{r_{i,i+1} r_{i+1,i+2}}\right) \tag{2.13}$$

$$\tilde{l}_p = -\frac{b}{\ln\left[\coth \frac{l_p}{r_0} - \frac{r_0}{l_p}\right]} \tag{2.14}$$

where we have approximated the unknown quantities $\tilde{l}_p$ and b by their nominal values l_p and r_0.

2.2.2 Chain dimensions

The length of the chain is called **contour length** (L_c) and can be approximated by

$$L_c \approx N \cdot < b > \tag{2.15}$$

where N the number of monomers and $\langle b \rangle$ is the average bond length between successive beads.

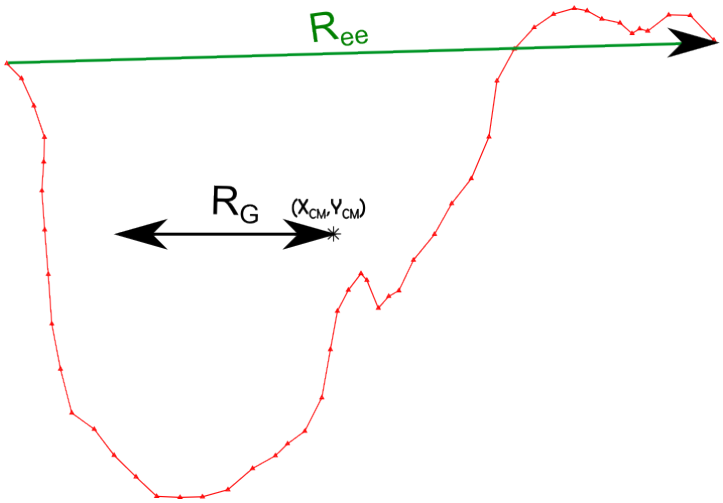


Figure 2.6: *Definition of R_{ee} (in green) and of R_G (in black) for a linear chain of $N = 50$ monomers (in red)*

In order to characterize the size of a polymer, we can consider either the **end-to-end distance**, R_{ee} (in green in *Figure* 2.6) or the **radius of gyration** R_G (in black in the same figure). The former is the distance between the first and the last beads of the chain

$$\vec{R}_{ee} = \vec{R}_{N-1} - \vec{R}_0. \tag{2.16}$$

while the latter is the root mean square distance of a bead from the centre of mass of the system

$$R_G = \sqrt{\frac{\sum_{i=0}^{N-1}(\vec{R}_i - \vec{R}_{CM})^2}{N}} \tag{2.17}$$

where $\vec{R}_i$ and $\vec{R}_{CM}$ are the position vector of the i-th bead and of the centre of mass, respectively.

This last quantity can be rewritten as

$$R_G = \sqrt{\frac{\sum_{i<j}(\vec{R}_i - \vec{R}_j)^2}{N^2}} \tag{2.18}$$

In this work we will generally use the radius of gyration rather than the end-to-end distance, because, when dealing with rings, the latter does not make sense.

In particular, let us also define the parallel ($R_{G,\parallel}$) and perpendicular ($R_{G,\perp}$) radius of gyration of a chain in a channel as

$$R^2_{G,\parallel} = \frac{\sum_{i=0}^{N-1}(X_i - X_{CM})^2}{N} \tag{2.19}$$

$$R^2_{G,\perp} = \frac{\sum_{i=0}^{N-1}\left[(Y_i - Y_{CM})^2 + (Z_i - Z_{CM})^2\right]}{N} \tag{2.20}$$

with the natural consequence that $R^2_G = R^2_{G,\parallel} + R^2_{G,\perp}$, where X is the coordinate along the channel axis and Y and Z are the ones in the constrained directions.

For ideal polymers, i.e. when the excluded volume effects are neglected, these two quantities scale both as the square root of the number of monomers

$$\sqrt{\langle R^2_{ee}\rangle} \sim \sqrt{\langle R^2_G\rangle} \sim \sqrt{N} \tag{2.21}$$

while, for self-avoiding chains it has been shown [13, 14] that they generally scale as

$$\sqrt{\langle R^2_{ee}\rangle} \sim \sqrt{\langle R^2_G\rangle} \sim N^{\nu} \tag{2.22}$$

where $\nu = 0.588$ is a universal exponent valid in three dimensions.

On the other hand, for a Kratky-Porod chain with finite contour length L_c and persistence length l_p a good approximation of $\langle R^2_{ee}\rangle$ is given[6] by the formula:

$$\left\langle R^2_{ee}\right\rangle \approx 2 l_p L_c \left[1 - \frac{l_p}{L_c}\left(1 - e^{-\frac{L_c}{l_p}}\right)\right] \tag{2.23}$$

which derives from equation 2.11.

Chapter 3

Molecular Dynamics

In this chapter we will give some methodological details on the Molecular Dynamics (MD) scheme and, in particular, on its approach by the Langevin stochastic equation. We will phenomenologically motivate, below, the use of this equation and discuss how it can be exploited in a computational framework.

Molecular Dynamics is a numerical integration scheme in which the system is modelled as an ensemble of interacting particles under specific internal and external constraints. Given the initial conditions of the ensemble and the interactions playing a role in the system, this method integrates numerically the equation of motion (overdamped Langevin equation here), providing new sets of coordinates and velocities at each integration time step [15, 16]. From the obtained particles trajectories, one can then calculate all the global properties of the system, described in Section 2.2.

3.1 Brownian motion

In 1827, Robert Brown, a Scottish botanist, observed a phenomenon that he was not able to explain: he saw through a microscope some pollen grains randomly moving in apparently still water, just as if they were alive. Many years later, in 1905, Einstein [17] explained this phenomenon on the basis of the molecular-kinetic theory of heat, providing the link between microscopic dynamics and macroscopic observable phenomena. We know that a molecule of water has a size of about 1 Å, whereas the particles observed by Brown measured a few micrometres. Hence, Einstein postulated that the Brownian motion of a particle in a liquid is due to the instantaneous imbalance of the forces exerted by the collisions between the particle and the much smaller molecules of the solvent, which are in chaotic thermal motion.

The DNA chain immersed in a solvent behaves exactly like the pollen grains observed by Brown, chaotically diffusing in the surrounding fluid.

3.2 The stochastic Langevin equation

A few years after the treatment of the Brownian motion problem by Einstein, Paul Langevin gave his own description of the same problem [18]. He modified the standard Newton's equation of motion by adding two terms of very different nature, but both

reflecting the effect of the thermal agitation of the solvent: a deterministic damping term (the same as in Stokes' law) describing the average viscous drag effect of the thermal motion of the solvent and a random force representing the contribution of the continuous collisions of the fluid molecules with the mesoscopic particle in the reference frame in which the particle is at rest.

The one-dimensional Langevin equation for a particle of mass m in a solvent is then [19]

$$m\ddot{x}(t) = -\gamma\dot{x}(t) + G\,(t) \tag{3.1}$$

where $\dot{x}$ and $\ddot{x}$ are, respectively, the first and second derivatives of the particle position, G is the thermal random noise and γ is the viscous damping coefficient appearing in Stokes' law, which, for a bead of diameter r_0, is equal to $3\pi\eta r_0$, being η the solvent viscosity.

We want now to define the statistics of the stochastic process $G(t)$: since it considers only the random effects of the collisions in an isotropic and homogeneus space, the average over the realizations must be zero:

$$\langle G(t)\rangle = 0 \tag{3.2}$$

Since the fluid is supposed to be in a stationary state, the correlation function of the random force between t_1 and t_2, $\langle G(t_1)G(t_2)\rangle$, depends only on the difference $t_1 - t_2$ and, in general, will exponentially decrease with the growing of its argument:

$$\langle G(t_1)G(t_2)\rangle \sim e^{-\frac{|t_1-t_2|}{\tau_c}} \tag{3.3}$$

where τ_c is the correlation time of the random force.

For $|t_1 - t_2| >> \tau_c$, collision events occurring around t_1 can be considered statistically independent from the ones occurring around t_2: this assumption, called two-time-scales hypothesis, is valid if the collision correlation time τ_c is much smaller than all the other characteristic times of the problems, such as the relaxation time $\tau_R = \frac{m}{\gamma}$ of the polymer velocity v. In this case one can formally assume that:

$$\langle G(t_1)G(t_2)\rangle = A\delta(t_1 - t_2) \tag{3.4}$$

where A is a measure of the strength of the fluctuating force and δ is the Dirac delta function.

Often, for simplicity, one supposes that $G(t)$ is a Gaussian process: this hypothesis is justified by the big number of elementary collisions that are taken into account in the computation of G and that justifies the application of the central limit theorem. If $G(t)$ is a Gaussian process, then the constant A is determined by the variance σ^2 of the process:

$$\langle G(t_1)G(t_2)\rangle = \sigma^2\delta(t_1 - t_2) \tag{3.5}$$

The key-point, now, is that the macroscopic quantity γ can be related with the variance of the microscopic random process $G(t)$ by the fluctuation-dissipation theorem:

$$\sigma^2 = 2k_B T\gamma \tag{3.6}$$

This is based on the assumption that the stationary state coincides with the state of thermodynamic equilibrium and so that the theorem of equipartition of energy holds.

We can now rewrite the Langevin equation as

$$m\ddot{x}(t) = -\gamma\dot{x}(t) + \sqrt{2k_B T\gamma}\eta\,(t) \tag{3.7}$$

with the following conditions on the stochastic process η

$$\langle\eta(t)\rangle = 0 \tag{3.8}$$
$$\langle\eta(t_1)\eta(t_2)\rangle = \delta(t_1 - t_2) \tag{3.9}$$

so that its variance is $\sigma^2 = 1$.

In presence of external forces, F_{ext} we only need to add a term to equation 3.7, obtaining:

$$m\ddot{x}(t) = -\gamma\dot{x}(t) + F_{ext}(t,x) + \sqrt{2k_B T\gamma}\eta\,(t) \tag{3.10}$$

3.2.1 Ballistic and diffusive regimes

From equation 3.7 we can analytically deduce that

$$\left\langle (x(t) - x_0)^2 \right\rangle = \left(\frac{\sigma}{\gamma}\right)^2 t + \frac{m\sigma^2}{2\gamma^3}(4e^{-\gamma t/m} - 2e^{-2\gamma t/m} - 3) + \frac{m^2 v_0^2}{\gamma^2}\left(1 - e^{-\gamma t/m}\right)^2 \tag{3.11}$$

where x_0 and v_0 are the initial position and velocity, respectively.

This solution can be easily approximated in two extreme cases: at very small time scales ($t \simeq 0$) by developing its Taylor expansion we find the *ballistic regime*

$$\left\langle (x(t) - x_0)^2 \right\rangle_{t\simeq 0} \simeq (v_0 t)^2 \tag{3.12}$$

characterized by the direct proportionality between the root mean square displacement and the increase of time. In the other extreme case ($t \to \infty$) the *diffusive regime* appear, being characterized by the equation:

$$\left\langle (x(t) - x_0)^2 \right\rangle_{t\to\infty} \simeq \frac{2k_B T}{\gamma} t \tag{3.13}$$

The previous equation is more usually expressed as

$$\left\langle (x(t) - x_0)^2 \right\rangle \simeq 2Dt \tag{3.14}$$

where D is the diffusion coefficient and where we made use of the so-called Einstein's re lation

$$D = \frac{k_B T}{\gamma} \tag{3.15}$$

This result, written here in 1-dimension, can be easily extended to d dimensions by writing:

$$\left\langle (\vec{x}(t) - \vec{x}_0)^2 \right\rangle \simeq 2dDt \tag{3.16}$$

Morover, it is possible to show that the stochastic 1-dimensional variable $x(t)$, for $t \to \infty$, follows a normal distribution centred at zero, whose variance grows with time as $var = 2Dt$ giving the following probability distribution function for x at time t:

$$p(x,t)_{x0} = \left(\frac{1}{4\pi Dt}\right)^{1/2} e^{\frac{(x-x_0-v_0/\gamma)^2}{4Dt}} \tag{3.17}$$

3.2.2 The overdamped case

Equation 3.10 can be simplified if we consider a situation where the friction forces are large and damp the movement of the diffusive particle in such a way that, on the considered time scale, its acceleration becomes negligible. This regime is called *overdamped* and, under these conditions, the Langevin equation reduces to:

$$\dot{x}(t) = \frac{F(t,x)}{\gamma} + \sqrt{\frac{2k_B T}{\gamma}}\eta\,(t) \tag{3.18}$$

In this way we reduce to a stochastic differential equation of the first order in the position x that can be numerically integrated as shown in the following sections.

3.3 Dynamical properties of DNA chains

In the specific case of the dynamics of a DNA segment in a solvent, the problem is much more complex then for a single bead: the random motion of each bead is restricted by the chain connectivity and by the interactions with the other monomers, obeying each to an overdamped Langevin equation (3.18) for each coordinate. For a bead $\vec{F} = \left(\vec{F}_{int} + \vec{F}_{ext}\right)$, with $\vec{F}_{int}$ and $\vec{F}_{ext}$ being respectively the total internal and external force acting on that bead: while $\vec{F}_{int}$ is due to the intra-chain energy (as in equation 2.2), $\vec{F}_{ext}$ may come from external factors as could be the presence of an electric field or the repulsion of a confining wall, as will happen in Chapter 4.

In the following subsections we will analyse two kinds of problems: the first considers the overall motion of the polymer, mainly consisting in the diffusion of its centre of mass, the second refers to the motion of a given bead.

3.3.1 Diffusion of the centre of mass

The motion of the full chain centre of mass is the result of the sum of individual adjustments and, in bulk conditions, each of its coordinates diffuses as

$$\left\langle (X_{CM}(t) - X_{CM}(0))^2 \right\rangle = 2D^0_{CM}t \tag{3.19}$$

whose diffusion coefficient D^0_{CM} (the same for each coordinate) is related to the single bead diffusion coefficient by [14]

$$D^0_{CM} \approx \frac{D^0_{bead}}{N} \tag{3.20}$$

In particular, let us notice that, from this equation, we expect longer DNA chains to diffuse slower.

The motion of the centre of mass is commonly reffered to as g_3, defined as

$$g_3(t) = \left\langle \left(\vec{R}_{CM}(t) - \vec{R}_{CM}(0) \right)^2 \right\rangle \tag{3.21}$$

while we will call $g_{3,\parallel}$ its longitudinal component

$$g_{3,\parallel}(t) = \left\langle (X_{CM}(t) - X_{CM}(0))^2 \right\rangle \tag{3.22}$$

appearing also in equation 3.19.

In Figure 3.1 one can see, as an example, the chaotic trajectory of the centre of mass of a 50-bead chain in bulk as a function of time.

Under confinement things will become more complicated: for a polymer strongly confined in a narrow channel, the available space is no more isotropic and this asymmetry leads to a different behaviour between the confined coordinates and the longitudinal one (i.e. the one along the channel axis), which is the only one still showing a diffusive behaviour. Most of all we are interested in the possible differences in the diffusion coefficient between linear and circular chains confined inside narrow channels.

3.3.2 Single monomer motion

In addition to the overall motion of the polymer, it is interesting also to look at the motion of some internal degrees of freedom of the chain as the motion of some particular beads.

In particular the internal dynamics of the polymer motion can be suitably characterized by monitoring the following bead mean-square displacement:

$$g_1(t) = \frac{\left\langle |\vec{R}_i(t) - \vec{R}_i(0)|^2 \right\rangle}{N} \tag{3.23}$$

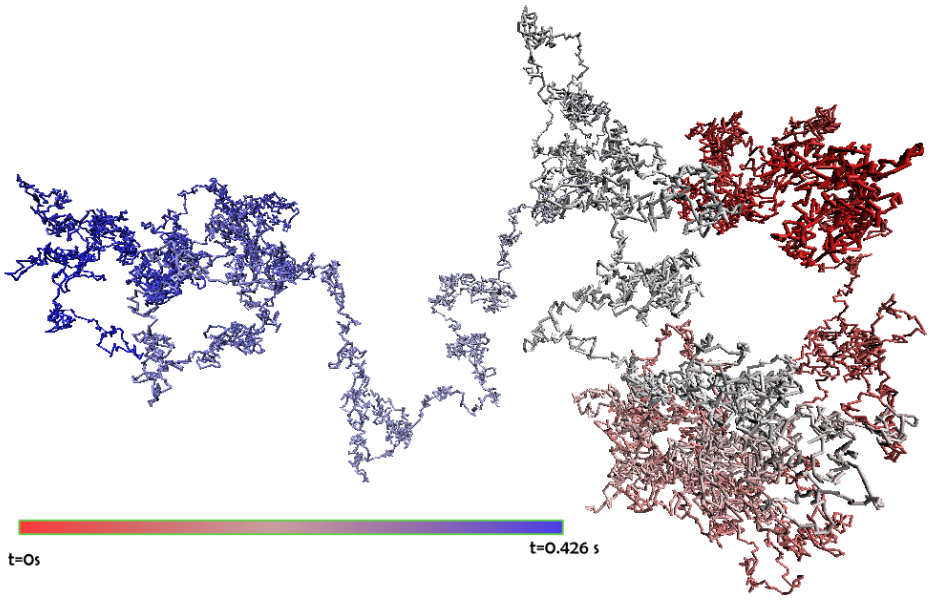


Figure 3.1: *Diffusive 3D-trajectory of the centre of mass of a chain of $N = 50$ beads in bulk as a function of time (red is initial time; blue final time)*

where i may be any given monomer in the chain and $\langle\rangle$ is the average over different realizations. In presence of a channel it is also interesting to study its parallel component

$$g_{1,\parallel}(t) = \frac{\langle |X_i(t) - X_i(0)|^2 \rangle}{N} \tag{3.24}$$

where X is the coordinate along the channel axis.

By comparing this last definition with equation 3.19 we expect that, when t is large enough

$$g_{1,\parallel}(t) \approx \frac{g_{3,\parallel}(t)}{N} \tag{3.25}$$

because the relative distance between the i-th monomer and the centre of mass will be much smaller than the distance covered in a time t by any monomer. On the contrary, for small t, a subdiffusive behaviour is expected, with a power law dependence

$$g_{1,\parallel}(t) \propto t^{\alpha} \tag{3.26}$$

where $\alpha < 1$ [20].

The crossover time between behaviour 3.25 and 3.26 is called ***Rouse time*** (τ_R).

3.4 Iteration algorithm and its implementation

For the simulations we have developed a C++ Molecular Dynamics code running on single-processor machines.

In the initial stage devoted to parameter settings, random numbers were generated using *gasdev* [21], a *C* function returning Gaussian distributed numbers with average 0 and variance 1. All simulations were initialized with the same seed, in order to enable crosschecks of different setups.

In the stage of the production of final trajectories we decided to resort to a uniformly distributed random numbers generator, that, being computationally less expensive, allowed us to speed up the simulations. This choice is justified by the fact that in a Molecular Dynamics simulation it is reasonable to use arbitrary distributions of random numbers, provided that their moments have the correct limiting behaviour [22]. In particular, simulations were run by using the Mersenne-Twister random number generator provided by Intel Math Kernel Libraries.

For the numerical integration of equation 3.18, we chose Euler's algorithm, according to which the new position of each bead is computed as:

$$x_i(t+\Delta t) = x_i(t) - \frac{\Delta t}{\gamma}\left[\vec{\nabla}_i U_{in}(\{\vec{R}_i\})\right]_x + \sqrt{\frac{2k_BT\Delta t}{\gamma}}\vec{\eta}(t) \tag{3.27}$$

where i is the bead index, and x denotes each one of its three coordinates. The square root of the time increment in the random force term is due to the fact that the considered random force is a Wiener process[23, 24].

In particular let us notice that the form of $\vec{\nabla}_i U_{in}(\{\vec{R}_i\})$ for the beads at the chain termini is quite different between linear and circular topologies because of the closure constraint, which adds one more bond and modifies also the bending energy.

3.5 Parameters

3.5.1 The Debye–Hűckel theory and the DNA effective thickness

The value of 2 nm that we gave in the previous sections for the bare width of a DNA chain considers only the steric hindrance of the molecule. On the other hand, DNA is a highly charged anionic polyelectrolyte, feeling a strong electrostatic self-repulsion that forbids bending radia even bigger than the bare diameter, coming to an increase in the excluded volume effect. The typical length scale over which the electric potential is effective strongly depends on the salt concentration of the solution: ions of positive charge accumulate around the molecule and screen its negative charge causing a reduction of the range of the electrostatic potential.

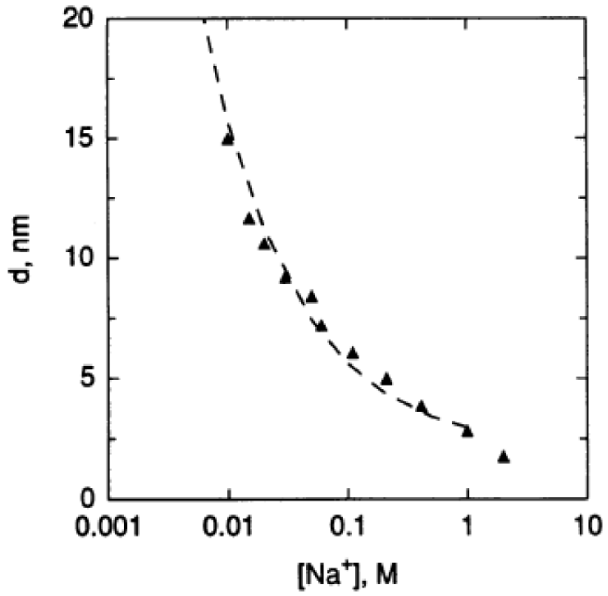


Figure 3.2: *DNA effective diameter as a function of solution Na^+ concentration adapted from ref. [26]*

The Debye-Hűckel theory [25] provides a convenient mean-field framework for estimating this effect and defines a screened electrical potential that, in the case of a point-like charge, is

$$V_{DH} = \frac{q}{4\pi\varepsilon_0 r} e^{-\frac{r}{\lambda}} \tag{3.28}$$

where r is the distance from the point-like charge, ε the permeability of the medium, q the nominal charge and λ the Debye screening length, that is the distance over which the potential begins to be strongly screened by the surrounding counterions.

The Debye-Hűckel theory can also be extended to the case of a linear distribution of charge, as in our case, giving similar results: this shows that the effective potential vanishes much quicker than the standard electrostatic one, resulting in the total screening of the charge from a certain and relatively short distance. We then introduced for the chain an effective diameter r_0 bigger than the physical one and strongly dependent on the concentration of counterions; in detail, as shown in *Figure* 3.2, at the physiological concentration of Na^+, i.e. 0.15M,

$$r_0 \sim 5nm[26] \tag{3.29}$$

which is more than twice the geometric diameter of the double helix. This estimate is in agreement with alternative ones made with very different criteria [27, 28, 29, 30].

In Figure 2.1 it is shown that average longitudinal distance between two consecutive base pairs in a dsDNA helix is 0.34 nm and so, with the present choice of the effective

diameter, each bead encompasses

$$n_{bp} = \frac{5nm}{0.34nm} \approx 15\text{base-pairs}. \tag{3.30}$$

3.5.2 Numerical values

The physical quantities of interest cannot be expressed directly in international system of units (SI) during the Molecular Dynamics simulations because their numerical values would be either very small or very large and this can lead to overflow of underflow as a result of floating-point operations. It is therefore necessary to represent all quantities in units such that their numerical values are almost unitary.

Then we chose to put all the energies in units of the thermal energy (per mole) k_BT, all the length in units of the chain effective diameter r_0 and to take γ as unitary.

We can summarize this information by saying that, in the chosen system of reference at the room temperature of $T = 300K$ and by using water as a solvent (whose viscosity is $\eta = 10^{-3}$Pa s), we have:

$$\text{Unit of energy} = \quad k_BT \quad = 0.596\text{kcal/mol} \tag{3.31}$$

$$\text{Unit of length} = \quad r_0 \quad = 5\text{nm} \tag{3.32}$$

$$\text{Unit of friction} = \quad \gamma \quad = 4.7 \cdot 10^{-11} Js/m^2 \tag{3.33}$$

The other parameters appearing in the model are expressed in derived units and their value can be found in Table 3.1; the chosen values are coherent with those already used in literature for similar studies on DNA such as [11, 31].

In particular the choice in equations 3.31 3.32 3.33 implies that the unit of time is:

$$\text{Unit of time} = \frac{\gamma r_0^2}{k_BT} = 0.284\mu s \tag{3.34}$$

After some preliminary tests used to verify the stability of the integration algorithm, we chose an integration timestep of:

$$\Delta t = 10^{-4} = 28.4ps \tag{3.35}$$

3.5.3 Autocorrelation time of the radius of gyration

It is important to notice that the choice of a Molecular Dynamics method assures both the possibility of investigating the dynamics of DNA chains by analysing its time evolution and the opportunity of testing its equilibrium properties by basing on the ergodic assumption that a long time average is equivalent to an equilibrium ensemble average.

Table 3.1: *Chosen values for the system parameters in simulation and experimental units*

Parameter	Simulation units	Experimental units
k_BT	1	0.596 kcal/mol
r_0	1	5 nm
γ	1	$4.7 \cdot 10^{-11} Js/m^2$
ε	1	0.596 kcal/mol
R_{max}	1.5	7.5 nm
k	30	0.715 $\frac{\text{kcal}}{\text{mol nm}^2}$
l_p	10	50 nm
$r_{i,i+1}^{min}$	0.9609	4.805 nm
Δt	10^{-4}	28.4 ps

Then, our prior goal is to evaluate the autocorrelation time of a characteristic observable, for example of the radius of gyration, in order to estimate after how many steps a configuration has lost its memory and can be considered independent from the previous one.

Then we can define the autocorrelation function of the radius of gyration as

$$C_{R_G}(\tau) = \langle (R_G(t) - \langle R_G \rangle)\,(R_G(t+\tau) - \langle R_G \rangle) \rangle \tag{3.36}$$

that in bulk shows the following exponential decay

$$C_{R_G}(t) \approx e^{-\tau/\tau_c} \tag{3.37}$$

where τ_c is just the autocorrelation time.

In Figure 3.3 we plot $C_{R_G}(t)$ in bulk for linear and circular chains of $N = 50$ beads (average on 30.000 independent configurations) with the superposition of the exponential fit forecast by equation 3.37. The same analysis has been performed for $N = 25$ beads and all the results are reported in Table 3.2.

	$\tau_C(\mu s)$			
	N=25		**N=50**	
	Linear	**Circular**	**Linear**	**Circular**
Bulk	8.95	0.142	84.35	11.93

Table 3.2: *R_G autocorrelation time in bulk for different topologies and number of beads*

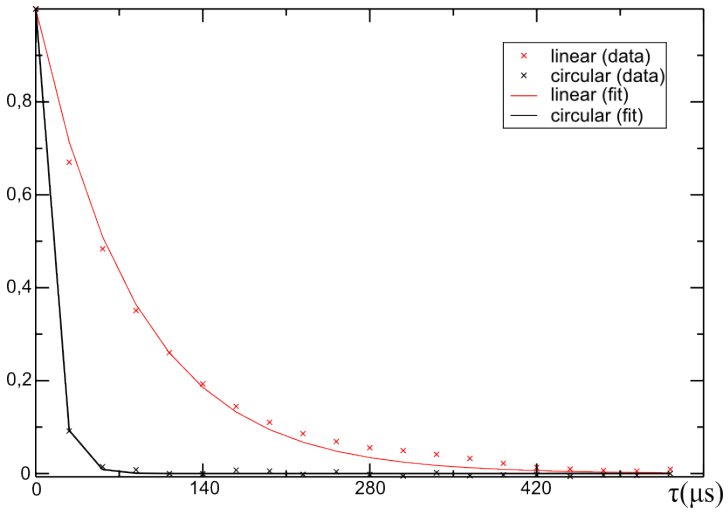


Figure 3.3: *R_G autocorrelation function for linear (red) and circular (black) chains in bulk with data (crosses) and fit (solid lines)*

In light of these results we chose as intermediate sampling time:

$$\tau_{sampling} = \begin{cases} 7.10\mu s & \text{if } N = 25 \\ 28.4\mu s & \text{if } N = 50 \end{cases} \tag{3.38}$$

which means that we record one configuration every 250000 or 1000000 time step Δt.

3.6 Validation of the code

After writing the code, we test it by analysing some quantities whose properties are known in literature.

3.6.1 Tangent-tangent correlation function

First of all let us study the tangent-tangent correlation function, C_{tt}, for a linear chain of $N = 50$ beads and, by averaging on 30.000 independent configurations, we find the set of points labelled by the blue crosses in *Figure* 3.4 as a function of the bonds distance $< b > j$.

We fit the dataset with the expected exponential behaviour described in Equation 2.11 and find an estimate of the persistence length $\tilde{l}_p^{fit} = (46.45 \pm 0.05)$ nm.

This estimate is a bit lower than the nominal value introduced in formula 2.7 and must be compared to the effective value for discrete chains, $\tilde{l}_p$, of Equation 2.12; for the simulated chain the bond length is not constant, but we can replace b with the average of the bond length over the realizations, $\langle b \rangle$, without appreciably affecting the reliability of the equation.

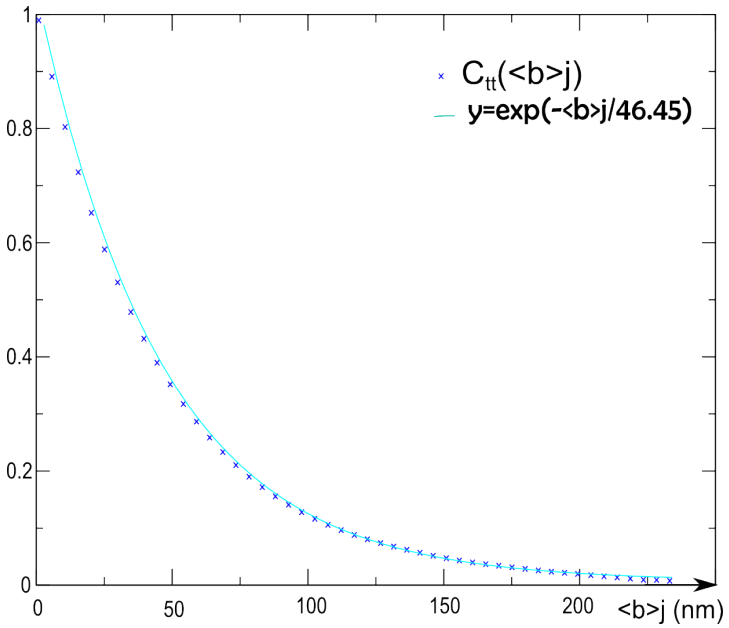


Figure 3.4: *Tangent-tangent correlation function between bonds separated by an index j for a 50-bead linear chain as a function of the bonds distance $< b > j$: data (crosses) and fit (solid line)*

By averaging on 30.000 independent configurations, we find $\langle b \rangle = (4.8548 \pm 0.0005)$ nm which is bigger than the value of minimal energy found in Section 2.1.1 at $r_{i,i+1} = 4.8030$ nm, as expected from the asymmetry in the bond length distribution.

By substituting $\langle b \rangle$ and the nominal value of l_p into Equation 2.12, we obtain an expected value for the effective persistence length of $\tilde{l}_p = 46.05$ nm.

We notice that $\tilde{l}_p^{fit}$ and $\tilde{l}_p$ are in good agreement and the slight excess of $\tilde{l}_p^{fit}$ may come, besides to statistical reasons, from the self-avoidance of the chain (which, in the simulations, is not ideal as in the Kratky-Porod model), whose excluded volume effect tends to make the chain stay stiffer, forbidding it to turn back on itself. Anyway, a 50-bead chain is only about five l_p long and so the excluded volume does not affect the results too much, because too compact configurations will already be strongly unfavoured by the high bending energy penalty: the excluded volume effect becomes much more important for larger contour lengths.

3.6.2 Chain dimensions

Another check that we can do to validate the code concerns $\langle R_{ee}^2 \rangle$: by inserting the corrected value of the persistence length for discrete chains, $\tilde{l}_p$, and the value of the contour length $L_c = N \cdot \langle b \rangle$ in equation 2.23 we obtain, for a 50-bead ideal Kratky-Porod chain, an expectation value of $\langle R_{ee}^2 \rangle \approx 1.81 \cdot 10^{-14} m^2$.

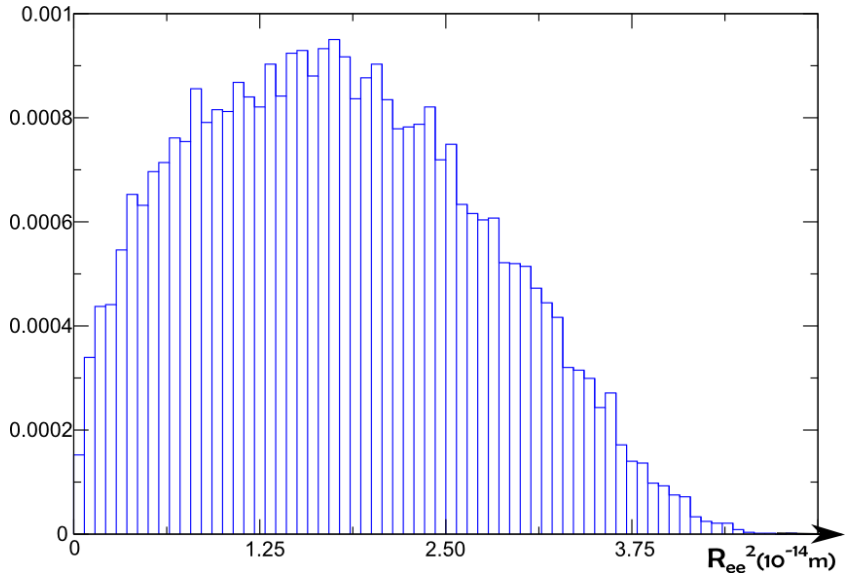


Figure 3.5: *Distribution of the R^2_{ee} of a 50-bead linear chain from 30.000 uncorrelated configurations*

On the other hand, from our data, the average on 30.000 uncorrelated configurations gives $\langle R^2_{ee,data} \rangle = 1.79 \cdot 10^{-14} m^2$.

To understand if our estimate is viable, we can analyse the distribution of R^2_{ee} in the considered data sampling, shown in *Figure* 3.5. The normalized histogram shows a very wide distribution that seems to confirm that our estimate of $\langle R^2_{ee,data} \rangle$ is compatible with the theoretical value for Kratky-Porod chains; once more we see that, for so short chains, the excluded volume effect has almost no concequences in bulk conditions.

3.6.3 Motion of the centre of mass

We can take advantage of the dependence of the motion of the centre of mass only on the stochastic forces (let us remember that the sum of internal forces in a system is always zero) and use it as a crosscheck in order to test the code: let us run two simulations with the same initial seed for the random number generator and develop the motion of two chains, one linear and one circular, of fixed number of beads $N = 50$. By fixing the seed we put ourselves in a deterministic framework because the stochastic force felt by any bead along any direction at time t is identical in both simulations and then we expect the motion of the centre of mass in the two cases to be exactly the same (the closure constraint is still an internal force).

As shown in Figure 3.6, the motion of the centre of mass is the same for both cases, as expected.

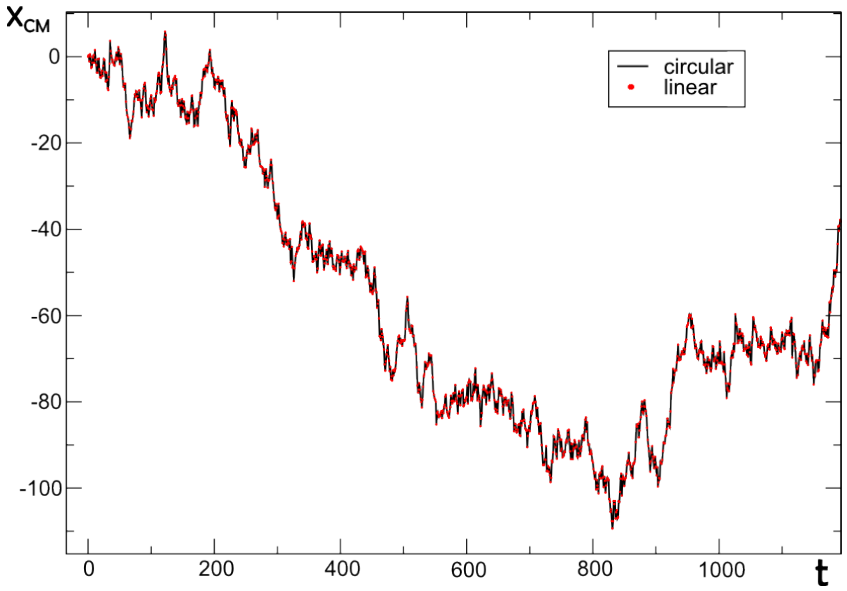


Figure 3.6: *Motion of the centre of mass of a linear (red points) and circular (straight black line) chain of $N = 50$ beads along x direction from simulations initialized with the same seed for the random number generator.*

Chapter 4

Confinement in a nanochannel

Previous studies have shown how metric and equilibrium properties of polymers are affected by the dimensionality and size of the confining region (see [3, 6] and references therein). Inspired by the recent interest in using nano-channels to manipulate and sort biomolecules, here we shall report on a numerical investigation of the equilibrium and kinetic properties of confined DNA chains. The study is carried out using a stochastic molecular dynamics scheme, with no treatment of hydrodynamic interactions. The latter choice is motivated by the study of Tessier et al. [5], who resolved the different electrophoretic mobility of channel-confined linear and circular chains by using stochastic simulations with no hydrodynamic effects. Based on this findings we choose to used the simplest, and hence most transparent, Langevin simulation scheme to characterize and compare the kinetic (and equilibrium) properties of the confined chains.

We point out that confinement in nanochannels is very common in biological systems and occurs, for example, during the translocation of DNA through a membrane channel. Nowadays it can also be easily reproduced in biotechnological experiments: chief examples are the long square silicon micro and nano-channels designed to sort DNA by contour length [1, 2].

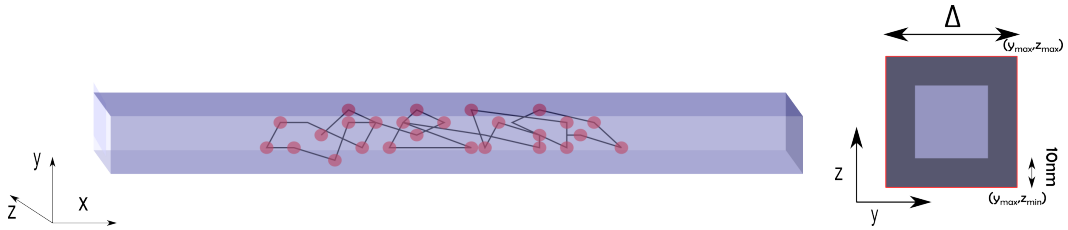
4.1 Modelling DNA confinement in a nanochannel

As shown in *Figure* 4.1a, we will consider a square channel of infinite length along its axis (identified with $\hat{x}$ axis). The transverse size along the $\hat{y}$ and $\hat{z}$ directions, indicated with Δ, will be varied in the range 32.7-65.4 nm.

We will model the channel as if it were composed of four flat surfaces, all endowed with a repulsive Lennard-Jones potential. The excluded volume interaction between the chain beads and the channel walls is therefore described as:

$$\begin{aligned} U_{channel} = \; & 4\varepsilon \sum_{s=y,z} \sum_{i=0}^{N-1} \{ \left[\left(\tfrac{r_0}{s_{max}-s_i} \right)^{12} - \left(\tfrac{r_0}{s_{max}-s_i} \right)^{6} + \tfrac{1}{4} \right] \Theta(2^{\frac{1}{6}} r_0 - (s_{max} - s_i)) \\ & + \left[\left(\tfrac{r_0}{s_i - s_{min}} \right)^{12} - \left(\tfrac{r_0}{s_i - s_{min}} \right)^{6} + \tfrac{1}{4} \right] \Theta(2^{\frac{1}{6}} r_0 - (s_i - s_{min})) \} \end{aligned} \tag{4.1}$$

where ε is the same as in Section 3.5 and $y_{max} = z_{max} = -y_{min} = -z_{min} = \Delta/2$ represent the coordinates of the four confining walls as shown in Figure 4.1b.



(a) 3D image **(b)** Section

Figure 4.1: *3-Dimensional image and section of the channel*

The presence of the channel introduces an external force field along the $\hat{y}$ and $\hat{z}$ axes, so that the motion of the centre of mass along the confined directions depends also on these external forces. On the other hand, the polymer diffusion along the channel axis is ruled only by the sum of the stochastic forces, because, in absence of hydrodinamic effects, the repulsive forces exerted by the channel walls have no components along $\hat{x}$ and then, if we want to see a difference between linear and circular topologies in the diffusive behaviour of the chain, we need to conceive a more realistic channel as we will do in the following sections.

4.2 Modelling the channel roughness

We therefore modelled the channel by taking into account also the intrinsic roughness of its walls, which has not been analysed so often in literature, yet, but is ubiquitous in nature. It is well accepted that the roughness can dramatically affect electroosmotic flows in nanochannels [32, 33, 34], so it is natural to expect that it may cause some changes also in the motion of confined chains, at least when hydrodynamics is taken into account.

The roughness tends to keep the chain trapped in the proximities of the walls once it gets close enough [32] and, in order to mimic this effect, we ideally divide the channel into a core and an external layer (Figure 4.1b). This latter is characterized by an effective damping coefficient $\tilde{\gamma} = 2\gamma$ so that (as can be easily seen by looking at the Langevin overdamped equation 3.18) in this region, the beads experience higher friction.

Since for practical reasons it is presently very hard to produce nanochannels with roughness lower than some nm we will therefore fix the amplitude of the external layer at the value 10 nm.

4.3 Chain dimensions for different channel size

The equilibrium configurational properties of DNA chains confined in a channel are controlled by the competition between different length scales: the radius of gyration of

the chain in bulk (R_G^0), the channel calliper size (Δ), the chain persistence length (l_p) and its contour length (L_c).

When $\Delta >> \langle R_G^0 \rangle$ the average longitudinal size ($\langle R_{G,\|} \rangle$) and the average extensions along the two constrained directions ($\langle R_{G,\perp} \rangle$) are all equal to $R_G/\sqrt{3}$, but when $\Delta << \langle R_G \rangle$, $\langle R_{G,\perp} \rangle$ reduces to $\approx \Delta$ and $\langle R_{G,\|} \rangle$ is expected to increase as Δ decreases [3].

In particular, the metric properties of chains subject to moderate confinement are well described by the ***de Gennes' blob theory***[13], which provides the scaling law for the longitudinal size ($\langle R_{G,\|} \rangle$) as a function of the channel width, Δ. In this theory moderately confined chains are described by a 1-dimensional succession of "blobs" having the same metric properties as the self-avoiding chains in bulk, but with a diameter given by the channel width, Δ.

On the contrary, when the channel size is much smaller than the DNA l_p, this picture does not work any more and the physics is dominated by the interplay between the geometrical confinement and the intrinsic elasticity and topology of the chain, as successfully described by ***Odijk's deflection theory*** [35].

Moreover, when $\Delta \approx l_p$ a transition regime appears with still different properties, recently theorized by Odijk himself [7].

In this work we will consider three different sizes for the channel:

- the narrowest channel width, Δ_1, is the radius of gyration of the circular chain of $N = 50$ beads calculated in bulk conditions, that is $\Delta_1 = \tilde{\Delta} = 32.7$ nm;
- the second channel has width $\Delta_2 = 1.5\tilde{\Delta} = 49.1$ nm;
- the third channel has width $\Delta_3 = 2\tilde{\Delta} = 65.4$ nm.

4.4 Channel confinement and $R_{G,\|}$ autocorrelation time

In order to estimate the time interval between two statistically independent configurations and so to have a reliable control on the equilibrium and kinetic properties of our system, we need to estimate the $R_{G\|}$ autocorrelation time for confined chains.

In this case the autocorrelation of $R_{G,\|}$ does no more present a trivial exponential decay: under confinement the autocorrelation function for a linear chain is given by the superposition of two exponentially decreasing functions, the fastest depending on the autocorrelation time, the slowest accounting for the autocorrelation of special configurations happening when the chain folds back on itself, as in Figure 4.2, causing a drastic decrease of its $R_{G,\|}$. For a circular chain this second component is strongly damped because of the closure constraint that makes every monomer equivalent to the others, while for linear chains both exponential decays can be clearly identified.

The $R_{G,\|}$ autocorrelation function for a chain of $N = 50$ beads (data and fit for both linear and circular topology) in a channel of width $\Delta = 32.7$ nm is shown in Figure 4.3

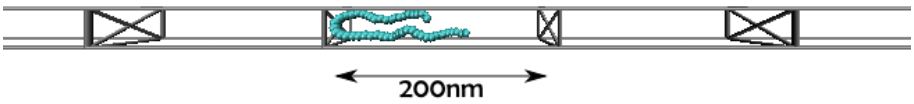


Figure 4.2: *Backfold of a linear chain of $N = 50$ beads in a channel of width $\Delta = 32.7$ nm.*

and all the estimated autocorrelation times for the considered channel widths and number of beads are reported in Table 4.1. In particular let us notice that, for any analised degree of confinement and at fixed N, the autocorrelation time is larger for linear chains than for circular ones in agreement with what found also by other studies [36].

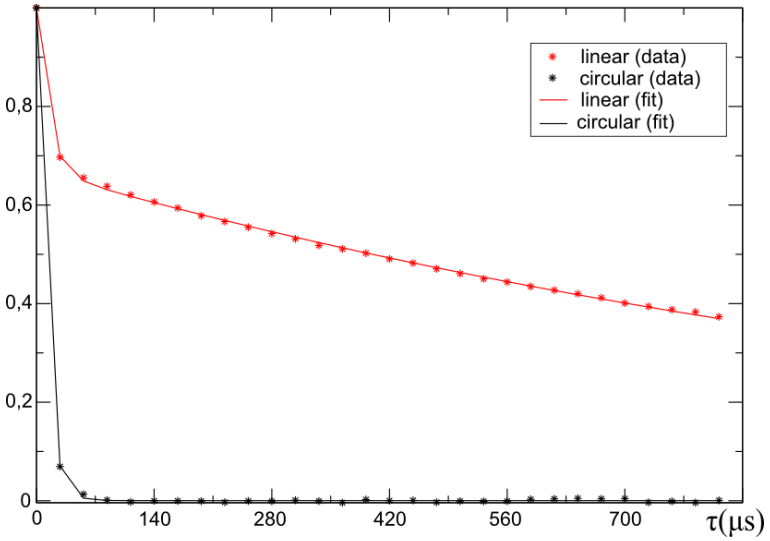


Figure 4.3: *$R_{G\|}$ autocorrelation function for linear and circular chains in a square channel of width $\Delta = 32.7$ nm: data (point) and fit (solid line)*

$\tau_c(\mu s)$				
Channel width	**N=25**		**N=50**	
	Linear	**Circular**	**Linear**	**Circular**
2Δ	10.37	3.62	58.50	28.97
1.5Δ	7.74	2.77	50.55	19.03
Δ	4.26	1.49	13.92	10.71

Table 4.1: *$R_{G,\parallel}$ autocorrelation time for different topologies, channel widths and number of beads*

Chapter 5

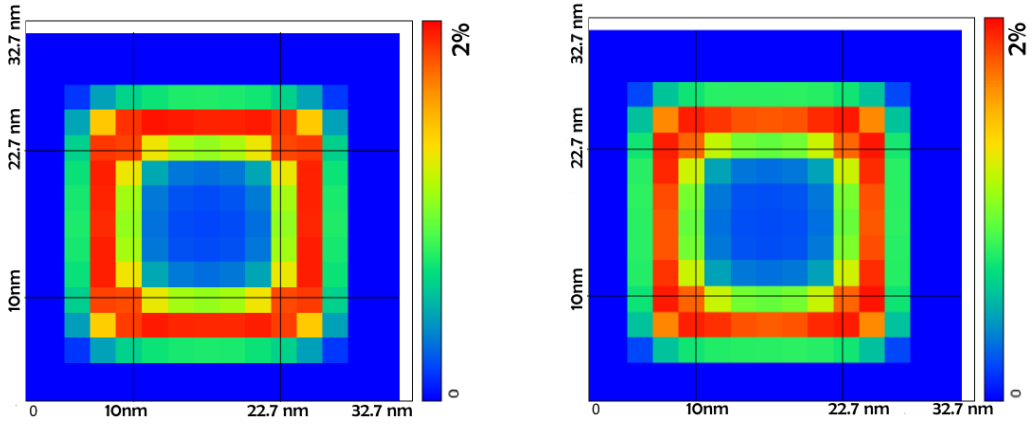
Results and discussion

5.1 Equilibrium properties

In this section we will present and discuss our results concerning the equilibrium properties of confined DNA chains.

5.1.1 Transverse distribution of the chain monomers

We first considered the transverse distribution of the chain monomers in the channel for both linear and circular chains (Figure 5.1 shows the case with $N = 50$ and $\Delta = 32.7$ nm): they are very similar to each other and, as expected, both have a concentric square shell structure. The most external layer is the least populated, because of the channel



(a) Linear chain **(b)** Circular chain

Figure 5.1: *Transverse monomer distribution for linear and circular chains of $N = 50$ beads in a square channel of width $\Delta = 32.7$ nm computed on 68.000 uncorrelated configurations.*

walls repulsion; the middle one has the highest concentration of monomers, because of the high value of the effective friction $\tilde{\gamma}$ and the innermost part has an intermediate

concentration of monomers. The same analysis has been repeated for all the considered channel widths and chain contour lengths, obtaining density profiles qualitatively similar to fig. 5.1, without showing appreciable differences due to topology.

5.1.2 Tangent-tangent correlation function

Another interesting quantity to discuss is the tangent-tangent correlation function, C_{tt}, for linear chains at different degrees of channel confinement. From fig. 5.2 we can see that, while C_{tt} in bulk approximately follows an exponential decay depending on the persistence length (see Section 3.6.1), for increasing confinement the correlation along the polymer do not decay exponentially any more but levels off at a constant non-zero value. This is due to the limitation of the possible orientations caused by the geometrical constraint. For strong confinement, in fact, the polymer is only allowed to roughly align with the channel axis, giving rise to constant non-zero values for the tangent-tangent correlation also between distant monomers. As already established in literature [37], we find a short initial range where the correlation function decays as in the unconstrained case, before being affected by the presence of the channel and reaching an approximately constant value.

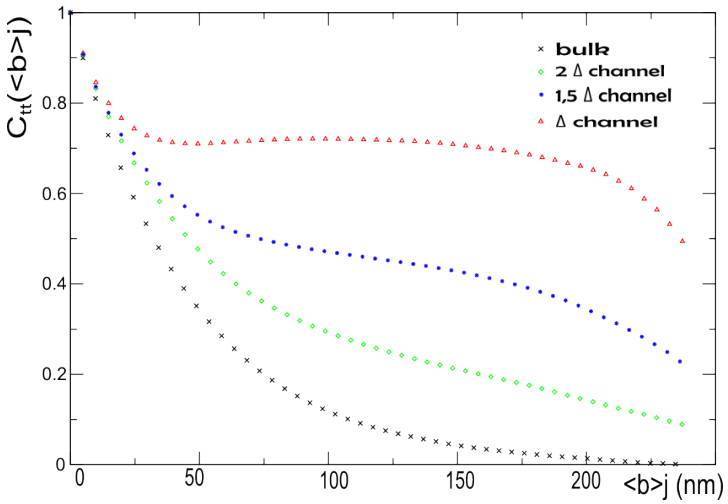


Figure 5.2: *Tangent-tangent correlation function between bonds separated by an index j for linear chains of $N = 50$ beads at different degrees of confinement*

We notice that, for the considered strongest confinement, there is a further decrease of C_{tt} concerning the monomers at the largest separation along the chain. This may be due to the fact that in a narrow channel the chain termini have more conformational freedom (and hence directional disorder) than the rest of the chain.

In order to better understand this effect, we optimally aligned 10000 independent configurations and obtained the envelope shown in Figure 5.3: both extremities are characterized by an enhancement of the envelope size, showing the expected disorder.

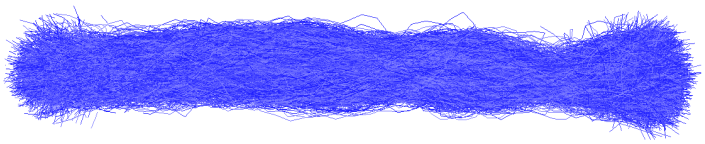


Figure 5.3: *Superposition of 10000 optimally aligned independent configurations of 50-bead linear chains in a channel with* $\Delta = 32.7$ *nm*

5.1.3 Average chain size under confinement

We now discuss the scaling properties of the average chain size in channels of different Δ. Figure 5.4 shows the scaling of $\langle R_G \rangle$, $\langle R_{G,\parallel} \rangle$ and $\langle R_{G,\perp} \rangle$ for linear and circular chains of 25 and 50 beads, as a function of the inverse of the channel width Δ. As expected from section 4.3 and in agreement with what found by prior experimental and computational studies [1, 38], $\langle R_{G,\perp} \rangle$ decreases for higher confinement, while $\langle R_{G,\parallel} \rangle$ progressively increases.

Concerning $\langle R_G \rangle$, only the linear chain of 50 beads (that is the one with biggest size in bulk $\langle R_G^0 \rangle$) shows a significant modulation with the channel size, while the ones with smaller bulk size are characterized by an almost constant average radius of gyration. We finally notice that, also with respect to $\langle R_{G,\parallel} \rangle$ and $\langle R_{G,\perp} \rangle$, the linear chain of 50 beads is the configuration most affected by confinement.

In order to further clarify what happens for the different topologies and lengths, one can look at the relative variation of the chain size with respect to their bulk values $\langle R_{G*}^0 \rangle$ as functions of the inverse of Δ (Figure 5.5). From these plots one can clearly see that the linear configurations are more affected by the presence of the channel than their circular counterparts and that, for fixed topology, longer chains are more influenced than shorter ones.

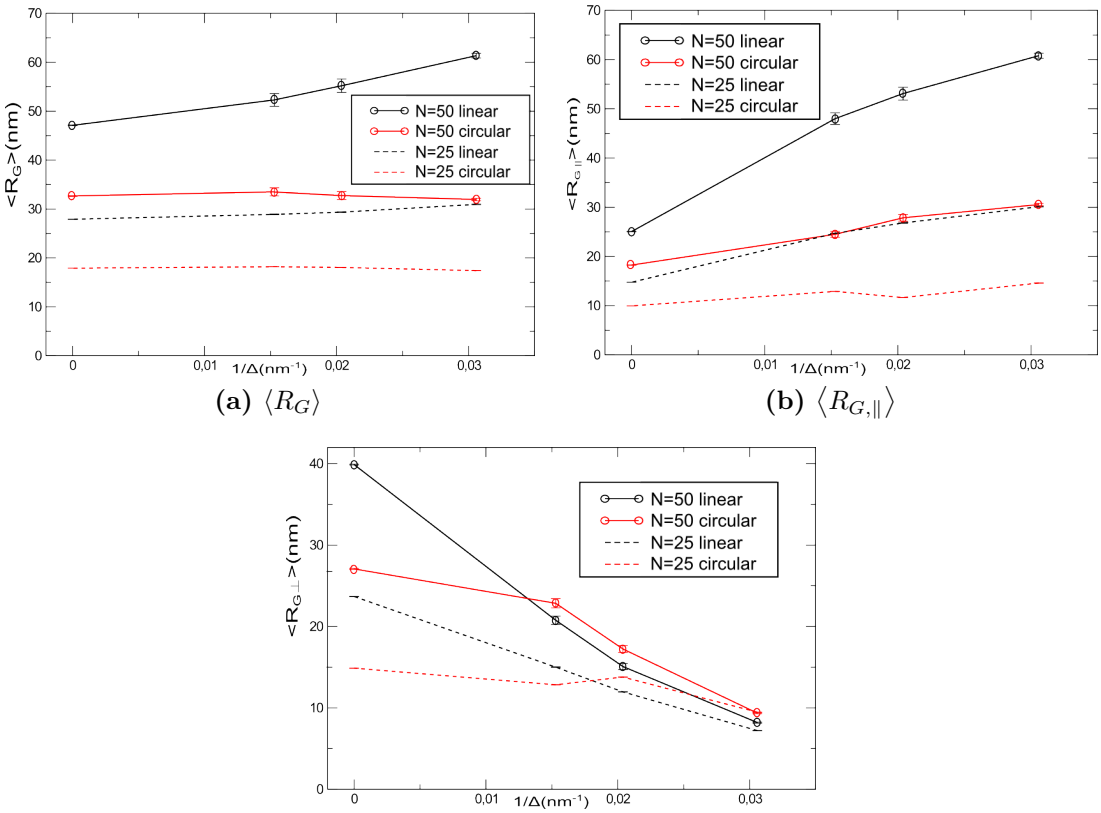
5.2 Dynamical properties

In this section we discuss the single monomer motion and the diffusion of the centre of mass for chains of different topologies and contour lengths within channels.

5.2.1 Single monomer motion

As said in Section 3.3.2, the motion of a single monomer in the chain can give important information about the chain dynamics.

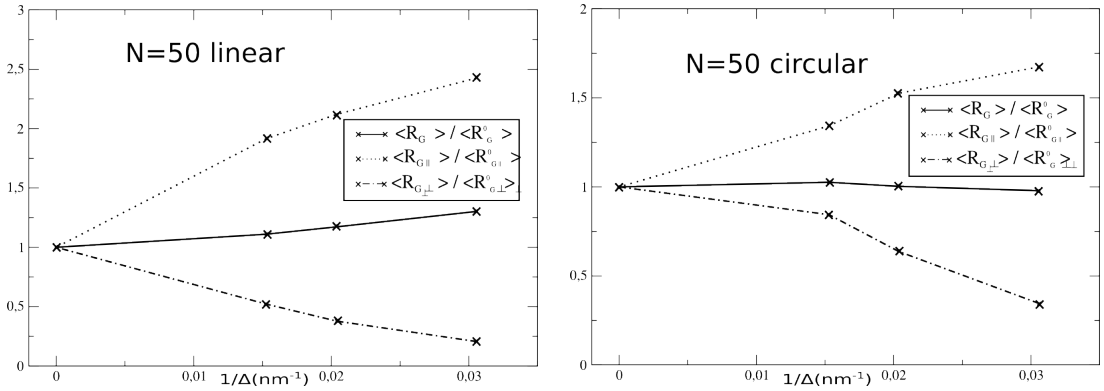
We will therefore analyse the behaviour of $g_1(t)$ for different contour lengths, topologies and channel widths. For closed chains we consider the monomer labelled in the simulations with number 0, but this choice is totally unimportant, because each monomer



(a) $\langle R_G \rangle$ **(b)** $\langle R_{G,\parallel} \rangle$

(c) $\langle R_{G,\perp} \rangle$

Figure 5.4: *Chain size scaling for different degrees of confinement as a function of the inverse of the channel width* Δ*. The ensemble averages are calculated on a large number of independent configurations (the precise number varies from case to case between 45000 and 150000) and the error bars have been associated by considering the semidispersion (i. e. the semidifference between maximal and minimal values) of 4 uncorrelated groups of configurations containing at least 12000 configurations each.*



(a) Linear chains of 50 beads

(b) Circular chains of 50 beads

(c) Linear chains of 25 beads

(d) Circular chains of 25 beads

Figure 5.5: *Relative variation of* $\langle R_G \rangle$, $\langle R_{G,\parallel} \rangle$ *and* $\langle R_{G,\perp} \rangle$ *with respect to the bulk values* $\langle R^0_{G*} \rangle$ *as a function of the inverse of the channel width* Δ *for different topologies and contour lengths*

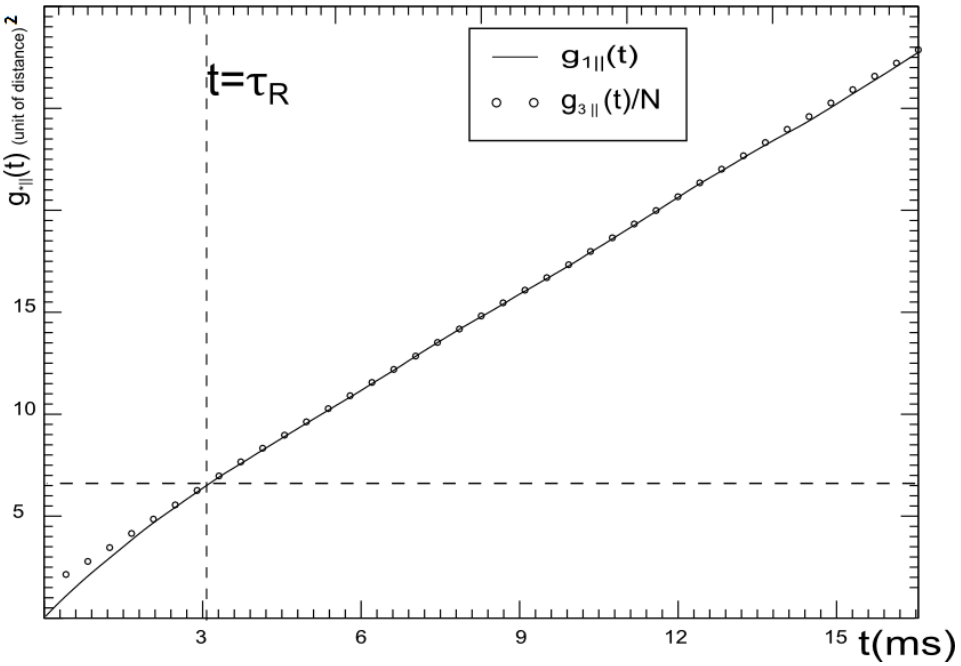


Figure 5.6: *Determination of the Rouse time, τ_R, through the comparison between $g_{1,\|}(t)$ and a vertical translation of $g_{3,\|}(t)/N$ for a closed chain of $N = 50$ beads in a channel of width $\Delta = 32.7$ nm. The monomer taken into account for the determination of $g_{1,\|}(t)$ is the one labelled with 0 in the simulations.*

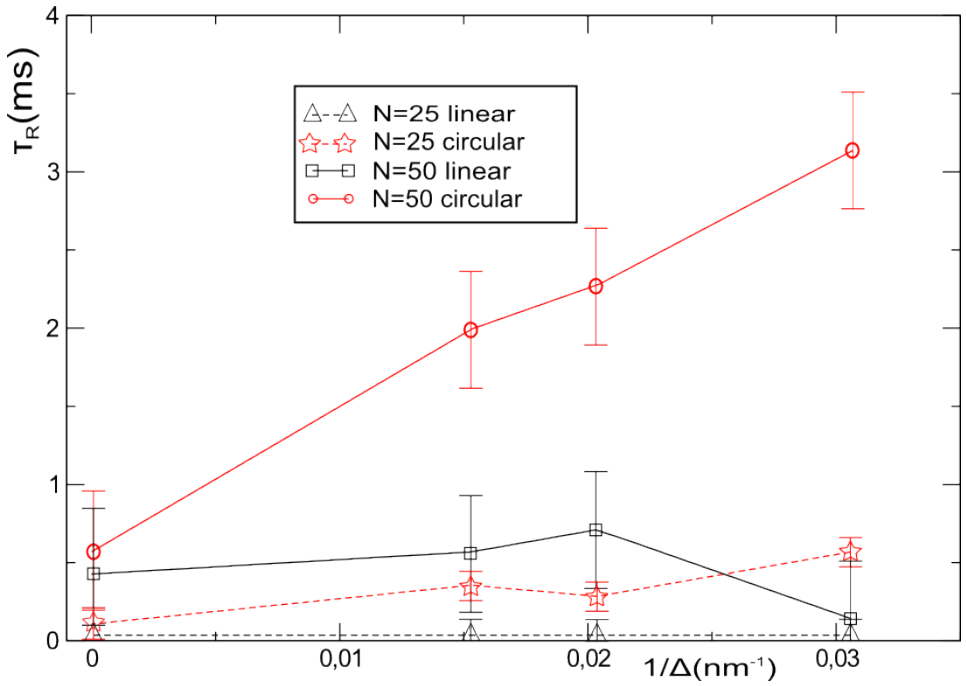


Figure 5.7: *The Rouse time, τ_R, for chains of $N = 25$ and $N = 50$ monomers as a function of the inverse of the channel width, Δ. The error values are equal to the time interval between two subsequent points in figure 5.6, because the time resolution can not exceed this limit.*

is topologically equivalent to the others and has, on average, the same distance from the centre of mass. For linear chains, on the contrary, different choices of the monomer will lead to different values for the crossover time between the short-time and long-time behaviours of $g_{1,\|}$, depending on the distance of the considered monomer from the chain midpoint. The Rouse (or relaxation) time for linear chains can be defined [11] by considering $g_{1,\|}(t)$ for the innermost monomer of the chain and analysing its evolution for growing confinement.

In figure 5.6 one can find the procedure used to determine τ_R, shown, as an example, for a circular chain of $N = 50$ beads in a channel of size $\Delta = 32.7$ nm: $g_{3,\|}(t)/N$ is superimposed on $g_{1,\|}(t)$ and vertically translated in order to have the better possible overlap between the two curves for time large enough. In this way it is easier to distinguish when the two curves have the same slope: τ_R is then the threshold time at which the two curves start to superimpose on each other.

Finally, figure 5.7 shows the variation of τ_R for different degrees of confinement: for all the analysed chain lengths and topologies, τ_R increases almost monotonically with decreasing Δ, in agreement with what found in some experimental studies [1]. In particular, for circular chains the Rouse time increases as Δ decreases with a higher rate than their linear counterparts.

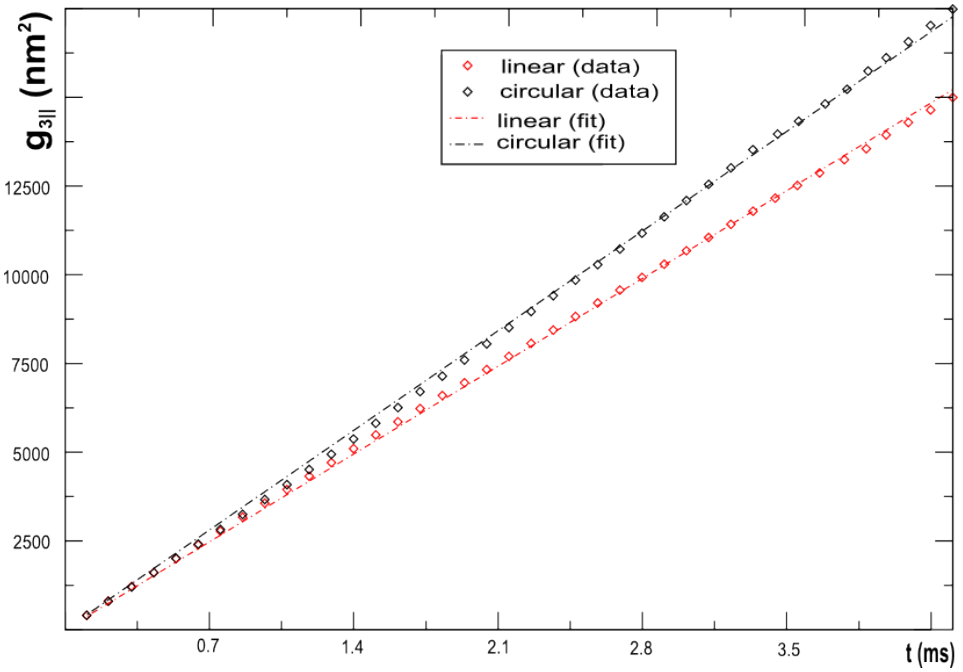


Figure 5.8: *$g_{3,\parallel}(t)$ for linear and circular chains of 25 beads and respective linear regression obeying equation $g_{3,\parallel}(t) = 2Dt$ in a channel of width $\Delta = 32.7$ nm.*

5.2.2 Diffusion of the centre of mass

The estimates of the diffusion coefficients are obtained by evolving for a very long time a single initial configuration and then extracting at regular time intervals $i \cdot \tau$ (with $i = 1, 2, ...number\ of\ points\ in\ the\ plot$) the quantity $g_{3,\parallel}(i \cdot \tau)$, defined in section 3.3.1, which is expected to obey equation 3.19. The linear regression on the obtained points as a function of time provides then an estimate for the parameter D, (an example is shown in Figure 5.8).

Figure 5.9 and Table 5.1 show the diffusion coefficients for linear and circular chains of 25 and 50 beads. The first thing to notice is that progressive confinement reduces the diffusion coefficient for all the considered contour lengths and topologies. We also point out that, for all the considered topologies and degrees of confinement, the diffusion coefficient for the chain of 50 beads is numerically compatible with being 1/2 of the one for the 25-bead chain. This result is expected in bulk when hydrodynamic interactions are neglected (as it is in our case) because of equation 3.20. In particular, the bulk

diffusion coefficients D^0 (first line of Table 5.1) are compatible with the theoretical values

$$D^0_{25beads} = \frac{k_B T}{25\gamma} = 0.352 \mathring{A}^2/ns \tag{5.1}$$

$$D^0_{50beads} = \frac{k_B T}{50\gamma} = 0.176 \mathring{A}^2/ns \tag{5.2}$$

Notice that the scaling relation

$$D \propto N^{-1} \tag{5.3}$$

is strictly valid for ideal chains only, but in our model it can be generalized also to self-avoiding semi-flexible chains in bulk, because in unconstrained conditions the excluded volume potential and the bending energy, being external forces, cancel in the reference frame of the centre of mass. This result is also in agreement with what found in Monte Carlo simulations for diluted polymers in bulk by Muller et al. [39].

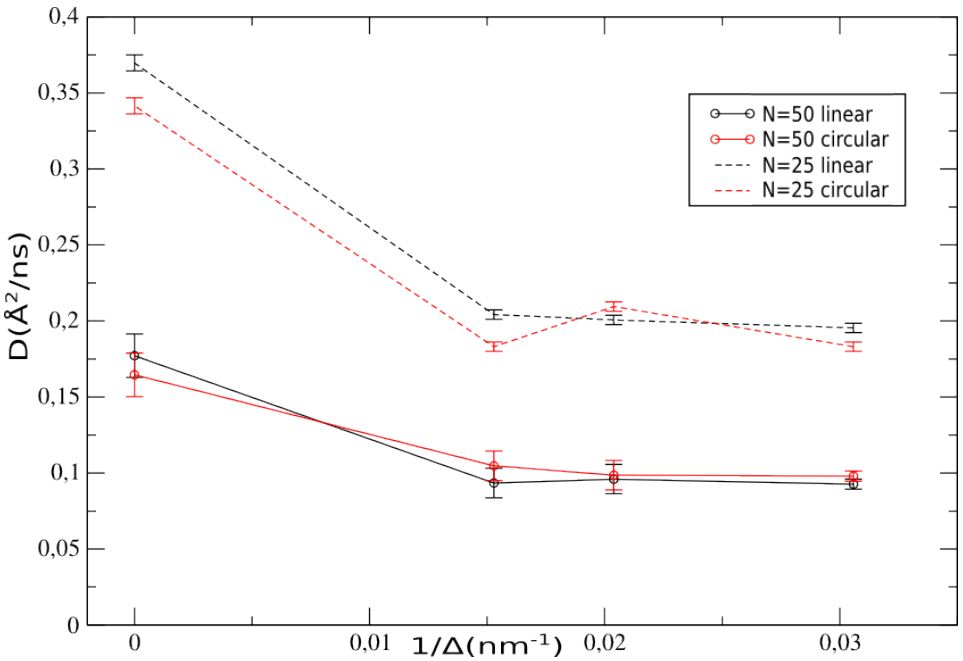


Figure 5.9: *Diffusion coefficient D for different contour lengths and topologies as a function of the inverse of the channel width Δ. The error bars have been estimated by dividing the set of configurations into four uncorrelated subsets and taking the semidispersion (i. e. the semidifference between maximal and minimal values) of the four values obtained for D in the four subsets.*

Under confinement, by contrast, the internal forces interplay with the roughness of the channel and equation 5.3 is not a priori guaranteed to apply. It is therefore most interesting that it continues to hold to a very good extent.

$D(\mathring{A}^2/ns)$				
Channel width	**N=25**		**N=50**	
	Linear	**Circular**	**Linear**	**Circular**
Bulk	0.370 ± 0.005	0.342 ± 0.005	0.18 ± 0.01	0.165 ± 0.001
2Δ	0.204 ± 0.002	0.183 ± 0.002	0.093 ± 0.002	0.105 ± 0.002
1.5Δ	0.201 ± 0.003	0.210 ± 0.002	0.10 ± 0.01	0.0986 ± 0.0004
Δ	0.1954 ± 0.0004	0.1831 ± 0.0004	0.0926 ± 0.0007	0.098 ± 0.003

Table 5.1: *Diffusion coefficients for different topologies, channel dimensions and contour lengths*

For what concerns topology, with the present data we do not observe appreciable difference in the diffusion coefficient, because, at fixed number of beads and channel width, the values of D for linear and circular chains are generally consistent and the little differences can be attributed to statistical noise. For example let us notice that the noisiest case is the 25-bead chain in bulk, whose values are identical, as explained some lines above.

So, with the present model, linear and circular chains exhibit very similar diffusive properties in the reference system of the centre of mass. This fact is not so realistic because, at fixed N, the different topologies implies a very different average size and it has been shown from several prior theoretical and experimental studies [14] that this feature plays a fundamental role in the overall motion of the polymer. A possible reason that may explain why our model can not capture the consequences of topology on the overall motion of the polymer is that they may emerge only when hydrodynamics plays an important role and is consequently taken into account or when other factors introduce some perturbations in the spatial symmetry of the problem (for example the non-homogeneous electric field in the electrophoresis Monte Carlo simulation of Tessier et al [5]).

On the other hand, we can neither completely exclude the presence of a topological effect in our model, because of the finite precision of our estimate of D; a more precise information may be attained by adding more statistics, but, in the light of the data obtained up to now, the possible topological effect that might be found would be much sligther than the difference between the diffusion coefficients of the 50-bead and the 25-bead chains.

Then, it is also possible that the topological differences are more important for long chains than for short ones (because of technical reasons experimental data generally concern longer filaments than the ones simulated in this work), so it could be interesting, in the future, to push the simulations to longest chains, even if it is computationally much more expensive (let us remember, for example, that the updating of LJ potential

needs a computational time proportional to the square of the number of beads N and that the autocorrelation time of the chain R_G scales as N^2, too).

Chapter 6

Conclusions and perspectives

In this work we presented a numerical study of the behaviour of linear and circular short DNA chains (125-250 nm) when confined inside nanochannels of width comparable with their size in bulk.

We modelled the DNA filaments as coarse-grained semi-flexible self-avoiding chains of beads, put them in square nanochannels (32.7-65.4 nm) and let them diffuse. The equilibrium and kinetic properties of the confined chains were studied by using Molecular Dynamics simulations with an overdamped Langevin approach. The channel roughness (≈ 10 nm) is taken into account by enhancing the effective friction acting on the chain beads, while the channel walls exert on the beads a short-range repulsive potential.

We first studied the scaling of the metric properties of open and closed chains as a function of their contour length and of the width of the confining region. In particular we considered the scaling properties of the chain average size and of its average longitudinal size (while the transverse one is clearly dictated by the channel confinement). Consistently with what established in previous experimental and numerical studies on chains of different length [3, 6], we observed that the above mentioned scaling properties differ significantly between open and closed chains at all considered levels of confinement. In both cases, however, upon reduction of the channel width, it is seen that the average radius of gyration remains fairly constant, while the transverse and longitudinal size vary considerably. This property has no counterpart with what observed for longer chains subject to channel confinement.

We completed the study of the equilibrium properties of the system by analysing the tangent-tangent correlation of confined chains. Besides observing the expected major qualitative differences between open and closed chains, the analysis aptly complemented the above-mentioned metric investigation by illustrating the increase of chain directional order as a function of progresisve confinement, as predicted on theoretical grounds [7].

Furthermore, in a rough channel, the chain monomers spread with higher probability in the proximities of the walls than in the very central part of the channel, being trapped by the channel intrinsic roughness that slows their motion down.

We next moved to consider the impact of chain topology, contour length and confinement on the kinetic properties. Such investigation is motivated by perspective applications in applicative contexts (nano-fluidic passive sorting of dsDNA). We first studied the motion of single monomers in the chain and extracted the Rouse time for linear and circular chains as a function of confinement. In agreement with some experimental studies, we

found out that the Rouse time, which at constant contour length is bigger for circular than for linear chains, increases as the channel size decreases for all the considered topologies and chain lengths.

Then, basing on analogies with previous stochastic simulations of elecrophoretic mobility of confined chains, we expected to observe different diffusion coefficients for linear and circular confined chains. At variance with this expectation, no prominent differences were observed in the diffusive kinetics of open and closed chains (while significant differences were observed as a function of chain length). Because dsDNA molecules of different topology are known to have different diffusivity properties in bulk, we must conclude that the employed numerical simulation setup, which represents the standard starting point of stochastic simulations, is not adequate to expose differences in the diffusivity of molecules of different length. For future developments it will therefore become necessary to include hydrodynamics in the model by approaching the problem, for example, with an extended Brownian Dynamics simulation method incorporating the solvent flow, as done by Jendrejack et al. for linear chains [40, 41] or by considering the fluid as made up of particles and coarse-grain the system onto the cells of a regular lattice as done by Malevanet and Kapral in [42].

Furthermore, it could be interesting to study the diffusive behaviour of longer chains and to analyse DNA molecules in melt, as did Robertson and Douglas in [43], in order to find results easier to test by real experiments. It could be of great interest also the analysis of the behaviour of more complex topologies, such as simple knots, but we must consider that all of this would need a higher amount of computational power.

Bibliography

[1] W. Reisner, K. J. Morton, R. Riehn, Y. M. Wang, Z. Yu, M. Rosen, J. C. Sturm, S. Y. Chou, E. Frey, R. H. Austin. Statics and dynamics of single dna molecules confined in nanochannels. *Phys. Rev. Lett.*, 94:196101, May 2005.

[2] J. O. Tegenfeldt, C. Prinz, H. Cao, S. Chou, W. W. Reisner, R. Riehn, Y. M. Wang, E. C. Cox, J. C. Sturm, P. Silberzan, R. H. Austin. The dynamics of genomic-length DNA molecules in 100-nm channels. *Proc. Natl. Acad. Sci. USA*, 101:10979–10983, 2004.

[3] D Marenduzzo and C Micheletti and E Orlandini. Biopolymer organization upon confinement. *Journal of Physics: Condensed Matter*, 22(28):283102, 2010.

[4] JL Harden and M. Doi. Diffusion of macromolecules in narrow capillaries. *The Journal of Physical Chemistry*, 96(10):4046–4052, 1992.

[5] F. Tessier, J. Labrie, and G.W. Slater. Electrophoretic separation of long polyelectrolytes in submolecular-size constrictions: a monte carlo study. *Macromolecules*, 35(12):4791–4800, 2002.

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

[7] Theo Odijk. Scaling theory of dna confined in nanochannels and nanoslits. *Phys. Rev. E*, 77(6):060901, Jun 2008.

[8] E. Ercolini, F. Valle, J. Adamcik, G. Witz, R. Metzler, P. De Los Rios, J. Roca, G. Dietler. Fractal Dimension and Localization of DNA Knots. *Phys. Rev. Lett.*, 98:058102, 2007.

[9] C. Frontali, E. Dore, A. Ferrauto, E. Gratton, A. Bettini, M. R. Pozzan, and E. Valdevit. An absolute method for the determination of the persistence length of native dna from electron micrographs. *Biopolymers*, 18(6):1353–1373, 1979.

[10] John D. Weeks, David Chandler, and Hans C. Andersen. Role of repulsive forces in determining the equilibrium structure of simple liquids. *The Journal of Chemical Physics*, 54(12):5237–5247, 1971.

[11] K. Kremer, G. Grest. Dynamics of entangled linear polymer melts: A molecular dynamics simulation. *J. Chem. Phys.*, 92:5057–5086, 1990.

[12] M. Rubinstein and R.H. Colby. *Polymer Physics*. Oxford University Press, 2003.

[13] Pierre-Gilles de Gennes. *Scaling concepts in Polymer Physics*. Cornell University Press, Ithaca, New York, 1979.

[14] M Doi and S F Edwards. *The theory of Polymer Dynamics*. Clarendon Press, 1986.

[15] D.C. Rapaport. *The art of molecular dynamics simulation*. Cambridge Univ Pr, 2004.

[16] D. Frenkel and B. Smit. *Understanding molecular simulation: from algorithms to applications*, volume 1. Academic Pr, 2002.

[17] A. Einstein. Uber die von der molekularkinetischen theorie der warme geforderte bewegung von in ruhenden fussigkeiten suspendierten teilchen. *Annalen der Physik*, 322(8):549–560, 1905.

[18] Paul Langevin. Sur la theorie du mouvement brownien. *C. R. Acad. Sci. (Paris)*, 146:530–533, 1908.

[19] R. Zwanzig. *Nonequilibrium Statistical Mechanics*. Oxford University Press, USA, 2001.

[20] Roman Shusterman, Sergey Alon, Tatyana Gavrinyov, and Oleg Krichevsky. Monomer dynamics in double- and single-stranded dna polymers. *Phys. Rev. Lett.*, 92:048303, Jan 2004.

[21] B.P. Flannery, W.H. Press, S.A. Teukolsky, and W. Vetterling. *Numerical recipes in C*. Press Syndicate of the University of Cambridge, New York, 1992.

[22] B. Dunweg and W. Paul. Brownian dynamics simulations without gaussian random numbers. *Int. J. Mod. Phys. C*, 2:817–827, 1991.

[23] M. Tuckerman. *Statistical mechanics: theory and molecular simulation*. Oxford University Press, USA, 2010.

[24] D.S. Lemons and P. Langevin. *An introduction to stochastic processes in physics: containing" On the theory of Brownian motion" by Paul Langevin, translated by Anthony Gythiel*. Johns Hopkins Univ Pr, 2002.

[25] P. Nelson. *Biological Physics (Updated Edition)*. WH Freeman, 2007.

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

[27] Dirk Stigter. Interactions of highly charged colloidal cylinders with applications to double-stranded dna. *Biopolymers*, 16(7):1435–1448, 1977.

[28] T. Odijk. Electrostatic interactions in a solution of linear micelles. *The Journal of Chemical Physics*, 93(7):5172–5176, 1990.

[29] N. M. Toan, D. Marenduzzo, C. Micheletti. Inferring the diameter of a biopolymer from its stretching response. *Biophys. J.*, 89:80–86, Jul 2005.

[30] N. M. Toan, C. Micheletti. Inferring the effective thickness of polyelectrolytes from stretching measurements at various ionic strengths: applications to DNA and RNA. *J. Phys. Condens. Matter*, 18:S269–S281, 2006.

[31] A. Rosa, E. Orlandini, L. Tubiana, and C. Micheletti. Structure and dynamics of ring polymers: Entanglement effects because of solution density and ring topology. *Macromolecules*, 44(21):8668–8680, 2011.

[32] D. Kim and E. Darve. Molecular dynamics simulation of electro-osmotic flows in rough wall nanochannels. *Physical review E*, 73(5):051203, 2006.

[33] J. Liu, M. Wang, S. Chen, and M.O. Robbins. Molecular simulations of electroosmotic flows in rough nanochannels. *Journal of Computational Physics*, 229(20):7834–7847, 2010.

[34] R. Qiao. Effects of molecular level surface roughness on electroosmotic flow. *Microfluidics and Nanofluidics*, 3(1):33–38, 2007.

[35] T. Odijk. The statistics and dynamics of confined or entangled stiff polymers. *Macromol.*, 16:1340–1344, 1983.

[36] P. Lin, C.C. Hsieh, Y.L. Chen, and C.F. Chou. Effects of topology and ionic strength on double-stranded dna confined in nanoslits. *Macromolecules*, 45(6):2920–2927, 2012.

[37] F. Wagner, G. Lattanzi, and E. Frey. Conformations of confined biopolymers. *Phys. Rev. E*, 75(050902), 2007.

[38] C. Micheletti and E. Orlandini. Knotting and metric scaling properties of dna confined in nano-channels: a monte carlo study. *Soft Matter*, 2012.

[39] M. Muller, J. P. Wittmer, M. E. Cates. Topological effects in ring polymers. II. Influence of persistence lenght. *Phys. Rev. E*, 61:4078–4089, 2000.

[40] R.M. Jendrejack, D.C. Schwartz, M.D. Graham, and J.J. De Pablo. Effect of confinement on dna dynamics in microfluidic devices. *The Journal of chemical physics*, 119:1165, 2003.

[41] R.M. Jendrejack, E.T. Dimalanta, D.C. Schwartz, M.D. Graham, and J.J. de Pablo. Dna dynamics in a microchannel. *Physical review letters*, 91(3):38102, 2003.

[42] A. Malevanets and R. Kapral. Mesoscopic model for solvent dynamics. *The Journal of chemical physics*, 110:8605, 1999.

[43] Rae M. Robertson and Douglas E. Smith. Strong effects of molecular topology on diffusion of entangled dna molecules. *Proceedings of the National Academy of Sciences*, 104(12):4824–4827, 2007.