Spherically Confined Polymers: Monte Carlo Simulations with Expanded Ensemble Density-of-States Method

by

Siriporn Pansri

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Physics

Waterloo, Ontario, Canada, 2018

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Andrey Milchev Professor
Supervisor(s)	Jeff Chen Professor
Internal Member	Bae-Yeun Ha Professor
Internal-external Member	Mohammad Kohandel Associate Professor
Other Member(s)	Wing-Ki Liu Adjunct Professor

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In this thesis, the Expanded Ensemble Density-of-States (EXEDOS) method - a combination of the Wang-Landau and Expanded Ensemble Monte Carlo algorithms is employed to investigate spatial conformations of a polymer chain under spherical confinement. The study focuses on flexible chains up to 600 monomers and semi-flexible chains with various stiffnesses up to 300 monomers in length. Spatial conformations of the polymer are studied, using a simple pearl-necklace chain model of varied diameter and stiffness, as well as the model of fused-sphere chain.

To test the applicability of the EXEDOS method, the confinement free energy was calculated for ideal and non-ideal flexible chains inside spheres of sizes smaller than their unconfined size. For ideal chains, the power-law dependence of the free energy on a confining radius is in excellent agreement with previous theoretical predictions. For self-avoiding chains at intermediate concentrations, the dependence of free energy on concentration deviates from that predicted by the blob scaling theory, most likely due to the finite size effects. At high concentrations, a stronger dependence of free energy on concentration is observed, compared to that obtained at intermediate concentrations. The density profile of a self-avoiding flexible chain was also studied, showing that at sufficiently high concentrations, excluded volume interactions push the chain close to the confining surface, leading to an oscillation in monomer number density near the surface.

In semi-flexible chains, bending energy experiences largest changes at low densities as the polymer folds to conform the confining sphere, and at high density its growth slows down as the chain starts forming ordered layer near the surface. We observe isotropic-nematic (I-N) transition for all considered polymer chains. The I-N transition of more flexible chains happens at higher densities than that of stiff chains. All chains form disordered to imperfect helicoildal structures, and at densities above the I-N transition, the structure with four $+1/2$ defects is observed in all considered chains. However, the polymer spatial arrangement is far from an ideal tetrahedral and tennis ball structure.

The EXEDOS algorithm is further extended to investigate the effects of steric hindrance on the structure in a semi-flexible chain, spherically confined at various concentrations. Semi-flexible chains modeled as pearl-necklace chains with ratio of diameter to bond length $d/a \leqslant 0.5$ did not develop ordered structures at any

considered concentrations, while chains with $d/a = 0.8$ and $d/a = 1$, formed imperfect helicoildal structures. On the contrary, a semi-flexible fused-sphere chain with monomer overlap ($d/a = 2$) forms distinct helicoildal structures, when confined inside a small sphere of the same size as the pearl-necklace chains with $d/a = 0.8$ and 1. The evolution of ordered parameters with concentration suggests that during the transition from disordered to ordered configuration, the fused-sphere chain with $d/a = 2$ and pearl-necklace chains with $d/a = 1$ and 0.8 may approach tetrahedral configuration before shifting to a helicoildal arrangement. Four $+1/2$ defects are observed in these chains confined at concentrations above the I-N transition, forming in places, where horizontal and vertical stands of polymer intersect. In the fused-sphere chain, the two $+1/2$ defects nearly merge in each pair to form a $+1$ defect at each pole.

Acknowledgements

First of all, I would like to thank my supervisor, Professor Jeff Chen, for giving me an opportunity to join his research group and to access computing resources at Sharcnet. Without these resources, lengthy simulations needed for this research would have never been possible. I also would like to thank the members of the supervisory committee, Professors Robert Wickham, Bae-Yeun Ha and Wing-Ki Liu for all the assistance and helpful feedback provided throughout my Ph.D. at the University of Waterloo. I would especially like to thank Professor Andrey Milchev and Mohammad Kohandel for some helpful comments during my defense.

Also, I would like to thank the Royal Thai Government for all financial support under the program Strategic Scholarship for Frontier Research Network (SFR) of Thailand's Commission on Higher Education. The scholarships financial supports, including tuition fees, cost of living, health plan, and air fare, made it possible for me to have diverse and valuable experiences while living in Canada.

Lastly, I would like to thank my parents, my brothers and all my relatives for all the long distance support.

Dedication

This is dedicated to Thai people.

Table of Contents

List of Figures

List of Tables

Nomenclature

MC	Monte Carlo
MD	Molecular Dynamic
WL	Wang-Landau
DOS	density of states
EXEDOS	Expanded Ensemble Density-of-States
DFT	density functional theory
LC	Liquid Crystal
LCP	Liquid Crystalline Polymer
dof	degree of freedom
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
LJ	Lennard-Jones
WLC	worm-like chain
BOO	bond orientational order
PEG	polyethylene glycol
$C_2H_2O_4$	ethylene oxide
fcc	faced-center cubic
hcp	hexagonal-closed pack
Si	Silicon
Ge	Germanium
C	Carbon
$\hat{n}$	unit vector in the direction of a nematic director
P_2	second-order Legendre polynomial

P_4	fourth-order Legendre polynomial
$\mathbf{T}$	nematic tensor
Q	nematic order parameter
q_l	l order of bond orintational order parameter
q_3	tetrahedral order parameter
Y_{lm}	spherical harmonics of order lm
r_c	confining radius
r_G	unconfined size of polymer
l_p	persistence length
$\vec{u}$	bond vector
η	volume fraction
ρ	density
n	number density
F	free energy
E_b	bending energy

Chapter 1

Introduction

1.1 Motivation

In the recent years, the topic of polymer confinement inside a sphere has received a lot of interest, as it has turned out relevant to problems of molecular storage and transport in biological cells [1–3], and to some technological applications [4–11].

Polymers have some remarkable properties. They can be tailored for high mechanical strength, flexibility, or chemical resistivity [12]. Their properties also make polymers of interest to liquid crystal technology [11]. Confinement of the polymers can induce formation of orientationally ordered structure, similar to that observed in molecules of liquid crystalline (LC) phase. The ability to control orientation and ordering of the molecules in LC phase can allow one to fine-tune physical properties of materials for certain LC based devices [13, 14]. For example, a molecular configuration with two +1 defects formed in microscopic droplets can be controlled by weak external field to modulate the direction and intensity of light for display or beam steering applications [15]. Another interesting example is application of molecules in LC phase in laser technology: liquid crystalline textures can form four +1/2 defects on a spherical surface, and express properties similar to those of sp^3 hybridized atoms such as Carbon (C), Silicon (Si) and Germanium (Ge) [16]. These defects can be functionalized for molecular attachment to form photonic crystals for manipulating and trapping light [17]. Knowing mechanism of defect formation can help one to create, remove or alter such defects so that they would better serve the purpose of specific devices. This makes the topic of spherically confined polymer

highly relevant to technology.

In biology, many biomolecules such as deoxyribonucleic acid (DNA), ribonucleic acid (RNA) and proteins are confined inside organelles of sizes many times smaller than their unconfined sizes [18]. The interaction of these molecules with confining boundaries strongly influences their molecular architecture [19–21], which is related to their function [18]. Therefore, understanding how this molecular organization is formed and maintained can lead to a better explanation of cellular functions and processes such as gene expression and DNA/RNA packing, which can help in development of novel medical applications.

Examples above show that understanding molecular organization under confinement is crucial for development of LC based and medical devices and for fundamental understanding of cell biology. Study of biomolecular organization began since 1920s with organisms that have simple life-cycle, such as bacteriophages, which are viruses that infect bacteria and replicate within them, to understand fundamental biology, role of DNA and nature of genetic code [22]. One of the most widely studied model systems, that represent DNA packed inside a confining sphere is bacteriophage T7. Its spool-like structure was revealed by cryo-electron microscopy in 1997 [23]. However, limitations of the experiment prevent uncovering the formation mechanism of this structure. Therefore, several computer simulations [24, 25] and theoretical studies [26] were used to give some insight to processes of DNA packing. It was theoretically predicted that DNA forms ordered structure in the early stage of packing [27]. Simulations of semi-flexible chains, often used as models of biomolecules, show that inside a sphere, the chain usually undergoes transition from coil-like to toroidal structure, when radius of the confining sphere becomes comparable to the chain persistence length (l_p) [3, 28–30]. With increasing concentration, a spool-like structure becomes preferred [31]. It was predicted by computer simulations [24, 25] that formation of the spool-like structure in bacteriophage DNA seems to depend on packing dynamics. Spakowitz *et al.* using MD simulations, predicted that this structure would only be observed if the DNA is twisted when injected into the cavity [24]. On the contrary, particle dynamics simulations by Fathizadeh *et al.* predicted that the spool-like structure would be displayed in DNA injected inside a sphere without twisting, while a tennis ball structure, with four $+1/2$ defects would form if the DNA is compressed into a sphere [25]. The discrepancy between results of different simulations shown above, indicates that the simulation methods still need

refinement and fine-tuning to a particular system of interest. It is interesting that the structure with four +1/2 defects was also observed in experiments [32, 33], and simulations [34] of rigid molecules confined in a thin uniform spherical shell. In a thick spherical shell and in a droplet, previous experiment [32, 33] and simulation studies [34] showed that the structure with two +1 defects was observed instead. However, recent MD simulations of multiple rigid chains (each chain having 32 monomers and bending stiffness of 32 in a unit of thermal energy) confined inside a sphere at density $\rho \geq 0.4$ showed the presence of a tennis ball structure [35]. For a system of less stiff chains (each chain having 32 monomers and bending stiffness of 24 in a unit of thermal energy), the tennis ball structure has been observed only at sufficiently high density ($\rho = 0.7$) [35].

To my knowledge, the only study showing the presence of a tennis ball structure in a single semi-flexible chain densely confined inside a sphere, was the particle dynamics simulation [25] conducted by Fathizadeh *et al.* In this study [25], the tennis ball structure was visualized and defect locations were identified through calculation of local nematic order parameter (Q_{local}). The current study will investigate the mechanisms of spatial organization in a semi-flexible chain confined inside a sphere using Monte Carlo simulations. I will use various order parameters, including the tetrahedral order parameter (q_3) to analysis polymer organization. The parameter q_3 was successfully used in a number of studies that addressed crystallization problems [36–38], to probe for the arrangement of the structure with four +1/2 defects at different levels of confinement. In those studies, the value of q_3 was non-zero, when a structure with tetrahedral symmetry was presented, but nearly vanished when the structure was arranged in non-tetrahedral fashion [36–38]. We will also calculate the Q_{local} values to identify defect locations in each polymer configuration and compare our results with those already reported [25].

The most popular simulations that apply classical mechanics to a system of interest are molecular dynamics (MD) and Monte Carlo (MC) simulations. In MD simulations, a particle moves following Newton's laws of motion. Thus, MD simulations are suitable for calculation of time - dependent physical properties such as diffusion coefficient, or interaction rates [39]. In MC simulations, a particle moves following a probabilistic algorithm, leading a system of interest to equilibrium. Therefore, MC simulations are better suited when static properties of a system such as ensemble average of the order parameters, or end-to-end distance and free energy are

of interest [40, 41]. However, the traditional MC algorithm called the Metropolis method, is not efficient when applied to systems at low temperatures, because of inadequate sampling [41]. The Metropolis method also loses efficiency when applied on systems with phase transitions. In a system with second order phase transition, the Metropolis method also needs exceedingly long time to sample the states close to transition region [40, 41]. In a system with the first order phase transitions, two phases usually coexist, and the Metropolis method needs long time to tunnel from one minimum of the Gibbs potential to the other. To overcome the potential barrier between phases, one needs the method which is able to generate a biased random walk in configurational space to modify the acceptance probability and increase the likelihood of transition between minima of the Gibbs potential.

Wang and Landau [42, 43] introduced a new MC algorithm, in which the acceptance probability is inversely proportional to the density of energy states modified at every simulation step. In this algorithm, all accessible states (favorable or less favorable) seem to be sampled more efficiently. Good performance of this algorithm was confirmed by many authors, especially on systems with discrete energy spectrum such as Ising model and Potts model [42, 43]. However, when the algorithm was used on systems with continuous energy spectrum, like proteins [44], polymers [45, 46] and liquid crystals [47], its efficiency was relatively low and modification of the original WL algorithm was needed. In systems with continuous energy spectrum, the number of states is infinitely large, and even for a small system, one needs large computational cost to sufficiently sample the whole energy range. Instead of sampling along energy space, Kim *et al.* [48] suggested to apply the WL algorithm [42, 43] to other reaction coordinates such as order parameter and distance between molecules, using the approach of expanded ensemble [84]. This new algorithm was named the expanded density of states (EXEDOS) [48], and proven to be efficient when applied to various systems and processes such as solidification of particles [49], interaction of counterions lying between charged surfaces [50], and transition between coil-bridge states in a self-avoiding chain attracted to a surface [51].

MC simulations are most suitable for this study, because we are only interested in calculation of static properties. Due to the limitations of the Metropolis MC method, here we apply the EXEDOS method to systems of spherically confined polymers to investigate the effects of confinement on polymer spatial organization and free energy. We will first consider an ideal flexible chain confined inside a sphere,

calculate the free energy, and compare it to available theoretical and simulation data to test validity of the method. Then, we will expand the EXEDOS algorithm to systems of self-avoiding flexible chains and semi-flexible chains confined inside a sphere of various sizes.

1.2 Background

1.2.1 Coarse-grained polymer model

Polymers are composed of a number of repeating structural units, called monomers, covalently bonded to each other [12]. The units can form linear chains as well as three dimensional networks, depending on the way monomers are linked to each other. For example, DNA is composed of repeating units of nucleotides which are linearly linked to each other, and a 3-dimensional network of polysaccharide hydrogels is formed by units of saccharide isomers [12].

Polymers are highly functional materials which can be tailored to have specific physical, chemical and mechanical properties; thus, they are extensively used in various industries [10]. When dissolved in a good solvent, a polymer molecule forms a coiled structure with a large number of internal degrees of freedom. Extremely large number of the degrees of freedom makes theoretical study of polymers at microscopic level difficult. Therefore, coarse-grained polymer models, in which one particle represents a group of atoms, typically replace the atomic model of polymers [41].

Coarse-grained polymer models can be divided into ideal and non-ideal chains. Here we define an ideal chain as a chain that allows monomer-monomer overlap, and non-ideal chain as a chain in which the overlap is not allowed. Apart from monomer-monomer overlap, flexibility of a chain can be modeled by specifying the correlation between directions of bonds and/or torsion angles so that the flexible chain becomes stiffer. Below we describe the models of linear polymers used in the thesis.

1.2.1.1 Ideal chain

The polymer chain is called ideal, when the monomers are treated as infinitely small particles with no excluded volume [52]. Consider a freely-jointed chain of N segments with position of each segment defined by vector $\vec{u}_i$ making a random walk with a step length $a = |\vec{u}_i|$. Direction of each segment is uncorrelated with that of other segments in the chain. Interactions between non-adjacent segments are neglected (no excluded volume interactions).

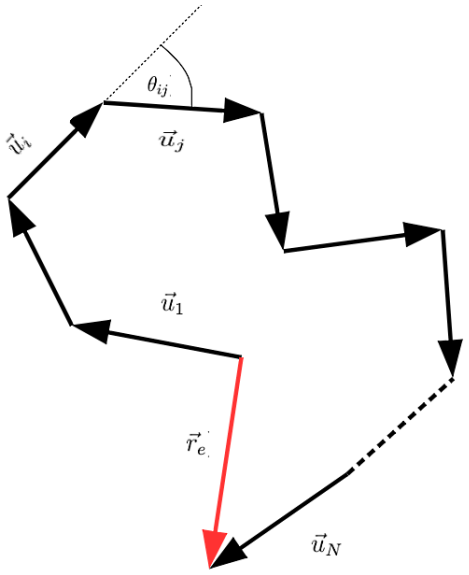

Figure 1.1: An ideal chain.

The end to end vector ($\vec{r}_e$) of this chain is given by:

$$\vec{r}_e = \Sigma_{i=1}^{N} \vec{u}_i, \tag{1.1}$$

and the average of this vector is $\langle \vec{r}_e \rangle = 0$. Instead of looking at $\vec{r}_e$, which has an average value of zero, one can look at the mean square end-to-end distance $\langle \vec{r}_e^2 \rangle$ to define a chain size.

$$\langle \vec{r}_e^2 \rangle = Na^2. \tag{1.2}$$

Generalizing Eq. 1.2, the size r_e depends on number of segments N as:

$$r_e \sim aN^{\nu}, \tag{1.3}$$

where ν is called the size exponent which is equal to 1/2 for ideal chain [53].

1.2.1.2 Non-ideal chain

The excluded volume interaction is the simplest form of interactions that makes a chain non-ideal. To model such a chain, one can introduce monomer-monomer interactions in the model by considering the size of the monomers and the volume they occupy. The volume occupied by the monomers is proportional to the excluded volume, and the interaction is called the excluded volume interaction.

Strength of the excluded volume interaction depends on temperature. At high temperature, the repulsive interactions between pairs of monomers dominate; forcing the chain to increase in size. At transition temperature between coil and globular configuration, attractive and repulsive interactions between molecules compensate each other, and the chain behaves similar to the ideal chain. At temperatures lower than the transition temperature, the chain collapses into a globular configuration.

Flory used the mean field approach to estimate the free energy resulting from the addition of excluded volume interaction in different solution conditions [53]. Following Flory's approach [53], one can estimate size of a real chain in various solvents as $r_e \propto N^{\nu}$, where ν is the Flory exponent (ν =3/5, 1/2, and 1/3 for the chain in good solvent, θ solvent and bad solvent, respectively, and N is the number of bonds in the chain.) Solvent quality also depends on temperature. At low temperature, the solvent may have poor quality, while at high temperature the same solvent may become a good solvent. At temperature that is equal to a particular θ temperature, a non-ideal chain behaves similar to an ideal chain. For a real chain in good solvent, the value of ν was determined in experiment [54] and theory [55] as: $\nu \approx 0.588 \pm 0.001$.

1.2.1.3 Semi-flexible chain

Many biopolymers such as DNA, RNA and actin are semi-flexible in nature, as their backbone cannot freely rotate due to steric hindrance, as shown in Fig 1.2 (a). A semi-flexible chain model can be continuous or discrete.

The continuous model describes backbone monomers as a continuous curve $r(s)$, as shown in Fig 1.2 (b), and it is mostly used in analytical studies. Bending rigidity is introduced in this chain through bending stiffness used for calculation of bending

energy cost (expressed in Eq. 1.6) between tangent vectors at segment s ($\hat{u}(s)$) and s' ($\hat{u}(s')$) in the chain ($|\hat{u}(s)| = |\hat{u}(s')| = 1$).

A discrete model has been used in a number of simulation studies [35, 56], and successfully predicted the behavior of biopolymers (such as DNA) under geometrical constraint [35]. This model considers a polymer chain as a string of $N+1$ monomers, each monomer having diameter d, and connected to its neighbors by a bond of length a (Fig. 1.2 (c)). The chain has contour length $L = Na$. The tangent-tangent correlation function ($\Phi_{\hat{u}\hat{u}'}$), which is the central quantity for a semi-flexible chain, is expressed as [12, 57]:

$$\Phi_{\hat{u}\hat{u}'}(i,j) = \langle \hat{u}(i) \cdot \hat{u}(j) \rangle = exp(-\frac{|i-j|}{l_p}), \tag{1.4}$$

where $\hat{u}(i)$ and $\hat{u}(j)$ are tangent vectors in the directions of bond i and j, respectively, ($|\hat{u}(i)| = |\hat{u}(j)| = 1$), and l_p is the decay length called persistence length. The mean square of polymer size is given as [12]:

$$\langle r_e^2 \rangle = \langle \vec{r}_e \cdot \vec{r}_e \rangle = 2l_pL(1 - \frac{l_p}{L}(1 - e^{-L/l_p})). \tag{1.5}$$

If $l_p << L$, the chain becomes flexible, and Eq. 1.5 simplifies to $\langle r_e^2 \rangle \cong 2l_pL = a_kL$, where $a_k = 2l_p$ and a_k is effective Kuhn length, which is defined only for long chains. On the other hand, if $l_p >> L$, the chain behaves as a stiff rod with $\langle r_e^2 \rangle \cong L^2$. Here, we will focus only on the discrete model as it will be applied in our study.

To model a semi-flexible chain, one needs to introduce a bending potential associated with the angle between two adjacent bonds. For a semi-flexible chain, it is assumed that rotation of two adjacent bonds relative to each other introduces a bending energy cost (E_b) defined as [58, 59]:

$$
\begin{aligned}
E_b &= \sum_{i=1}^{N} \frac{l_p k_B T}{2a} |\hat{u}_i - \hat{u}_{i+1}|^2 \\
&= \sum_{i=1}^{N} \frac{l_p k_B T}{a} (1 - cos\ \theta_{i,i+1}) \\
&= \sum_{i=1}^{N} k_b (1 - cos\ \theta_{i,i+1}),
\end{aligned}
\tag{1.6}
$$

where $k_b = l_p k_B T/a$ is bending stiffness, $\theta_{i,i+1}$ is the angle between tangent vector $\hat{u}_i$ and $\hat{u}_{i+1}$, as seen in Fig. 1.2 (c) and a is a bond length ($a = d$, d is monomer diameter).

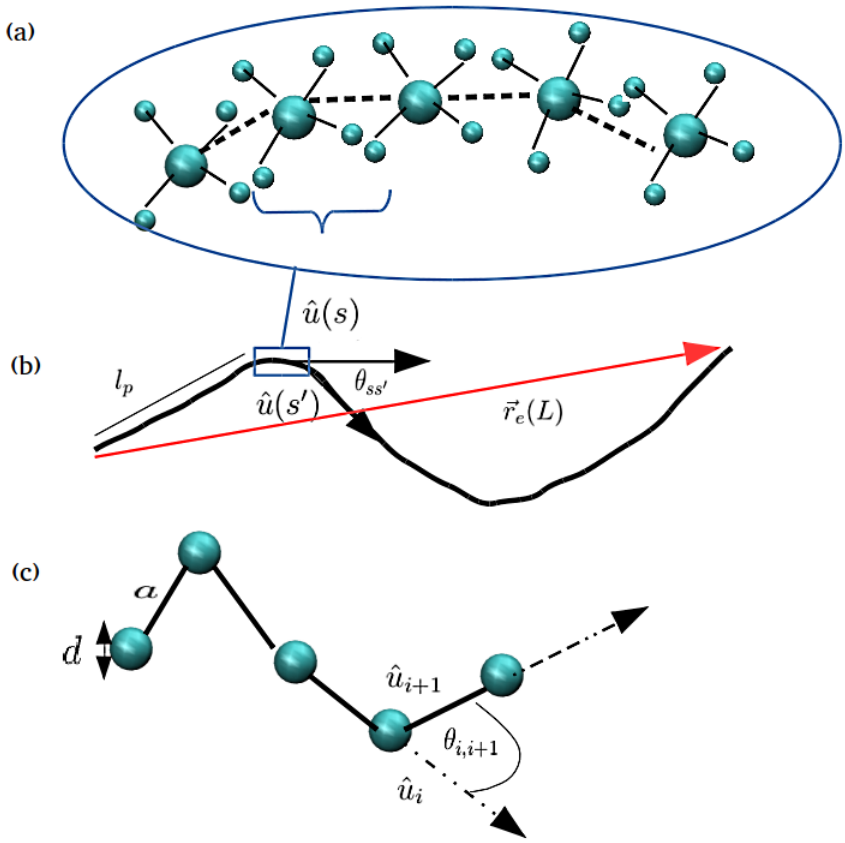

Figure 1.2: (a) Schematic representation of a semi-flexible chain with steric repulsion between backbone limiting rotational freedom of the chain. (b) The semi-flexible chain is shown as a continuous space curve $r(s)$. $\hat{u}(s)$ and $\hat{u}(s')$ are tangent vectors at segments s and s', and $\vec{r}_e$ is end-to-end vector. (c) A discrete semi-flexible chain is composted of beads, each bead having diameter d and being connected to its adjacent neighbors by a bond of length a. $\hat{u}_i$ and $\hat{u}_{i+1}$ are unit vectors in the directions of bonds i, and $i+1$, $\theta_{i,i+1}$ is the angle between these two bonds.

1.2.2 Confinement free energy of a flexible polymer under spherical confinement

To characterize the statistical properties of a confined polymer in canonical ensemble, we use the Helmholtz free energy (F), which is the thermodynamic potential that measures the work obtainable from a closed thermodynamic system at a constant temperature and volume. The Helmholtz free energy is defined as:

$$F = U - TS, \tag{1.7}$$

where U is internal energy of a system, T is an absolute temperature and S is the system's entropy, which is related to conformational multiplicity Ω or a number of system microstates as:

$$S = k_B \ln \Omega. \tag{1.8}$$

Here, k_B is the Boltzmann constant. For athermal chain $U = 0$, the free energy is equal to $-k_B T \ln \Omega$.

Consider a flexible polymer chain with unconfined size $r_G = aN^{\nu}$ (N is number of segments and a is segment length), and with bulk conformational multiplicity Ω. When confined inside a sphere of diameter $D >> r_G$, the chain does not feel the confining sphere, and Ω does not noticeably change. However, as the confining diameter D decreases, the chain starts to interact with the confining boundary, leading to a decrease in the number of accessible conformations, affecting Ω, and thus the entropy. Consequently, the free energy increases. One can use the blob model, which will be described below, to understand a concept of the confinement free energy.

An important concept allowing one to describe statistical properties of a chain of N segments in the presence of perturbation is the concept of blob [60], which defines the length scale below which segments are unaffected by perturbation created by either an extension force or confining surface. In blob concept, a system is discretized into blobs, each blob having size $\xi \sim D$. If the chain is confined in a sphere of size $D < r_G$ (Fig. 1.3), the segments inside each blob are unaffected by the confining sphere, but on the scale larger than ξ they are affected by the surface. If each blob would include g segments, the chain of N segments will consist of N/g blobs. Since, the segments inside each blob are unperturbed, we can apply statistics of an isolated

chain to these segments. In this case, the size ξ of each blob in ideal chain will be:

$$\xi^2 = a^2 g \tag{1.9}$$

and in the non-ideal chain it will be:

$$\xi^2 = a^2 g^{2\nu}. \tag{1.10}$$

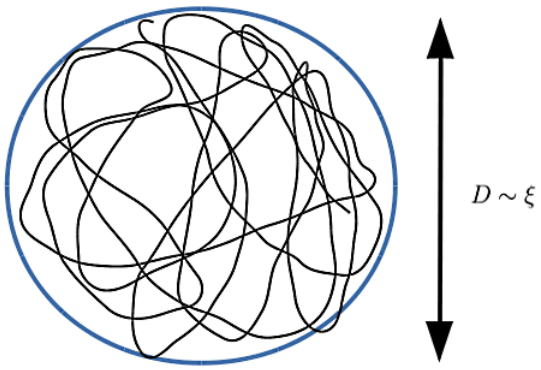

Figure 1.3: A polymer system divided into blobs, each blob having size of ξ, is confined inside a sphere of size $D < r_G$, where $r_G = aN^{\nu}$ is unconfined size of the chain.

Consider an ideal chain of N segments, with each segment having length a, confined inside a sphere of diameter $D < r_G = a \cdot N^{1/2}$. In blob theory, the blob size ξ is taken to be equal to the confining diameter D; thus, Eq. 1.9 can then be written as:

$$D^2 \approx a^2 g. \tag{1.11}$$

It is assumed that each blob has a particular direction associated with it; thus, it is restricted to one degree of freedom. Therefore, the confinement free energy [61, 62] which is the work needed to bring the chain from bulk ($D >> r_G$) to a confined state increases by $k_B T$ per blob, and can be written for the ideal chain as:

$$F_{id} \backsimeq k_B T \frac{N}{g} \backsimeq k_B T (\frac{r_G}{D})^2. \tag{1.12}$$

In a non-ideal chain or a self-avoiding chain confined inside a sphere of size $D < r_G$, the confinement free energy is no longer extensive in N, since changes in the chain length affect volume fraction of the system [63]. We know that the confinement free energy depends only on r_G and and polymer concentration ($\sim D^3$). Therefore, scaling law for the confinement free energy can be written as:

$$F \backsimeq k_B T(\frac{r_G}{D})^{3/(3\nu-1)} = N k_B T \eta^{1/3\nu-1}, \tag{1.13}$$

where $\eta = Na^3/D^3$. Free energy expressed by Eq. 1.13 can also be obtained using blob concept. We know that inside a blob of size ξ, behavior of the monomers is similar to that in bulk: $\xi = ag^{\nu}$. Volume fraction can also be expressed as: $\eta = ga^3/\xi^3$. Combining these two relations, the blob size ξ can be written as:

$$\xi = a(\frac{D^3}{a^3 N})^{\nu/(3\nu-1)}. \tag{1.14}$$

Under assumption that each blob has one degree of freedom, the confinement free energy can be rewritten in terms of $k_B T$ per blob as [64]:

$$F = k_B T \frac{D^3}{\xi^3} \tag{1.15}$$

As the confining diameter becomes much smaller than unconfined size of the polymer chain ($D \ll r_G$) and the blob size becomes approximately the same as monomer size $\xi \backsimeq a$, corresponding to concentrated regime, excluded volume interaction is screened beyond the length scale a. Thus, the confinement free energy again can be described by Eq. 1.12 [65].

We should note that the investigation of free energy properties of a semi-flexible chain is not the main interest of this thesis. The readers may consult the work of Sakaue [66], if the free energy function of a semi-flexible chain confined inside a sphere is of interest.

1.2.3 Liquid crystals

The liquid crystalline phase is an intermediate phase between solid and liquid. Molecules in this phase can flow like ordinary liquid, but at the same time have anisotropic properties, such as optical birefringence, electric and magnetic anisotropy

including molecular ordering which is characteristic of solid state [67–69]. Due to a combination of molecular ordering and fluidity, liquid crystals (LCs) carry very unique physical properties and receive a lot of attention in both technology and science.

In liquid phase, molecules can move freely with respect to each other. Their positions and orientations are randomly distributed (as seen in Fig. 1.4 a)). Theoretically, a phase transition is defined only in thermodynamic limit i.e. a system with infinite number of particles and volume, but finite density, and only in equilibrium [70]. Such system exists only in theory but it is in a good approximation for bulk system in reality where the system's surface is negligibly small with respect to system volume. Using this approximation, transition from isotropic phase into liquid crystalline phase of materials can be achieved by changing temperature or concentration, or both. Cholesterol and DNA are examples of materials that undergo transition from isotropic into liquid crystalline phase due to changing temperature and concentration, respectively.

With respect to molecular ordering, liquid crystalline phases are often divided into three-major classes: nematic (Fig. 1.4 b)), cholesteric (Fig. 1.4 c)) and smectic phase (Fig. 1.4 d)) [67–69]. In nematic phase, the arrangement of molecules exhibits long-range orientaional order but no positional order. The molecules in this phase tend to align their long-axes parallel to each other, along a preferred direction, called director ($\hat{n}$ in Fig. 1.4). Lack of correlation in positional order allows these molecules to move like a liquid. The molecules in cholestic phase also have long-range orientational order but no positional order: the director of the system varies as it moves through the medium. Arrangement of molecules in smectic phase differs from that in nematic and cholestic phases in that the molecules are arranged in layers and exhibit both orientational and positional ordering [67–69].

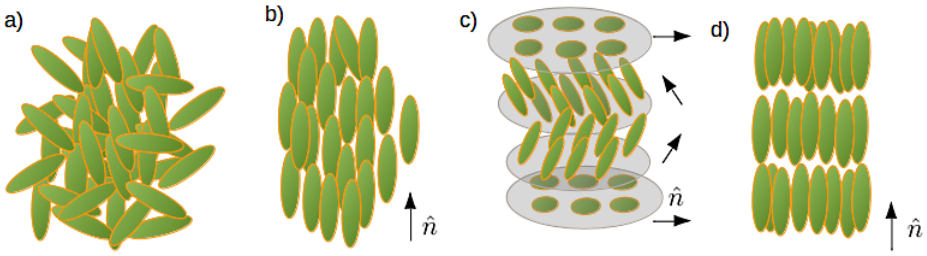

Figure 1.4: Schematic representation of a) isotropic b) nematic, c) cholestic and d) smectic phase. Unit vector $\hat{n}$ denotes an averaged director of molecules. $\hat{n}$ and $-\hat{n}$ are indistinguishable.

Materials that pass through more than one phase as they undergo phase transition from liquid to solid, are called polymorphous. In these materials, decreasing temperature or increasing concentration can progressively reduce translational and orientational entropy of molecules. The molecules reorient themselves, forming more and more ordered structures. Transition from isotropic liquid phase into solid phase usually follows the sequence: *isotropic* $\longrightarrow$ *nematic or cholesteric* $\longrightarrow$ *smectic* $\longrightarrow$ *solid*, as the temperature decreases or concentration increases.

1.2.3.1 Identification of phases

Nematic order parameter

Due to simplicity of molecular organization in nematic phases, compared to other liquid crystalline states, transformation between isotropic fluid and nematic liquid crystal has been extensively studied to gain fundamental understanding about this phase transition. LC phase can be identified using a concept of order parameter, a quantity that vanishes in one phase, but has a finite value in another one. The main distinction between molecules in isotropic and nematic phase is in their symmetry. In an isotropic phase, molecules can move freely and all orientations are equally possible, while, molecules in nematic phase tend to align along a preferred direction called director ($\hat{n}$) and become correlated in space. Thus, molecular symmetry should be identified as an order parameter to distinguish isotropic fluid from nematic liquid crystal.

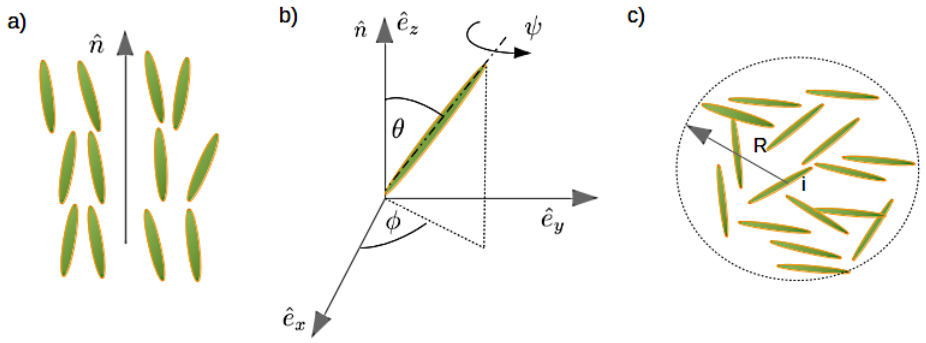

Figure 1.5: a) Schematic representation of nematic liquid crystals ($\hat{n}$ here is a nematic director), b) orientation of LC molecule represented through Eulerian angles; θ is the angle between long molecular axis of the molecule and the director $\hat{n}$, ϕ is the angle the molecule rotates in azimuthal direction and ψ is the angle the molecule rotates about its long axis c) particles (neighbors) within a cut-off distance R, relative to particle i.

The orientation of a single molecule in nematic phase can be described using three Eulerian angles: (θ, ϕ, ψ) as illustrated in Fig. 1.5 b). Angle θ is the angle between long-axis of the molecule and the director $\hat{n}$ aligned with the $z - axis$ of a fixed Cartesian coordinate system. Angles ϕ and ψ describe the rotation of the molecule in azimuthal direction and about its long molecular axis, respectively. In general, physical properties of molecules can be different in all three principal directions $\hat{e}_x, \hat{e}_y, \hat{e}_z$. However, nematic phase has cylindrical symmetry, with physical properties measured along the director different from those measured perpendicular to it. If we look closely at orientation of a single molecule in Fig. 1.5 a), we can see that having cylindrical symmetry the molecules need only angle θ to describe their degree of ordering with respect to the director $\hat{n}$. In isotropic phase there is no preferred angle, all angles are equally probable, while in nematic phase the most probable angle is θ. Thus, the angle θ could be used as an order parameter to differentiate isotropic and nematic phases. However, it is not convenient to measure the angle θ itself, and projection of the molecule along $\hat{n}$ ($cos\theta$) is measured, instead. Directions $\hat{n}$ and $-\hat{n}$ are indistinguishable; therefore, $cos^2\theta$ is often used to describe the molecular orientation. Traditionally, order parameter should be equal to unity in perfectly ordered phase and zero in isotropic phase. The term:

$$\langle P_2 \rangle = \frac{3\langle cos^2\theta \rangle - 1}{2} \tag{1.16}$$

satisfies this condition and is often used as an order parameter for nematic liquid crystals. $\langle ... \rangle$ denotes an average of a quantity of interest. The values for $P_2 = 0,\ 1,\ -0.5$ correspond to the molecules that are perfectly isotropic, parallel to the director or perpendicular to it, respectively. Expression of P_2 in Eq. 1.16 coincides with the second Legendre polynomial, and has been used to study order-disorder transition in many systems, such as rod-like molecules and a semi-flexible chain confined inside a sphere [26, 71]. In some cases, higher order Legendre polynomials are used to achieve more reliable and complete characterization of molecular arrangement [72, 73]. For example, in the polarized Raman measurements of nematic alkylcyanobiphenyls, parameter P_4 has been used to show the level of molecular association (dimerization) in the molecules, which cannot be measured with P_2 order parameter [73].

The definition of P_2 in Eq. 1.16 can only be used, when the nematic director $\hat{n}$ is prior known. In practice, it is not aways easy to define $\hat{n}$ especially in a system of many particles. In this case, molecular structures can be characterized using the following equation, which does not require prior knowledge of $\hat{n}$ [74]:

$$\mathbf{T}_i = \frac{1}{2N_b} \sum_{j=1}^{N_b} (3\hat{u}_{j\alpha}\hat{u}_{j\beta} - \delta_{\alpha\beta}), \tag{1.17}$$

where $\hat{u}_{j\alpha}$ is an α component ($\alpha = x, y, z$) of bond j, which is a unit vector along long axis of molecule j, $\delta_{\alpha\beta}$ is Kronecker delta, N_b is the number of nearest neighbors for molecule i. Eigenvalues λ_+, λ_0 and λ_- are obtained by diagonalization of $\mathbf{T}$. The eigenvalue with the largest absolute value is considered to be a nematic order parameter (Q) and its corresponding eigenvector is called a nematic director ($\hat{n}$). If $\lambda_+ >> \lambda_0 \approx \lambda_-$, the structure has one preferred direction (uniaxial nematic). If $\lambda_+ = \lambda_0 = \lambda_- = 0$ (in an infinitely large system), the structure is an isotropic phase. For intermediate case where $\lambda_+ \neq \lambda_0 \neq \lambda_- \neq 0$, the structure is biaxial. Theoretically, the value of Q lays in the range $-0.5 \leq Q \leq 1$. The value $Q = 1$ corresponds to a perfectly ordered structure, with all molecules are perfectly aligned (which can never happen in practice). The value $Q = -0.5$ corresponds to a perfect oblate structure.

The nematic order parameter can be calculated both globally and locally. In the global case, the tensor $\mathbf{T}$ is built using all molecules in a system, while, in local calculation only molecules located within a cut-off distance from molecule i are used to construct the tensor $\mathbf{T}$.

Bond orientational order parameter

In order to describe crystalline or semi-crystalline structures in a state of molecular transition from liquid to solid phase, one would consider local bond ordering induced by dense packing inside a hard sphere or by directional bonding in molecular and atomic liquid. The bond orientation is locally unique in different close-packed formations or in different packing symmetries [75]. For example, atoms of group-IV elements such as Si and Ge have four covalently bonded nearest neighbors, and form tetrahedral symmetry [76]. With increasing pressure higher than 10 GPa, the tetrahedral network is broken and each atom of these group-IV elements becomes surrounded by six nearest neighbors, forming hexagonal symmetry [76].

Local bond ordering can be studied through the distribution of bonds joining a particle i positioned at $\vec{r}_i$ and its nearest neighbors [75, 94]. Expanding the distribution in spherical harmonic form, a rotationally invariant coefficient of the expansion at specific order (l) is unique to a specific symmetry as it depends on angles describing local bond orientations, as we shall see below. Thus, these coefficients can also be used as local bond orientational order parameters [75, 94].

Let us first consider particles located within the distance R relative to the particle i (Fig. 1.5 c)), which are the nearest neighbors of the particle. Each line connecting particle i with another one forms a bond j. If we assume that each particle is symmetric, rotation around its long axes should not cause any change. Therefore, orientation of every bond j can be described by spherical angles $(\theta_{ij}, \phi_{ij}) \equiv \Omega_{ij}$, as shown in Fig. 1.5 b). In this coordinate system, the angular distribution of neighboring bonds to the particle i ($\varrho(\vec{r}_i, \Omega)$), can be expressed in terms of spherical harmonics as:

$$\varrho(\vec{r}_i, \Omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} q_{lm}(\vec{r}_i) Y_{lm}^{*}(\Omega), \tag{1.18}$$

where Y_{lm} is a spherical harmonic function and the $q_{lm}(\vec{r}_i)$ are the expansion coefficients defined as the bond orientational order parameter. From the completeness

relation for spherical harmonics, $q_{lm}(\vec{r}_i)$ for particle i can be written as:

$$q_{lm}(i) = \frac{1}{N_b(i)} \sum_{j=1}^{N_b(i)} Y_{lm}(\vec{r}_{ij}(\theta, \phi)), \tag{1.19}$$

where N_b is the number of nearest neighbors of particle i, and $\vec{r}_{ij}$ is the vector from particle i to j. Eq. 1.19 can be rewritten for the rotational invariants of spherical harmonics as :

$$q_l(i) = \sqrt{\frac{4\pi}{2l+1} \sum_{m=-l}^{l} |q_{lm}(i)|^2}. \tag{1.20}$$

The value of q_l depends only on the relative angle between bonds, and it is therefore independent of the choice of coordinate system. The bond orientational order parameter q_l has been used on several systems having perfect crystalline symmetry [38, 75]. It was found that different l values (except $l = 0$) are sensitive to different crystalline symmetries [38, 75]. For example, in a system with ideal tetrahedral structure such as diamond, only the component of q_3 is nonzero ($q_3 \approx 0.745$) [38]. For systems with cubic symmetry, q_4 is nonzero and for those with icosahedral symmetry such as super cooled Lennard-Jones (LJ), q_6 is nonzero [75] .

1.2.3.2 Topological defects in nematic liquid crystals

Confinement of LC molecules on a curved surface causes distortion in their orientations, creating areas in which the directors are undefined and order parameters are close to zero, called defects [68]. Defects in nematic liquid crystals are commonly observed as points (zero dimension), lines (one dimension), and sheets (two dimensions). Sheet defects are usually unstable; therefore, they are not frequently observed or studied [77]. The point defects can be characterized by a number called charge, or winding number. Suppose we have a field of vectors Φ on a surface. If we draw a closed contour l on this surface in a clockwise direction and follow this contour, the direction of vector field will change as we move. Assuming that the field has no discontinuities along the contour, this change will be incremental over the contour l : $\Delta\Phi = \frac{d\Phi}{dl}\Delta l$. The number: $\frac{1}{2\pi}\int \frac{d\Phi}{dl} dl$ is called charge. If there is no defect inside the contour, this integral will be zero. For example, to draw a circle around the defect in Fig. 1.6 (a) in a clockwise direction, we have to rotate by angles

of 2π radians. The field also rotates by angles of 2π radians in clockwise direction, the defect in Fig. 1.6 (a) has a charge of $2\pi/2\pi = +1$, accordingly. In Fig. 1.6 (b), the field rotates by angles of π radians, so the defect has a charge of $+1/2$. The field in Fig. 1.6 (c) rotates by angles of π radians in anti-clockwise direction, so it has a charge of $-1/2$.

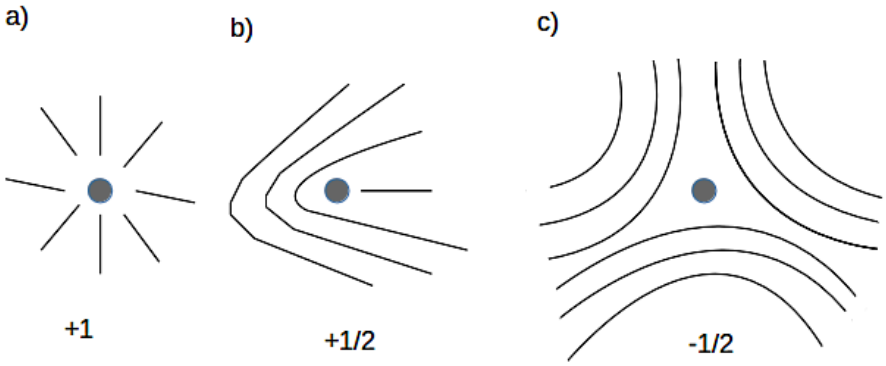

Figure 1.6: Defects with topological charge of +1 (a), +1/2 (b), and -1/2 (c).

Under geometrical constraint, the total topological charge on arbitrary surface is limited, according to the Euler [78] and Poincare [79] and Hopf theorems [80]:

$$\sum_i s_i = \chi = 2(1-g). \tag{1.21}$$

Here the index i runs over defects, s is topological charge of each defect, and g is total number of handles. A sphere does not have a handle; therefore, total topological charge of a spherical surface should be $+2$.

1.2.4 Computer simulations

A computer simulation forms a bridge between experimental observation and theoretical predictions by simulating the behavior of processes and systems. Molecular

dynamics and Monte Carlo simulations are among the most well known methods used to investigate physical properties of systems with many particles. In MD simulations, particles move following Newton's laws of motion to reproduce evolution of the system in time. This makes MD technique useful when dynamic properties, such as interaction rate, transport coefficients, and time-dependent responses to perturbations of a system, are of interest. Time averaged properties of the system can be obtained in MD simulations, and can be considered as an ensemble average, when sampling is assumed to be ergodic. In MC technique, particles move following a probabilistic algorithm. This makes MC method suitable for investigation of time-independent properties such as end-to-end and order parameter of a polymer chain.

Both MD and MC simulations can be divided into atomistic and coarse-grained simulations. In atomistic approach, interactions between all atoms are taken into account. This method is only suitable for systems with small number of atoms. In coarse-grained simulation, atoms are grouped into blocks and each block is then treated as a single particle. Therefore, the coarse-grained simulation is usually employed to investigate physical properties of larger systems [81].

In this thesis, we are interested in calculating the time-independent properties of polymers without considering atomistic details such as interaction between hydrogen-oxygen and hydrogen-carbon. Therefore, coarse-grained MC simulation was our chosen option.

The most basic approach used in MC simulations is the Metropolis algorithm, which directly generates a new state from an old state of a system with a carefully-designed transition probability to brings the system closer to equilibrium. The method has been often used to investigate physical properties of various systems. However, when dealing with phase transitions or critical phenomena at low temperatures, the Metropolis algorithm becomes inefficient due to inadequate sampling [41]. For example, at low temperature, where the system has low energy, the Metropolis algorithm is easily trapped in local energy minima since the Boltzmann factor is suppressed ($e^{-\Delta E/k_B T} \longrightarrow 0$, E is system energy.), when trying to move from low to higher energy state. In second order phase transitions, the Metropolis algorithm takes prohibitively long time to achieve accurate estimate of system behavior near the transition region, as standard deviation of the canonical energy distribution function ($\sigma_E = \sqrt{\langle E^2 \rangle - \langle E \rangle^2}$) is very large, corresponding to the maximum

specific heat ($C_v = \sigma^2/k_B T^2$), at critical temperature. Thus, one has to sample a large energetic domain [40, 41]. This problem has been mediated by conducting cluster correlated moves, which accelerate the convergence [82]. First order phase transitions are usually presented by coexistence of two phases and the energy barrier between these phases increases with system size. In these cases, the Metropolis MC algorithm also needs lengthly time to tunnel from one energy minimum to the other and easily gets trapped. Therefore, convergence may not be achieved at all.

Wang and Landau [42, 43] introduced a new MC algorithm to produce flat energy histograms. The energy states are sampled with probability proportional to the reciprocal of the density of energy states, which is self-adjusting and continuously updated throughout the long simulation (until it is considered invarient). In this algorithm, all accessible energy states (both more and less favorable) seem to be sampled more efficiently, compared to the conventional MC algorithm. The WL algorithm has been reported to be very efficient when applied to systems with discrete energy spectrum such as Ising model and Potts model [42, 43]. However, its efficiency decreases, when it is used on systems with continuous energy spectrum such as proteins, polymers and liquid crystals [44–47, 83], as the number of states of systems with continuous energy spectrum becomes extremely large. Kim *et al.* [48] combined the WL algorithm [42, 43] and the expanded algorithm [84] into the combined algorithm, called the expanded density of states (EXEDOS) [48], which could be applied on any reaction coordinates, not just energy. Good performance of the EXEDOS algorithm was reported when tested on various problems such as solidification of particles [49], calculation of potential of mean force between counterions and charged surfaces [50], and transition between coil-bridge states of a self avoiding chain attracted to a surface [51]. More detailed description of the EXEDOS algorithm is given in the next chapter.

1.3 Literature review

1.3.1 Nematic liquid crystals under spherical confinement

When nematic liquid crystals are confined to a curved surface, their directors become distorted in certain regions due to the geometrical constraint, leading to formation of spots where their directors are undefined, called defects. Lubensky and Prost [85]

theoretically predicted three possible molecular arrangements of nematic LCs on a spherical surface: splay (Fig. 1.7 b)), toroidal (Fig. 1.7 a)) and tetrahedral (Fig. 1.7 c)). The splay structure has two +1 defects located at each pole. The nematic field radiates from one pole and circles the sphere merging at the opposite pole. Toroidal structure also has two +1 defects, but the nematic field spirals around the poles as it moves from one pole to another. Tetrahedral structure has four +1/2 defects at vertices of tetrahedron [85]. Such a tetrahedral structure was observed in MC simulations [34] and experiments [33] of a thin uniform nematic shell. In a thin and non-uniform nematic shells, locations of the four defects no longer coincide with the vertices of tetrahedron [32, 34]. At intermediate shell thicknesses, the structure with two +1/2 defects and one +1 defect was also occasionally observed [34]. In a droplet or a thick nematic shell, a structure with two +1 defects was observed [32, 34].

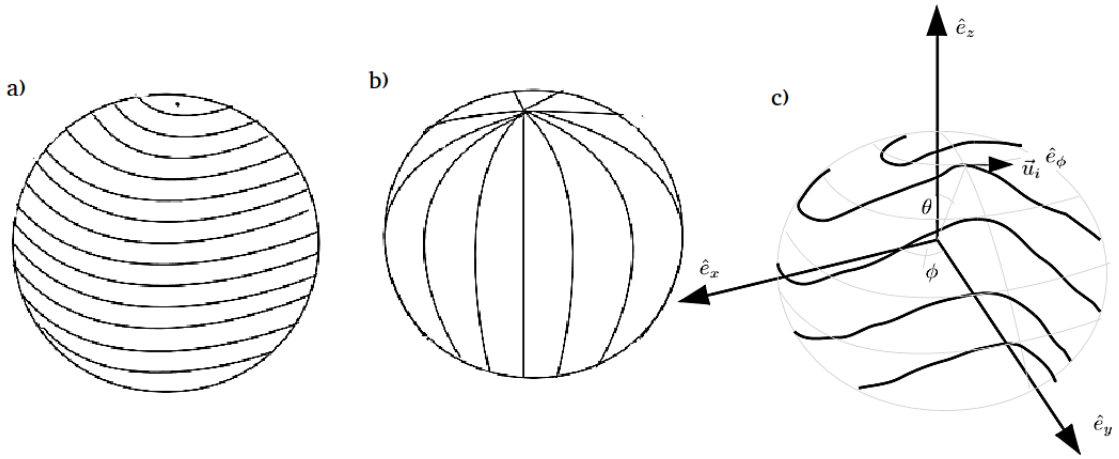

Figure 1.7: Possible molecular arrangements of nematic LCs on a spherical surface as theoretically predicted by Lubensky *et al.* [85]: a) toroidal, b) splay and c) tetrahedral structure.

In theory, it may be possible to use defect locations in tetrahedral structure as bonds, similar to sp^3 hibridized atoms of carbon group [16], for molecular attach-

ment. If achieved in practice, this property can be useful for biosensing applications to amplify signal for better detection. Confinement of LCs can also increase their surface to volume ratio, leading to an improvement of their elastic properties useful in microfluidic applications and in artificial muscles [86]. The examples above show that a problem of nematic LC confinement is a subject of interdisciplinary research relevant to applications in many areas of human life.

1.3.2 Spherically confined polymers

Confinement of a polymer chain to a restricted geometry greatly reduces its accessible conformations. The competition between entropic and enthalpic interactions determines the spatial organization and physical properties of the chain molecule, which can be very different from that in the bulk [87–90]. For example, human DNA, which is approximately 2 meters long in unconfined disordered state occupies the volume of about 0.2 mm radius (called radius of gyration), but nicely packs inside a nucleus with diameter of about 6 μm [18]. Stability of proteins also noticeably increases under strong confinement, compared to that in bulk [89, 90].

A flexible chain confined inside a sphere has been most often studied in a concept of scaling laws [91–94], which can be used to predict universal behaviors of molecules irrespective of their chemical constitution. This approach can be used in both experiments and simulations to elude the dominant mechanisms and relevant parameters. For example, the confinement free energy (F) of a self-avoiding flexible chain inside a sphere of approximately the same size as its unconfined size ($r_c \sim r_G$) is described by a power-law: $\beta\Delta F \sim (1/r_c)^x$, where x is found from blob theory to be 3.93 [53] and in MC simulations [91, 92], it falls in the range 3.44-3.8 for a chain in good solvent. At higher concentrations, the same chain needs more energy to pack inside a smaller sphere; thus, the value of x can increase up to 5.91, as was found in some MC simulations [91].

Confinement of a semi-flexible chain inside a small sphere is closely relevant to the problem of biomolecular confinement in a cell. As mentioned earlier, 3-d structure of the confined biomolecules is closely related to their function. Several studies were dedicated to investigation of packing thermodynamics and geometry of biomolecules in equilibrium [3, 28–30]. For example, simulation of a semi-flexible chain inside a sphere shows that the chain begins transitioning from coil-like to

toroidal structure, when the radius of a confining sphere is of an order of persistence length (l_p) [3, 28–30]. As the confining radius becomes much smaller than l_p, a spool-like structure becomes preferred [31]. The same structure was also observed inside the bacteriophage T7 DNA by cryo-electron microscopy [23], and predicted by simulations [24, 25]. On the contrary, particle dynamics simulations by Fathizadeh *et al.* predicted that the tennis ball structure with four +1/2 defects was formed when the T7 DNA was compressed into a sphere [25]. The tennis ball structure was also observed in MC simulations of a self-avoiding semi-flexible chain, densely confined on a sphere [95], similar to that found in other simulations [34], and experiments of rigid molecules confined in a thin spherical shell [32].

Investigation of monomer distribution inside a sphere is also related to a problem of chromosomes distribution, which affects DNA expression [96]. In lymphocytes, for example, gene rich chromosomes cluster in the interior of the nucleus, while inactive heterochromosomes (which are often involved in formation of tumor cells) aggregate at the surface [97]. Thus, understanding the mechanism of chromosome arrangement inside nucleus, may shed light on detection of early stages of formation of tumor cells [97].

1.4 Outline of the thesis

The thesis is composed of 6 chapters and a brief description of each following chapter follows below.

In Chapter 2, theory of the traditional Metropolis MC method and its weakness will be discussed. The original Wang-Landau (WL) algorithm which has been used to investigate physical properties of systems with discrete energy spectrum will be addressed, followed by description of the expanded ensemble density of states (EXEDOS) algorithm, which combines the WL and the expanded ensemble algorithm. Finally, we will address the simulation protocol and analysis methods that were applied in this work.

In Chapter 3, the EXEDOS algorithm will be used on a system of a flexible chain of different lengths confined inside a sphere of variable radius to test applicability and limitations of the algorithm. The physical properties obtained from the simulations will be analyzed using scaling laws, and compared to results from other

authors.

In Chapter 4, the EXEDOS algorithm will be applied to a more complex system of a self-avoiding semi-flexible chain confined inside a sphere to investigate the effects of density changes on molecular arrangement of the chain with different flexibilities using a concept of order parameters borrowed from studies of liquid crystals. Characteristics of polymer layer in the vicinity of a confining surface and defects formed in this layer in dense phases will also be addressed.

In Chapter 5, the same algorithm will be applied to the system of a semi-flexible polymer chain with fixed persistence length, and variable monomer diameter. The goal of this chapter is to investigate the influence of molecular flexibility, controlled by changes of monomer diameter, on the structure and defect arrangement formed in dense phase of a semi-flexible chain. Results of this chapter will be analyzed in the same framework as in Chapter 4.

Chapter 6 will conclude the findings of this thesis, and outline the prospects for the future work.

Chapter 2

Monte Carlo simulations

2.1 Monte Carlo simulations in statistical physics

The Monte Carlo method was proposed by Fermi, Ulam, von Neumann and Metropolis to study diffusion of neutrons through fissile material back in the 1940s [39]. Later, the method was used in many applications in statistical mechanics, because it allowed one to estimate the ensemble averages of physical quantities in a more efficient manner, than directly calculating them from theoretical principles of statistical mechanics.

Consider a closed system with a fixed number of particles (N), temperature (T) and volume (V). In theory, the average of a quantity A in equilibrium can be calculated from the Boltzmann distribution as:

$$\langle A \rangle = \sum_n A_n P_n, \tag{2.1}$$

where n denotes a state, A_n is a value of A in that state, and P_n is probability that the system would be in state n, expressed as:

$$P_n = \frac{e^{-E_n/k_B T}}{Z}, \tag{2.2}$$

where E_n is the energy of a system in state n, k_B is the Boltzmann constant, T is the absolute temperature, and Z is partition function, which can be written as:

$$Z = \sum_{\alpha} e^{-E_\alpha / k_B T}. \tag{2.3}$$

The sum runs over all configurations of the system. The number of configurations grows exponentially with a number of particles N, and for most systems of interest it is neither practical, nor possible to calculate the partition function numerically. The MC method samples a small group of states in equilibrium, instead of exploring all possible states of the system, to estimate the average value of quantity A. The average is determined as:

$$\langle A \rangle = \frac{1}{N_s} \sum_{i=1}^{N_s} A_i, \tag{2.4}$$

where N_s is the total the number of samples calculated during the simulation. MC algorithm uses probabilistic approach to generate configurations that lead the system close to equilibrium, and we will describe in the subsequent section how this probability is chosen.

2.2 Monte Carlo methods

In this section, we will show how a new configuration is updated to lead a system of interest towards the equilibrium, in different MC algorithms. Consider a system in state n. During the simulation, its evolution in phase space is governed by the master equation.

$$\frac{dP_n}{dt} = -\sum_{n \neq m} (\omega_{n \longrightarrow m} P_n(t) - \omega_{m \longrightarrow n} P_m(t)), \tag{2.5}$$

where $P_n(t)$ and $P_m(t)$ are probabilities for the system to be in states n, and m , respectively; t is simulation time, $\omega_{n \longrightarrow m}$ and $\omega_{m \longrightarrow n}$ are transition probabilities between these two states. In equilibrium, the system occupies states with highest probability of occurrence, and $dP_n/dt = 0$; thus, Eq. 2.5 becomes:

$$\sum_{n \neq m} (\omega_{n \longrightarrow m} P_n - \omega_{m \longrightarrow n} P_m) = 0. \tag{2.6}$$

The sum above can only be equal to zero if all its terms are also equal to zero; which gives the following equilibrium condition:

$$\omega_{n\longrightarrow m}P_n = \omega_{m\longrightarrow n}P_m. \tag{2.7}$$

Eq. 2.7 is called the detailed balance equation, and it shows that transition between states n and m in equilibrium are equally probable in both directions. When using a Markov chain of states, each new state m is directly created from the previous state n, and both n and m states have a common partition function Z. In this case, substituting Eq. 2.2 into Eq. 2.7 gives:

$$\frac{\omega_{n\longrightarrow m}}{\omega_{m\longrightarrow n}} = \frac{P_m}{P_n} = e^{\frac{-\Delta E}{k_B T}}, \tag{2.8}$$

where $\Delta E = E_m - E_n$. Eq. 2.8 states that the acceptance rate does not depend on the partition function Z, but only the Boltzmann factor.

2.2.1 The Metropolis Monte Carlo algorithm

In MC algorithms, a system transitioning from one state to another is simulated by proposing an update to a new state, and accepting this state or rejecting following a certain set of rules. The probability $\omega_{n\longrightarrow m}$ of a system transitioning from state n to m is written as:

$$\omega_{n\longrightarrow m} = W^p_{n\longrightarrow m} \cdot W^a_{n\longrightarrow m}, \tag{2.9}$$

where $W^p_{n\longrightarrow m}$ and $W^a_{n\longrightarrow m}$ are probabilities of proposing and accepting a move from state n to m. The Metropolis algorithm uses symmetric proposal probabilities; $W^p_{n\longrightarrow m} = W^p_{m\longrightarrow n}$. Thus, Eq. 2.8 gives us the following relation for acceptance probabilities :

$$\frac{W^a_{n\longrightarrow m}}{W^a_{m\longrightarrow n}} = \frac{P_m}{P_n} = e^{\frac{-\Delta E}{k_B T}}. \tag{2.10}$$

In Eq. 2.10, both probabilities of acceptance lie between 0 and 1. If the state m has higher energy that n, $E_m > E_n$, Metropolis algorithm sets the acceptance probability $W^a_{m\longrightarrow n}$ equal to 1. Probability $W^a_{n\longrightarrow m}$ then takes a value of $e^{-\Delta E/k_B T}$, which is less than 1. Then Eq. 2.8 can be rewritten with the Metropolis's choice as:

$$W^a_{n\longrightarrow m} = min(e^{\frac{-\Delta E}{k_B T}}, 1). \tag{2.11}$$

If an updated configuration has lower energy than the previous one, it is always accepted, while if it has higher energy, it may still be accepted with probability $e^{\frac{-\Delta E}{k_B T}}$. In practice, the decision for accepting or rejecting the new configuration is made by generating a pseudo-random number ranging from 0 to 1, and comparing it to the Boltzmann factor $e^{\frac{-\Delta E}{k_B T}}$. If the number is lower than or equal to $e^{\frac{-\Delta E}{k_B T}}$, the new configuration is accepted, otherwise it is rejected. If the new configuration is accepted, information about the previous configuration, such as coordinates and energy, is replaced by that representing the new configuration. After the system reaches equilibrium, the statistical average of any quantity A can be determined using Eq. 2.4.

As mentioned in Chapter 1, the Metropolis MC method is less efficient due to inadequate sampling when applied to systems at low temperature, systems with rugged energy landscape and those undergoing phase transition [41]. Several algorithms such as multicanonical [98–101] and entropic samplings [102] were proposed to overcome these problems by artificially altering potential to enhance probability of transition between energy states [98–102]. However, the potential or the weight factor of these algorithms is not known a priori and their implementation is very time consuming [103]. With these algorithms, the simulation usually starts with running several replica exchanges to obtain an estimate for density of states using the Weighted Histogram Analysis Method (WHAM) [104]. The density of states is then used as a weighing factor in a multicanonical simulations. This generates an energy histogram from which a new density of states is obtained by WHAM and the next cycle starts again. Once, a good estimate of the density of states is obtained, the final multicanonical simulation is conducted to produce thermodynamic quantities of interest.

2.2.2 Wang-Landau Monte Carlo algorithm

Wang and Landau [42, 43] introduced an algorithm in which the acceptance probability is inversely proportional to the density of energy states $g(E)$. This density-of-states does not have to be priorly known and converges self-consistently to a true density-of-states, which can be obtained within one simulation. The concept of Wang-Landau (WL) and muticanonical methods are similar: the canonical Boltzmann distribution ($P(E,T) \propto g(E) exp(-E/k_B T)$) is modified by a weight factor

$w(E,T)$ so that every energy level has equal probability to be visited during the simulation.

$$w(E,T)P(E,T) \sim h(E) \equiv constant, \tag{2.12}$$

where $h(E)$ is (ideally flat) energy histogram. Temperature in Eq. 2.12 does not have any meaning, and the energy distribution is always constant, independent of temperature. It is convenient to set $\lim_{T\longrightarrow\infty} P(E,T) \sim g(E)$ and thus $\lim_{T\longrightarrow\infty} w(E,T) \sim 1/g(E)$. With this weight factor and the Metropolis condition, the acceptance criterion becomes:

$$W^{a}_{n\longrightarrow m} = min(\frac{g(E_n)}{g(E_m)}, 1). \tag{2.13}$$

After each trial move, the density-of-states is changed by $g(E) = c^n \cdot g(E)$, where c is called modification factor. The value $c > 1$ is kept constant throughout the nth iteration, but decreases from iteration to iteration. Usually, c is reduced between iterations as: $c^n = \sqrt{c^{n-1}}$, where $n = 1, 2, 3, ..., I$ and $c^0 = exp(1) = 2.718$. The simulations continue until $c \to 1$ at iteration I, at which point $g(E)$ is considered well converged.

1. Define energy range of interest, $[E_{min}, E_{max}]$.

2. Divide the energy range into a number of bins.

3. Start with $g(E) = 1$ and $h(E) = 0$ for all E.

4. Generate a random walk in energy space with acceptance criterion: $W^a_{n \longrightarrow m} = min(\frac{g(E_n)}{g(E_m)}, 1)$.

5. Modify $g(E)$ and $h(E)$ after every trial move. Set $h(E) = h(E) + 1$ and $g(E) = g(E) \cdot c$, where c is modification factor, initially set to e^1.

6. Check flatness of the energy histogram. If $h(E)$ is sufficiently flat, set $h(E) = 0$ and $c = \sqrt{c}$.

7. Continue until $c \sim 1$.

Table 2.1: The Wang-Landau algorithm.

In practice, the WL algorithm is performed on a system of interest using the algorithm shown in Table 2.1. First, the range of energy spectrum is selected for a system. This energy range is then divided into a number of bins. Two histograms are constructed for density of energy states ($g(E)$) and number of visits ($h(E)$). The initial density-of-states ($g(E)$) is set to unity, and initial histogram is set to zero. Initial modification factor (c) is set to $exp(1)$. If the initial state of the system is described by a position in phase space r_n and energy E_n, new position r_m with energy E_m is obtained by changing system configuration. The new configuration m is always accepted if its density-of-states $g(E_m)$ is smaller than $g(E_n)$; otherwise, if $g(E_m) > g(E_n)$, it may be accepted with probability $g(E_n)/g(E_m)$. The decision of accepting or rejecting configuration m is made by generating a pseudo-random number in a range of [0,1), and comparing it to the factor $g(E_n)/g(E_m)$. If the number is smaller than or equal to $g(E_n)/g(E_m)$, the configuration m is accepted, otherwise it is rejected. If the configuration m is accepted, information about the configuration n, such as coordinates and energy, is replaced by that corresponding

to the configuration m. The histogram $h(E_m)$ is incremented by one, and the density-of-states ($g(E_m)$) is modified as $g(E_m) = g(E_m) \cdot c$. If the new configuration is rejected, the value of $h(E_n)$ is incremented by one, and $g(E_n)$ is modified as $g(E_n) = g(E_n) \cdot c$. The cycle is repeated until the histogram becomes sufficiently flat: $|\forall\ h(E) - \langle h(E) \rangle| \leq (1 - K) \cdot \langle h(E) \rangle$, where $K \in (0.5 - 0.99)$, and $\langle h(E) \rangle$ is an average number of visits over all histogram bins. The whole histogram is then emptied and the density-of-states is normalized by a constant value $g(E_0)$. The normalized $g(E_i)$ is then used as an initial density-of-states for the next cycle, with smaller modification factor c as: $c = \sqrt{c}$. The WL algorithm proceeds until $c \approx 1$, at which point $g(E)$ is considered to be converged.

From converged $g(E)$, it is possible to determine partition function as:

$$Z = \sum_E g(E) e^{-E/k_B T}. \tag{2.14}$$

Eq. 2.14 indicates that the partition function Z can be determined by summation over limited number of configurations in each energy level, instead of collecting all possible configurations in the ensemble, as done in Eq. 2.3. Once, Z is found, most thermodynamic quantities are readily calculated. For instance, the internal energy $E(T)$ can be determined as:

$$\langle E \rangle = \frac{\sum_E E g(E) e^{-E/k_B T}}{\sum_E g(E) e^{-E/k_B T}}. \tag{2.15}$$

The Wang-Landau samplings have proven to be very efficient, when performed on systems with discrete energy spectrum such as Ising model and Potts model [42, 43]. The method was also successfully employed to study small systems with continuous energy spectrum [44, 45, 47, 105]. Phase equilibrium between vapor and liquid in the Lennard-Jones fluid, for example, was successfully studied by the WL method [106, 107]. The method was also applied to study the coil-globe transition [83] and the probability distribution of end-to end distance [46, 108] of a flexible polymer. Several research groups also conducted WL simulations to study folding process of a simple peptide model [44, 45, 105].

The disadvantage of the WL method is its poor convergence when applied on large system. Due to a large number of states in systems with continuous energy

spectrum, the algorithm requires exceedingly long time to converge. Furthermore, when a system is sufficiently complex, the algorithm needs incorporation of advanced moves, special rules for reducing modification factor or adding configurational-bias algorithm, depending on physical properties of the system [45, 47, 109]. In the subsequent subsection, we will discuss the algorithm that combines the Wang-Landau [42] and the expanded density-of-states [84] algorithm to study the systems with continuous energy spectra.

2.2.3 Expanded ensemble density-of-states Monte Carlo algorithm

The expanded ensemble density-of-states (EXEDOS) [48] method is based on WL algorithm [42], but it can be applied on any reaction coordinate, rather than only energy, similar to the expanded ensemble algorithm [84]. The EXEDOS method was first proposed by Kim *et al* [48] to calculate potential of mean force (PMF) between a particle suspended in liquid crystal and a wall. Later, the EXEDOS has been successfully used to study systems with continuous energy spectrum, such as solid crystallization [49], proteins [110], and colloids [50, 111]. Due to these advantages of the EXEDOS method, we chose it for our work, and description of this method will be given below.

Consider a classical system of N particles with a volume V, at temperature T, characterized by equilibrium distribution $P_\xi(N, V, T, \xi(r))$, where $r = \{x_i, y_i, z_i\}$, and $\xi(r)$ is a reaction coordinate of interest divided into M segments: $\xi = \xi_1, \xi_2, ..., \xi_M$, each segment representing a single state. Partition function Z_ξ of this expanded ensemble can be expressed as:

$$Z_\xi = \sum_{m=1}^{M} Z_m(N, V, T, \xi_m) g_m, \tag{2.16}$$

where, g_m is weighting factor, and Z_m is a canonical partition function for a system in state ξ_m. The probability that the system will visit a state ξ_m can now be written as:

$$P_m = \frac{Z_m g_m}{Z_\xi}. \tag{2.17}$$

For a system transitioning from state ξ_n to ξ_m, the acceptance probability becomes:

$$W^a_{n\longrightarrow m} = min(\frac{g_n}{g_m} \cdot e^{\frac{-\Delta E}{k_B T}}, 1), \tag{2.18}$$

where $\Delta E = E_m - E_n$. For athermal system, in which $E = 0$, Eq. 2.18 becomes:

$$W^a_{n\longrightarrow m} = min(\frac{g(\xi_n)}{g(\xi_m)}, 1). \tag{2.19}$$

---

1. Define a range of values of the reaction coordinate of interest, $[\xi_{min}, \xi_{max}]$.

2. Divide the range into a number of bins.

3. Start with $g(\xi) = 1$ and $h(\xi) = 0$ for all ξ.

4. Generate a random walk in configuration space with acceptance criteria:
   $W^a_{n\longrightarrow m} = min(\frac{g(\xi_n)}{g(\xi_m)} \cdot e^{\frac{-\Delta E}{k_B T}}, 1)$, where $\Delta E = E_m - E_n$

5. Modify $g(\xi)$ and $h(\xi)$ after every trial move.
   Set $h(\xi) = h(\xi) + 1$ and $g(\xi) = g(\xi) \cdot c$, where c is modification factor, initially set to e^1.

6. Check flatness of histogram.
   If $h(\xi)$ is sufficiently flat, set $h(\xi) = 0$ and $c = \sqrt{c}$.

7. Continue until $c \sim 1$.

Table 2.2: The Expanded Ensemble Density-of-States algorithm.

Table 2.2 summarizes the EXEDOS algorithm. To perform the EXEDOS simulation, we first divide a reaction coordinate ξ into M bins, each bin corresponding to a single configuration state. Two histograms for the density of configuration states ($g(\xi)$) and the number of visits ($h(\xi)$) are constructed in the way, similar to that used in WL simulations. When the system is probed for transition from state ξ_n to ξ_m, Eq. 2.18 or Eq. 2.19 is used to decide if state ξ_m is accepted or not. If the state is accepted, the corresponding histograms $g(\xi_m)$ and $h(\xi_m)$ are updated as: $g(\xi_m) = g(\xi_m) \cdot c$ and $h(\xi_m) = h(\xi_m) + 1$. If the state ξ_m is rejected, $g(\xi_n)$ and $h(\xi_n)$

are modified as: $g(\xi_n) = g(\xi_n) \cdot c$ and $h(\xi_n) = h(\xi_n) + 1$. The modification factor c is initially set to $exp(1)$. The next simulation step follows WL method, until good convergence in $g(\xi)$ is achieved. This estimated $g(\xi)$ can be used in Eq. 2.18 for calculation of an acceptance probability during the production runs for measuring a quantity of interest.

2.3 General simulation protocol

In this section, we will describe how the EXEDOS simulations were conducted in present work. Before describing the simulation protocol in details, we will first describe interaction potentials and polymer models used in the our simulations.

2.3.1 Interaction potential

In order to add excluded volume interactions to a non-ideal chain, we used a hard-sphere potential (E_e) described as:

$$E_e = \begin{cases} 0, & \text{if } r_{ij} \geqslant d \\ \infty, & \text{otherwise,} \end{cases} \tag{2.20}$$

where r_{ij} is the distance between monomer i and j, and d is monomer diameter. For semi-flexible chains, bending potential (E_b) was also added, expressed here as [112]:

$$E_b = k_b \sum_{i=1}^{N-1} (1 - cos\theta_i). \tag{2.21}$$

If N is the total number of bonds in the chain of $N+1$ monomers, θ_i is the angle between $\vec{u}_i = \vec{r}_{i+1} - \vec{r}_i$ and $\vec{u}_{i+1} = \vec{r}_{i+2} - \vec{r}_{i+1}$, as illustrated in Fig. 2.1 ($\vec{r}_i$, $\vec{r}_{i+1}$ and $\vec{r}_{i+2}$ are position vectors of monomer i, $i+1$, $i+2$, respectively), and k_b is bending constant defined in elasticity theory for a polymer chain in 3-dimensions as: $k_b = l_p \cdot k_B T/a$ [35, 56, 58, 59].

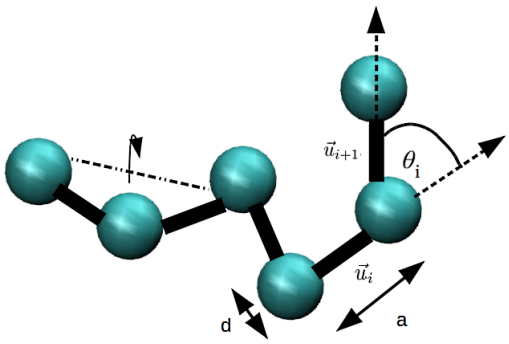

Figure 2.1: A freely-jointed chain. θ_i is the angle between bonds i and $i+1$, d and a are bead diameter and bond length, respectively.

2.3.2 Generating polymer chains

In this section, we will describe how polymer models were generated for our simulations. We will first describe a freely-jointed chain model used in Chapter 3-5, followed by a fused-sphere chain used in Chapter 5.

2.3.2.1 Freely-jointed chain

Consider a freely-jointed chain, comprised of $N+1$ hard-spheres of diameter d (indexed from $i = 1$ to $i = N+1$), connected with N bonds, each bond having length $l/a = 1$, as shown in Fig. 2.1. Position of each monomer is described by a vector $\vec{r}_i$. A unit vector describing bond i is defined by $\hat{u} = (\vec{r}_{i+1} - \vec{r}_i)/|\vec{r}_{i+1} - \vec{r}_i|$.

To generate the chain, the coordinate of the first monomer was assigned at $(x, y, z) = (0, 0, 0)$. The next monomer was then added by randomly selecting the azimuthal angle (ϕ) in the range of $(0, 2\pi)$ and the cosine value of the polar angle (θ) in the range of $(-1, 1)$. The angle ϕ and the $cos(\theta)$ value were then converted into x, y, z positions of a monomer, and the monomer was accepted if separation between monomers in the chain was larger than or equal to the diameter. For an ideal chain, monomer-monomer overlap was allowed. The process was repeated until required length of the chain was achieved.

To generate a discrete worm-like chain, the 10^7 (for chains with $l_p/a \leq 16$) - 10^8 (for chains with $l_p/a = 24$ and 32) Metropolis MC steps, each MC step made of

$N+1$ trial moves, were applied to the chains using bending potential (E_b) described by Eq. 2.21 as an interaction potential. For a self-avoiding semi-flexible chain, the hard-sphere potential in Eq. 2.20 was also added. A monomer in each chain was randomly selected and rotated around the axis connecting its adjacent neighbors, as shown in Fig. 2.1. The trial move was accepted, if new monomer position did not overlap with any other monomers (for non-ideal chain) and the criterion in Eq. 2.10 was satisfied. The simulations proceeded until the required convergence was achieved.

2.3.2.2 Fused-sphere chain

A chain of $N+1$ fused-hard spheres shown in Fig. 2.2 (each having diameter d), connected with N bonds (each having length $l/a = 1$), is used in Chapter 5. To generate a chain, positions of the first two monomers were assigned. The angle ϕ was randomly selected in the range of $(0, 2\pi)$. The cosine of θ was randomly selected in the range of $(0, cos(\theta_{max}))$, where θ_{max} is the maximum angle between two adjacent bonds formed by monomers i, $i+1$, and $i+2$ (achieved when spheres i and $i+2$ are touching each other, as seen in Fig. 2.2). The coordinate of monomer was then calculated from the selected ϕ and $cos(\theta)$. If there were no overlap between nonadjacent spheres, position of the new monomer was accepted. The process was repeated until the chain had a required length.

Bending stiffness was introduced into the chain, by using the bending potential (Eq. 2.21) as an interaction potential. In the Metropolis simulations, a monomer was randomly selected and moved by rotating around the axis connecting its adjacent neighbors. The trial move was accepted if nonadjacent monomers did not overlap and the criterion in Eq. 2.10 was satisfied.

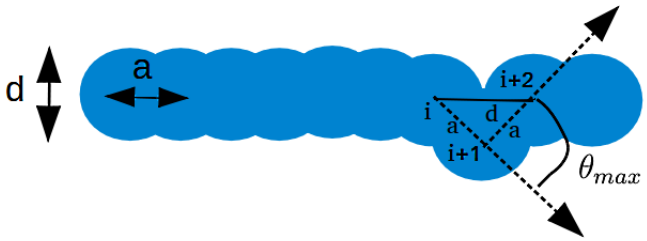

Figure 2.2: A fused-hard sphere chain. The symbols d, a are bead diameter and bond length, respectively, and θ_{max} is the maximum angle between 2 bonds.

2.3.3 Density-of-states in EXEDOS

After chain models were generated, the range of confining radii (r_c) was chosen as a reaction coordinate for each chain, and assigned specifically for each model. This range was then divided into a number of shells as shown in Fig. 2.3 a). Each shell reflects a separate configuration state. Each chain was then placed inside a sphere by putting the chain center of mass at the center of the sphere.

Two sets of histograms were constructed showing the number of times the furthermost monomer, relative to center of the sphere, of the confined chain visited each shell ($h(r_c)$) and density of corresponding configuration states ($g(r_c)$). Initially, the number of visits in each state was set to zero ($\forall\ h(r_c) = 0$), while the density-of-states in each shell was assumed to be unity. To avoid numerical overflow, the density-of-states was used in logarithmic form. Thus, the initial logarithmic density-of-states in all shell was zero, $\forall\ ln\ g(r_c) = 0$.

During the simulation, a single monomer was randomly selected and moved by rotating it about the axis connecting two neighboring monomers. For a non-ideal freely-jointed chain and a non-ideal discrete worm-like chain, the trial move was rejected if separation between monomers became closer than diameter d. For a non-ideal fused-sphere chain, the trial move was rejected if there was overlap between nonadjacent monomers. The listed cell method [39] was applied when calculating the separation distance between monomers, to save computational time. If the separation after the move was larger than or equal to diameter d, the trial move from old to new state was accepted only if the condition of Eq. 2.18 was satisfied. As

mentioned, the density-of-states was obtained in logarithmic form; hence, Eq. 2.18 was modified as:

$$\begin{aligned} P_{acc}(old \rightarrow new) &= min[e^{\frac{-(E_{new}-E_{old})}{k_B T}} \cdot e^{-(ln\ g_{new} - ln\ g_{old})}, 1] \\ &= min[e^{-\beta \Delta E - \Delta ln\ g}, 1]. \end{aligned} \tag{2.22}$$

Where β=$1/k_BT$, E_{new} and E_{old} are total energies of the system in the new and old states, respectively. The term ΔE was set equal to zero, when the algorithm was applied to athermal chains (Chapter 3). If Eq. 2.22 was satisfied, the number of visits to the bin corresponding to the new state was incremented by one $H_{new} = H_{new} + 1$, and the density-of-states was modified by a factor c: $ln\ g_{new} = ln\ g_{new} + ln\ c$. If the trial move was rejected, the number of visits in the old bin was incremented by one, $H_{old} = H_{old} + 1$, and the density-of-states was modified as: $ln\ g_{old} = ln\ g_{old} + ln\ c$. The trial move was also rejected if it was out of the confining range set for each chain. The initial modification factor was set to be $c = exp(1)$ or $ln\ c = 1$.

Once, the histogram $h(r_c)$ was sufficiently flat ($| \forall\ h(r_c) - \langle h \rangle| \leq (1 - K)\ \cdot \langle h \rangle$, where $K \in (0.6 - 0.85)$, is specified separately in each chapter. $\langle h \rangle$ is the number of visits averaged over all histogram bins); the histogram $h(r_c)$ was reset to zero, while the density-of- states $g(r_c)$ was normalized by a constant and became an initial condition for the next cycle. In each subsequent cycle, the modification factor stepped down as $c = \sqrt{c}$ or $ln\ c = 0.5 \cdot ln\ c$. The simulations continued until the modification factor was smaller than $ln\ c = 10^{-8}$ ($c \sim 1$). The obtained density-of-states was substituted into Eq. 2.22, and this equation was then used as an acceptance criterion during production runs.

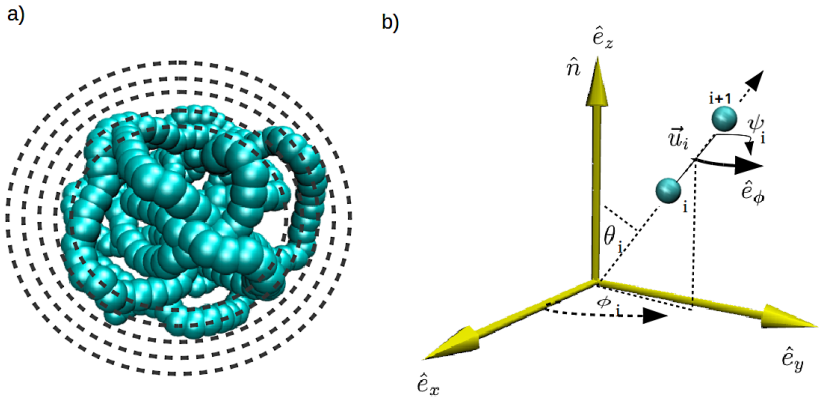

Figure 2.3: a) Polymer chain in a confining sphere, divided into a number of concentric shells. b) Coordinate system, used to define positions and orientation of bonds, $\hat{n}$ is nematic director, which is a unit vector in the direction of the largest eigenvalue of the tensor $\mathbf{T}$, constructed from all bonds in a chain. ψ_i is the angle between $\hat{\phi}$ and a bond $\vec{u}_i$, where $\vec{u}_i = \vec{r}_{i+1} - \vec{r}_i$, $\vec{r}_{i+1}$ and $\vec{r}_i$ are position vectors of monomer $i+1$ and i, θ is the angle between the bond $\vec{u}_i$ and $\hat{n}$, ϕ is the angle between $\hat{e}_x$ and the projection of $\vec{u}_i$ on x-y plane.

After convergence of the density-of-states, it was normalized so that $\int g(r_c) 4\pi r_c^2 dr_c = 1$. Then, the free energy of the chain confined inside a sphere of radius r_c, can then be determined using the relation:

$$e^{\frac{-F(r_c)}{k_B T}} \propto P(r_c), \tag{2.23}$$

where $P(r_c) = 4\pi r_c^2 g(r_c)$. Eq. 2.23 can be written as:

$$F(r_c) = -k_B T \ln P(r_c) + Constant, \tag{2.24}$$

For a semi-flexible chain, free energy difference between the configuration ξ_m and ξ_n can be calculated as (using Z_m from Eq. 2.17):

$$\begin{aligned} \Delta F &= -k_B T \ln \frac{Z_m}{Z_n} \\ &= -k_B T (\ln \frac{g(\xi_n)}{g(\xi_m)} + \ln \frac{P_m}{P_n}). \end{aligned} \tag{2.25}$$

When convergence of the density-of-states is achieved, the second term in Eq. 2.25 is zero since the probabilities of finding configurations ξ_m and ξ_n are equal.

2.3.4 Structural analysis

In this section, we will describe how we calculate the density profiles of the chains, and address the structural order parameters that could distinguish between disordered and ordered phases, and test the structure for a presence of different arrangements at various degrees of confinement.

2.3.4.1 Number density (n)

To find density profile (n) as a function of distance from center of a sphere (r), a confining sphere (r_c) was first divided into a number of shells (Fig. 2.3 a)), each shell having width of Ka, where a is a bond length, and $K \in (0.2 - 0.5)$ is a constant specified for each considered case. The number of monomers located in each shell was counted and divided by volume of the shell and then normalized to give $\int n(r) 4\pi r^2 dr = 1$.

2.3.4.2 Nematic order parameter (Q)

The nematic order parameter is used to test the liquid crystals the presence of orientational order as mentioned in Chapter 1. When a semi-flexible chain is confined inside a sphere of size much smaller than its persistence length, it arranges itself into an orientationally ordered structure, which is similar to that observed in nematic liquid crystals. We use Q to probe for orientational order in a spherically confined polymer in Chapter 4 and 5.

To obtain Q, a traceless second rank tensor $\mathbf{T}$ was formed as:

$$\mathbf{T} = \frac{1}{2N_b} \sum_{i=1}^{N_b} (3\hat{u}_{i\alpha}\hat{u}_{i\beta} - \delta_{\alpha\beta}), \tag{2.26}$$

where $\hat{u}_{i\alpha}$ is an α component ($\alpha = x, y, z$) of a bond i, defined as $\hat{u}_i = (\vec{r}_{i+1} - \vec{r}_i)/|\vec{r}_{i+1} - \vec{r}_i|$, where $\vec{r}_{i+1}$ and $\vec{r}_i$ are position vectors of monomers i and $i+1$ and $\delta_{\alpha\beta}$ is Kronecker delta. N_b is the total number of surface bonds, located between the confining surface and the shell holding the highest number of monomers, such

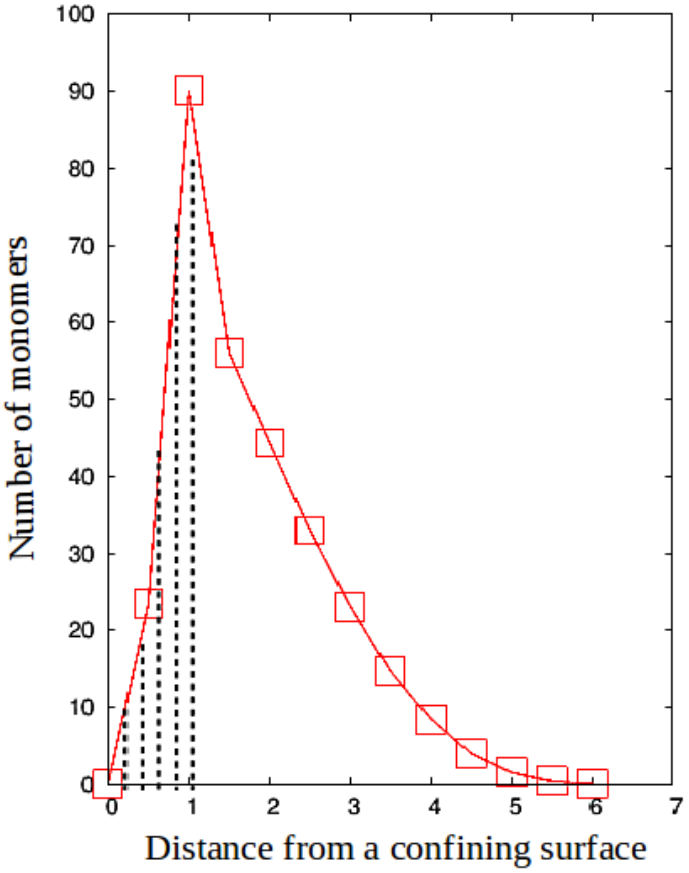

Figure 2.4: Number of monomers as a function of distance from the confining surface. Shaded area indicates the total number of surface bonds used in calculation of order parameters.

as that illustrated by the shaded area in Fig. 2.4. Eigenvalues λ_+, λ_0 and λ_- were obtained by diagonalization of $\mathbf{T}$, and the largest of λ_+, λ_0 and λ_- was considered to be a nematic order parameter (Q). The eigenvector corresponding to Q is called a nematic director ($\hat{n}$). Theoretically, the value of Q lies in the range of $-0.5 \leq Q \leq 1$. The value $Q = 1$ corresponds to a perfectly ordered structure, with all bonds perfectly aligned, which can never happen in practice. The value $Q = -0.5$ corresponds to a perfect oblate structure, and $Q = 0$ ($\lambda_+ = \lambda_0 = \lambda_- = 0$) corresponds to a disordered structure, which can be observed only in an infinitely large system.

To investigate order-disorder transition, the mean square fluctuation Σ^2 of Q was also calculated as:

$$\Sigma_Q^2 = \frac{1}{N_s - 1} \sum_{i=1}^{N_s} (Q_i - \langle Q \rangle)^2, \tag{2.27}$$

where $\langle Q \rangle$ is the average value of Q, calculated using Eq. 2.4 and N_s is the number of simulation snapshots. At a transition concentration, the structure usually reorients itself from disordered to more ordered states, and at this concentration the mean square variation of the order parameter between samples is the largest. It is noted that error bars (δ) in all the plots defined as standard error of each mean:

$$\delta = \frac{\Sigma}{\sqrt{N_s}}. \tag{2.28}$$

2.3.4.3 Helicoidal order parameter (P_2)

Helicoidal order parameter (P_2) was calculated to probe the structure for the presence of helicoidal ordering of surface bonds. To begin with, $\hat{e}_z$ was defined as a unit vector in the direction of the largest eigenvalue of the tensor $\mathbf{T}$ (Eq. 2.26) constructed from surface bonds of a chain. Vector $\hat{e}_\phi$ was then defined as: $\hat{e}_\phi = \hat{e}_z \times \hat{e}_r$, where $\hat{e}_r = \vec{r}_i/|\vec{r}_i|$ and $\vec{r}_i$ is position vector of monomer i. Then, P_2 was calculated using a second-rank Legendre polynomial, similar to that used to calculate nematic order parameter in liquid crystal system described in Chapter 1:

$$P_2 = \frac{1}{2N_b} \sum_{i=1}^{N_b} (3cos^2\varphi_i - 1), \tag{2.29}$$

where φ_i is the angle between bond $\hat{u}_i$ and $\hat{e}_\phi$ (Fig. 2.3 b)). The value of $P_2 = 1$ indicates a presence of a perfect helicoidal structure, $P_2 = -0.5$ implies that the structure is in splay arrangement, and $P_2 = 0$ is observed in a disordered state. The average value of the helicoildal order parameter ($\langle P_2 \rangle$) was calculated using Eq. 2.4, and mean square fluctuation of P_2 ($\Sigma^2_{P_2}$) was calculated using Eq. 2.27 by replacing Q with P_2.

2.3.4.4 Tetrahedral order parameter

Tetrahedral order parameter (q_3) belongs to a series of bond orientational order parameters q_l developed by Steinhardt *et al.* to study structures transitioning from supercooled liquid to glass phases [75]. Later, the q_3 order parameter was also successfully used to detect the presence of the tetrahedral structure in various systems such as in ice lattice [37, 38] and in crystallization of other solids [36]. In these studies [36–38], when the tetrahedral structure was formed, only the q_3 component was substantially larger than zero.

Since tetrahedral structure was predicted to be a ground state for nematic liquid crystals confined on a spherical surface [85], we expect to observe tetrahedral texture on the surface of a cavity, when a semi-flexible chain confined to a sphere much smaller than its persistence length. Therefore, we use q_3 to test for the presence of tetrahedral texture. In general, the order parameter q_l is written as:

$$q_l = [\ \frac{4\pi}{2l+1} \sum_{m=-l}^{l} |\bar{q}_{lm}|^2\ ]^{1/2}, \tag{2.30}$$

where $l = 3$, $-l \leq m \leq l$, and $\bar{q}_{lm,i}$ is averaged over N_b bonds, as shown below:

$$|\bar{q}_{lm}|^2 = \frac{1}{N_b} \sum_{i=1}^{N_b} q_{lm}(\vec{r}_i) q^*_{lm}(\vec{r}_i). \tag{2.31}$$

Here $\vec{r}_i$ refers to midpoint of a bond i. $q_{lm}(\vec{r})$ is expressed through spherical harmonics:

$$q_{lm}(\vec{r}) = Y_{lm}(\theta(\vec{r}), \phi(\vec{r})), \tag{2.32}$$

where $\theta(\vec{r})$ and $\phi(\vec{r})$ are polar angles of the bond, θ is the angle between $\hat{e}_z$, and $\vec{r}$, and ϕ is the angle between projection of midpoint of the bond on x-y plane and $\hat{e}_x$.

Here, $\hat{e}_z$, $\hat{e}_x$, and $\hat{e}_y$ are unit vectors in directions of eigenvector corresponding to the largest, second largest, and the smallest eigenvalues of the tensor $\mathbf{T}$ constructed from all bonds in a considered chain. For an ideal tetrahedral structure, we should expect to observe $q_3 \approx 0.745$ [38] and for disordered structure in an infinitely large system $q_3 = 0$ (when all θ angles are not equal to $\pi/2$ ($\forall\ \theta \neq \pi/2$), otherwise symmetric alignment of four peripheral vertices around central vertex on the same plane is observed [113].) We should note that q_3 was calculated only for surface bonds.

An average value of q_3 was calculated using Eq. 2.4, and mean square fluctuation of $\Sigma^2_{q_3}$ was determined using Eq. 2.27 by replacing Q with q_3.

2.3.4.5 Local nematic order parameter (Q_{local})

In experiments, the defect locations and charges are identified through observation of the nematic field in cross polarized light [32, 33]. In theoretical [85] and simulation [34] studies, they can be defined by calculating the local nematic order parameter (Q_{local}). The defect locations are identified as areas where the Q_{local} values are close to zero. The topological charge of each defect can then be determined by looking at the arrangement of the vector field around the defect, as described in Chapter 1.

To calculate Q_{local}, we use only surface bonds located closest to the confining surface. The Q_{local} value is calculated in the same manner as Q using Eq. 2.26, by selecting Q_{local} as the largest eigenvalue of $\mathbf{T}$, constructed from a selected surface bonds and their nearest neighbors, located within a distance $r/a = 1.2$ from them, where a is a bond length.

Chapter 3

Free energy and polymer structure of a spherically confined flexible chain

3.1 Introduction

A problem of a flexible polymer confinement inside a small sphere is often encountered in biological applications such as drug [114] and gene delivery [115], gel electrophoresis [116], and translocation of macromolecules across biomembranes [1, 2, 117]. In real systems, mechanisms governing these processes are very complicated due to biological factors such as pH of solution and interactions between molecules in the cell; thus, our understanding of these mechanisms is still far from complete. Molecular confinement has been studied both theoretically [61, 118] and with computer simulations [91, 119]. Most works address a simple model of a polymer confined inside a sphere, to gain an insight into the basic properties of this system, such as relation between the confinement free energy and driving force of molecular translocation [117], the free energy and translocation time [91], and the effect the confinement has on the structure of biomolecules [92, 93, 120]. The knowledge of the confinement free energy is crucial to these studies, and when this is known, it can be used to obtain a number of physical properties, such as osmotic pressure, translocation time, etc.

Placing a polymer chain into a sphere with a confining radius (r_c) smaller than

unconfined size of the polymer (r_G), greatly reduces its accessible conformations. This in turn, increases the free energy of the system. In the current study, we are interested in calculating the confinement free energy of a flexible chain confined inside a sphere of size smaller than or equal to its r_G in good solvent, and determining the relationship between this free energy with other parameters such as cavity size and osmotic pressure.

As mentioned in Chapter 1 and Chapter 2, the expanded ensemble density-of-states (EXEDOS) has been successfully used to study systems with continuous energy spectrum, such as solid crystals [49], liquid crystals [48], proteins [110], and colloids [50]. In this work, we apply the EXEDOS algorithm to a chain confined inside a sphere. We will start by using this algorithm on systems of an ideal flexible chain with various degrees of polymerization, confined in a sphere of variable sizes. We will address the power-law dependence of the confinement free energy on degree of confinement, and compare our findings to the relations in Eq. 1.12 [61]. Later, we will expand the algorithm to a system of a self-avoiding flexible chain in a sphere. We will also determine the relationship between the confinement free energy and other quantities such as polymer concentration, and osmotic pressure, and investigate monomer density profile of the confined chain near the confining surface. Finally, we will compare our findings to available theoretical, and simulation results.

3.2 Simulation description

3.2.1 Determination of density-of-states

MC simulation used to model a confined polymer, follows the approach described in Chapter 2. Below we focus on specific details of the polymer model and how the MC method was applied to it. The polymer was described as a freely-jointed chain with degree of polymerization $N+1$ varying from 100 to 600. For self-avoiding chains, each monomer was represented as a spherical bead with diameter $d/a = 0.8$, and in an ideal chain each monomer was considered as a point without volume, separated from adjacent monomers by a distance a. All polymer chains were equilibrated by conducting 10^8 traditional MC steps, each MC step consisting of $N+1$ trial moves [40]. The chain was confined to a sphere of radius varying from $r_c/a = r_{c_{min}}/a$ (de-

fined in Table 3.1) to $r_c/a = r_{c_{max}}/a \approx 4 \cdot r_G/a$ with an increment $dr_c/a = 0.5$. The upper (lower) limit on r_c/a was chosen to be large (small) enough, so that probability of polymer (being compressed) spreading beyond this radius would be negligibly small. In this case, $F(r_c/a = r_{c_{max}}/a) \approx F(r_c/a = \infty)$. Unconfined radius of a polymer chain is expressed as $r_G/a = N^{\nu}$, where $\nu \approx 0.588$ [54] for a self-avoiding chain in good solvent and $\nu = 0.5$ for an ideal chain [53]. To change the chain configuration, a randomly selected monomer was rotated by an arbitrary angle around the axis connecting the previous and following monomers in the chain. The new configuration was accepted if the condition of Eq. 2.19 was satisfied (and there was no overlap between monomers for self-avoiding chains). After each trial move, the algorithm checked the position of the furthermost monomer (from the center of a sphere). If it was located within the sphere, density of states and number of visits to this sphere were modified as described in Chapter 2. In this way, the distance of the furthermost monomer from center of a sphere served as a parameter determining the radius of confinement. There is a number of works that proved this distance to be an effective means to describe the physical size of the system [121]. For example, it was shown that scaling behavior of an effective spring constant in a cylindrically confined flexible chain better follows the predicted power-law behavior, when the furthermost distance is used as the chain's confined size, rather than the end-to-end distance [122]. This distance also appeared to give good agreement with theoretical predictions when used to calculate force needed to compress circular bacterial chromosome into micro-channel with optical tweezers [121]. Following these studies, we used the furthermost monomer from the center of a sphere to determine whether a chain occupied the considered sphere in EXEDOS simulations.

The simulations then followed the general method described in Chapter 2. For all chains, flatness constant was chosen equal to 0.7, and the initial and final modification factors were $ln\ c = 1$ and $ln\ c = 10^{-8}$, respectively.

Number of Monomers $(N+1)$	Confining radius $(r_{c_{min}}/a)$	
	Ideal chain	Self-avoiding chain
100	3.0	3.0
200	4.0	4.0
300	4.5	4.5
400	5.5	5.5
500	7.5	7.5
600	9.5	9.5

Table 3.1: Minimum confining radius $(r_{c_{min}}/a)$ of ideal and self-avoiding chains of different lengths $(N+1)$, used in simulations in this chapter.

3.2.2 Data analysis

Once the normalized density of states was found for each confined chain, free energy was then found from the relation:

$$e^{-\beta F(r_c/a)} \propto \Sigma_{r_i/a=r_{c_{min}}/a}^{r_c/a} g(r_i/a) \propto W(r_c/a), \tag{3.1}$$

where $g(r_i/a)$ is the density of states at a confining radius r_i/a, $r_{c_{min}}/a$ is the smallest radius (defined in Table 3.1). $W(r_c/a)$ is probability distribution of the furthermost monomer position relative to center of a sphere.

The density of configuration states was used as a weighting factor for acceptance criterion in Eq. 2.19 during production runs for determination of the number density (n), which was averaged over 8000 structures obtained during the production runs. To calculate the number density, a number of monomers was counted in a series of spherical shells, each occupying the radius between r/a and $(r+dr)/a$ (where $dr/a = 0.5$), divided by the volume of each shell $(4\pi(r/a)^2(dr/a))/3$, and normalized so that $\int n(r/a)4\pi(r/a)^2(dr/a) = 1$.

3.3 Results and discussion

3.3.1 Ideal chain

In this section, we test the performance of the EXEDOS algorithm when applied on spherically confined polymers. We will first determine the confinement free energy of an ideal chain confined inside a sphere of size smaller than its unconfined size. Then, we will compare the free energy to that obtained by theoretical prediction [61, 118].

3.3.1.1 Radial distribution of the furthermost monomer

Fig. 3.1 shows the probability distribution of the furthermost monomer position relative to the center of a confining sphere, for six ideal chains of different lengths, ranging from 100 to 600 monomers. The distribution functions are close to Gaussian: the probability that the furthermost monomer would travel to the radius much smaller or much larger than the chain unconfined size, is relatively small. Indeed, it is unfavorable to compress a chain into a sphere with diameter much smaller than its unconfined size and impossible to expand the chain more than its extended size. Thus, the chain loses a large fraction of its accessible conformations under strong compression or large extension. The longer chain, having bigger size, feels the confining surface sooner. Thus, longer chains lose more accessible conformations, compared to shorter chains, when confined in a sphere of the same size [96]. To avoid the loss of accessible conformations, the longer chain favors confinement in a bigger sphere, and peak of the distribution function shifts to larger confining radius, as seen in Fig. 3.1.

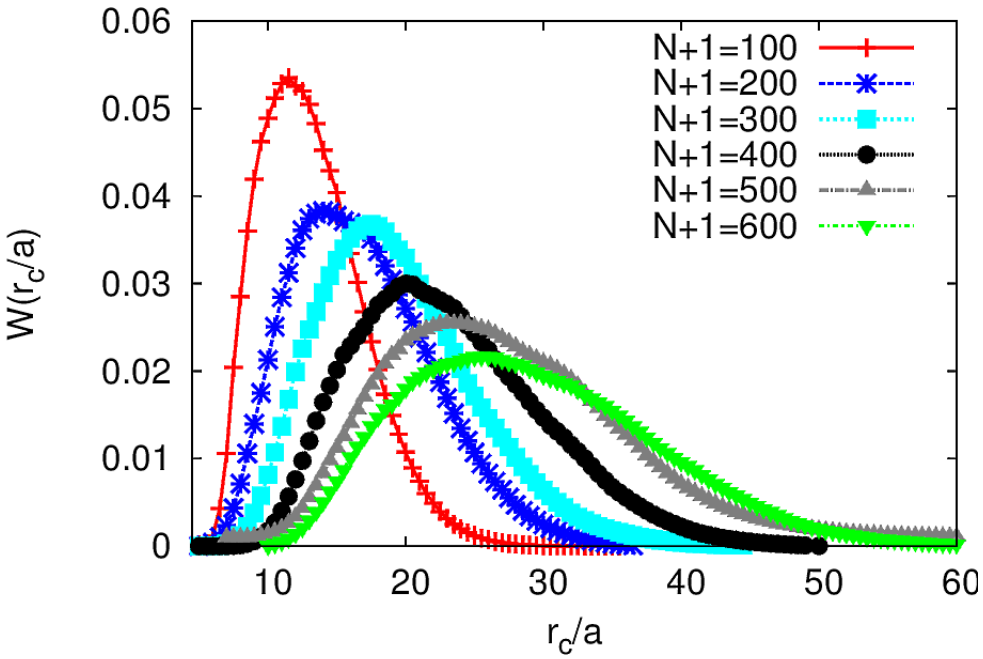

Figure 3.1: Radial distribution of the furthermost monomer position ($W(r_c/a)$) of an ideal chain of different lengths ($N+1$), labeled in the legends.

3.3.1.2 Confinement free energy

The confinement free energy of a chain can be calculated using the relation in Eq. 3.1. Changes in the confinement free energy (ΔF) of each chain in Fig. 3.2 were calculated as a difference in free energy of the chain confined inside a sphere of size r_c/a and $r_c/a \approx 4 \cdot r_G/a$. Fig. 3.2 shows that the confinement free energy increases monotonously as the confining radius decreases, indicating unfavorable packing process due to decreasing number of available configurations. The longer chain feels the confining surface sooner, compared to a shorter chain when compressed into a sphere of the same size. Thus, the longer chain loses more accessible configurations and has larger free energy cost, as shown in Fig. 3.2.

To determine the relationship between the confinement free energy and a confining radius for an ideal chain inside a sphere of size smaller than its unconfined size, changes in the confinement free energy were plotted as a function of the ratio of polymer size to a confining radius (r_G/r_c), in logarithmic scale, as shown in Fig. 3.3. The value ΔF follows the power law, with the slope ranging from 2.54 $\pm$0.12

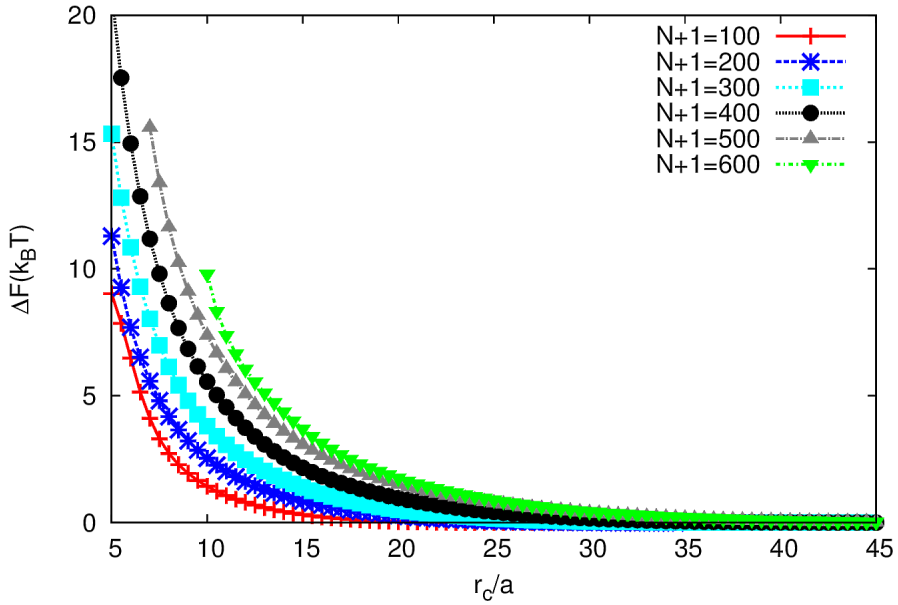

Figure 3.2: Changes in the confinement free energy (ΔF) of an ideal chain with different lengths ($N+1$) labeled in the legends, as a function of a confining radius (r_c/a). Here $\Delta F = F(r_c/a) - F(\infty)$, where $F(\infty)$ is the free energy of the chain confined inside a sphere of size $r_c/a \approx 4 \cdot r_G/a$, and $r_G/a = N^{\nu}$ is an unconfined size of a chain, a is a bond length, and $\nu = 0.5$.

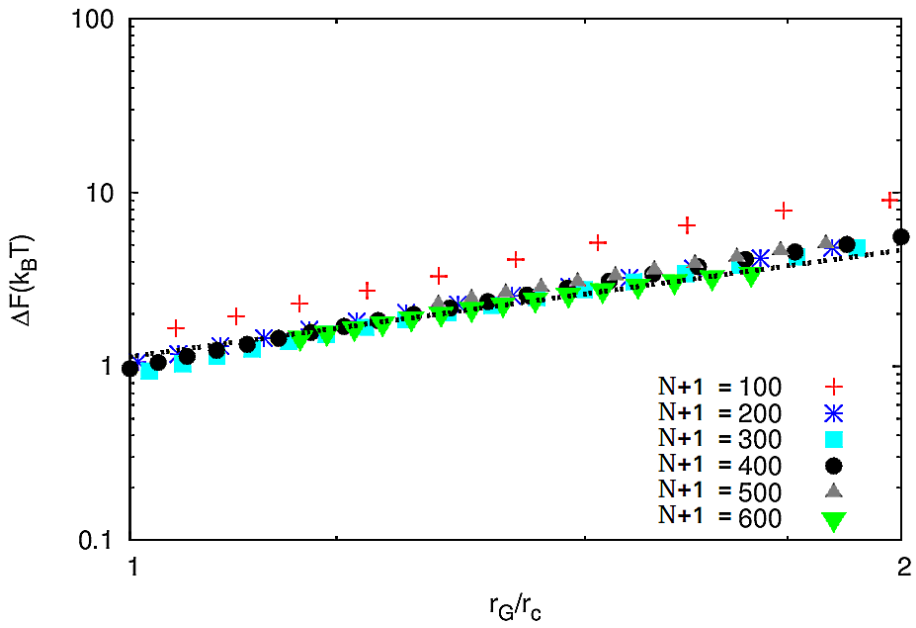

Figure 3.3: Changes in the confinement free energy (ΔF) of an ideal chain of different lengths ($N+1$) labeled in the legends, as a function of the ratio of polymer size to a confining radius (r_G/r_c) plotted in logarithmic scale. Here $\Delta F = F(r_c/a) - F(\infty)$, where $F(\infty)$ here is the free energy of the chain confined inside a sphere of size $r_c/a \approx 4 \cdot r_G/a$, and $r_G/a = N^{\nu}$ is an unconfined size of a chain, a is a bond length, and $\nu = 0.5$. The dashed line gives slope of 2.03 ± 0.05, obtained from the chain with 600 monomers.

for the chain of 100 monomers to 2.03 $\pm$0.05 for the chain of 600 monomers. The changes in slope are most likely due to effects of a finite chain length on scaling behavior. The power-law relation: $\Delta F \propto (r_c/a)^{-2.03\pm0.05}$, which was found for the chain with 600 monomers fits well with the theoretical findings reported by Cassasa [61] and those later reported by Gao *et al* [118].

Seeing a good agreement between our findings and theoretical works [61, 118], we expand the EXEDOS algorithm to systems of a self-avoiding flexible chain of various lengths, confined inside a sphere of variable sizes.

3.3.2 Self-avoiding chain

3.3.2.1 Radial distribution of the furthermost monomer position

Fig. 3.4 shows the probability distribution of the furthermost monomer position relative to the center of a confining sphere, for a self-avoiding chain of different lengths, ranging from 100 to 600 monomers. Each curve represents an average of 5 sets of data. Here, the distribution functions for the outermost monomer positions are also close to Gaussian, similar to those observed for ideal chains. The probability that the chain would spontaneously compress into a radius much smaller than its unconfined size or expand far beyond its extended size is vanishingly small. Longer chains occupy larger radius, compared to shorter ones. Thus, the longer chains lose more accessible conformations than the shorter chains, when confined inside a sphere of the same size. To minimize this loss, the longer chain prefers to occupy in a bigger sphere, and its distribution peaks at a larger confining radius, as shown in Fig. 3.4. It is noticeable that the probability distribution of the furthermost monomer position of the self-avoiding chain in Fig. 3.4 peaks at a larger confining radius, compared to that of an ideal chain in Fig. 3.1 of the same length, because of the added excluded volume of the self-avoiding chain.

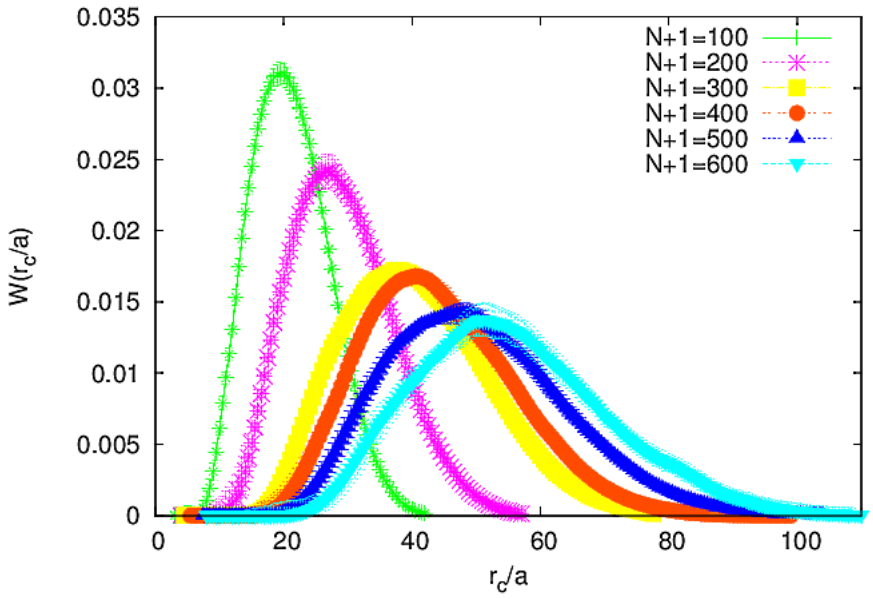

Figure 3.4: Radial distribution of the furthermost monomer position ($W(r_c/a)$) of a self-avoiding chain of lengths $(N+1)$, labeled in the legends.

3.3.2.2 Confinement free energy

The confinement free energy of the chain was calculated using the relationship in Eq 3.1. Fig. 3.5 shows the difference between free energy of a chain confined inside a cavity of radius r_c/a and the radius that is much larger than r_G/a ($r_c/a \approx 4 \cdot r_G/a$). This free energy difference shows the energy required to bring the molecule from unconfined state into a sphere of radius r_c/a: $\Delta F = F(r_c/a) - F(\infty) \approx F(r_c/a) - F(4 \cdot r_G/a)$. It is apparent that the confinement free energy decreases monotonously with an increase of a confining radius. With an increase in the degree of polymerization $(N+1)$, the unconfined size of the polymer becomes bigger and when confined, the bigger chain loses more accessible conformations, compared to the smaller one [96]. Therefore, the confinement free energy of longer chains is larger, compared to the shorter ones, when they are confined in a sphere of the same size, as seen in Fig. 3.5.

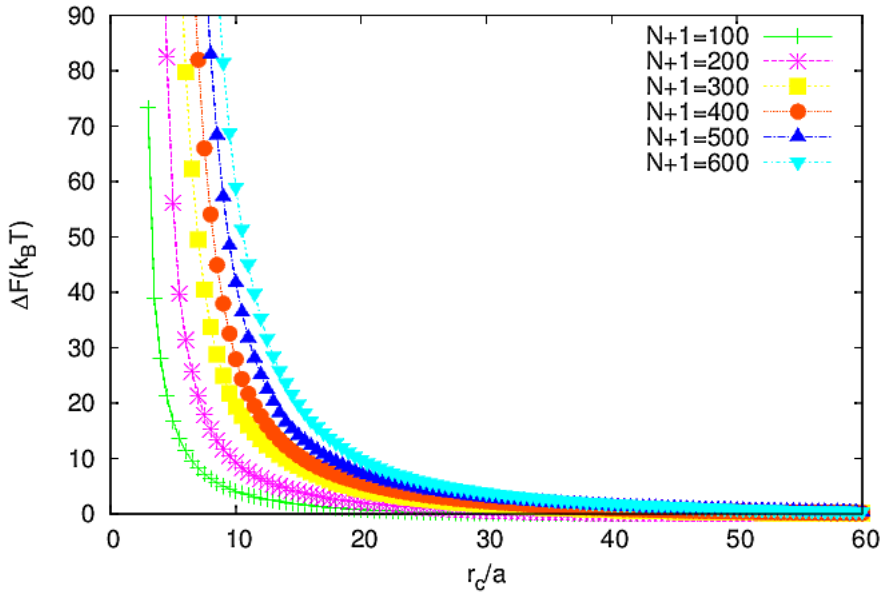

Figure 3.5: Changes in the confinement free energy (ΔF) of a self-avoiding flexible chain with lengths $(N+1)$, displayed in the legend, and $\Delta F = F(r_c/a) - F(\infty)$, where $F(\infty)$ here is free energy of the chain confined inside a sphere of size $r_c/a \approx 4 \cdot r_G/a$, and $r_G/a = N^{\nu}$ is an unconfined size of a chain, a is a bond length, and $\nu \approx 0.588$ [54].

Scaling analysis

The dynamic and static properties of a self-avoiding flexible chain in semi-dilute solution ($r_c \leqslant r_G$) are usually expressed as functions of certain parameters of interest such as polymer concentration (η) or a confining radius. The power-law dependence of the properties on these parameters can be determined to describe polymer behavior irrespective of its chemical constitution. In this section, we consider the scaling relationships for the free energy as a function of various parameters, and compare them to the previously reported findings [91, 92, 120].

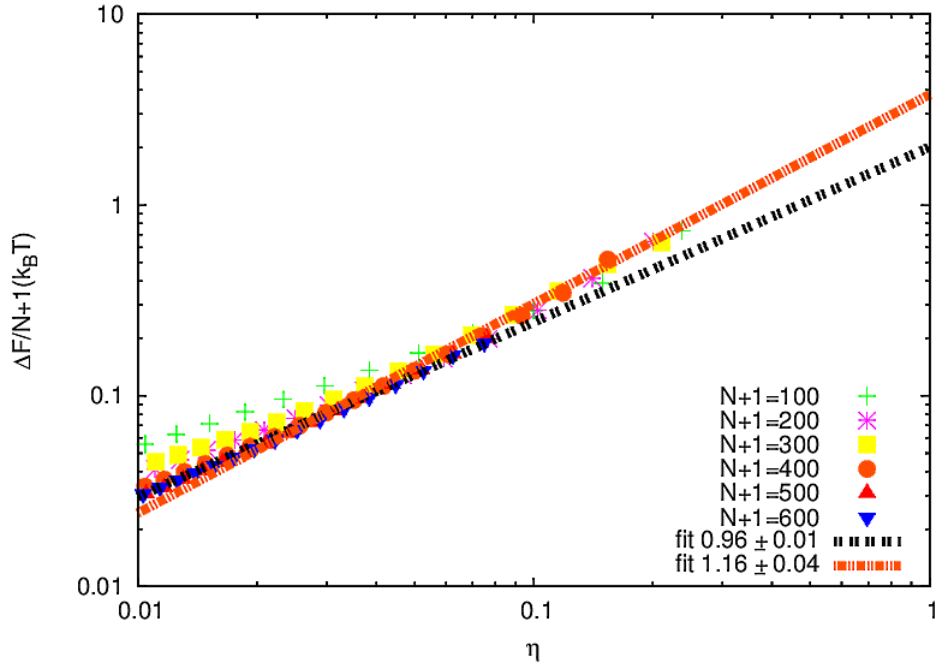

Figure 3.6: Free energy of confinement per monomer ($\Delta F/(N+1)$) as a function of polymer concentration (η). Orange line shows the power law fit obtained from the chain with 300 monomers for the dependence of $F/(N+1)$ on η at $0.15 < \eta < 0.3$: $\Delta F/(N+1) \propto \eta^{1.16\pm0.04}$, and black dashed line shows the power law fit obtained from the chain with 600 monomers for $0.01 \leqslant \eta \leqslant 0.15$: $\Delta F/(N+1) \propto \eta^{0.96\pm0.01}$. Here $\Delta F = F(r_c/a) - F(\infty)$, where $F(\infty)$ is free energy of the chain confined inside a sphere of infinite size (defined here at $r_c/a \approx 4 \cdot r_G/a$), $r_G/a = N^{\nu}$ is an unconfined size of the chain, a is a bond length, and $\nu = 0.588$ [54].

Fig. 3.6 shows the relation between the specific free energy of confinement ($\Delta F/(N+1)$) and the polymer concentration (η) in good solvent, where $\eta = (N+1) \cdot d^3/(2r_c)^3$. At a given concentration, the difference of $\Delta F/(N+1)$ decreases with an increase of chain lengths ($N+1$) due to the decrease in the finite size corrections of the scaling relationships. For all chains, the curves nearly overlap indicating that the confinement free energy is an extensive function in number of monomers ($N+1$). At moderate concentrations ($0.01 \leqslant \eta \leqslant 0.15$), the data of the chain with 600 monomers fits well with the power-law relation: $\Delta F/(N+1) \propto \eta^{0.96\pm0.01}$ (shown as a black dashed line in Fig. 3.6), which is slightly smaller than that predicted by the density functional theory (DFT) [120] ($\Delta F/(N+1) \propto \eta^{1.31}$) and other MC simulations [91] ($\Delta F/(N+1) \propto \eta^{1.28}$), when a polymer chain was in the same

concentration range.

At higher polymer concentrations ($0.15 < \eta < 0.30$), the dependence of the specific free energy on concentration becomes stronger and follows the relation: $\Delta F/(N+1) \propto \eta^{1.16\pm0.04}$ (from data of the chain with 300 monomers), shown as an orange dotted line in Fig. 3.6. This increase in the exponent value at high polymer concentration, was also observed in the DFT studies [120], MC simulations ($\Delta F/(N+1) \propto \eta^{1.97}$) [91] of a flexible polymer confined inside a cavity, and in experiments where polyethylene glycol partitioned in protein nanopores [124].

	Exponent (x)	Method	No. of Monomers ($N+1$)
$\Delta F \propto \eta^x$ ($\eta \leqslant 0.15$)	1.28 ± 0.06	MC [91]	2048
	1.31	DFT [120]	2000
$\Delta F \propto \eta^x$ ($0.15 < \eta < 0.35$)	1.97±0.07	MC [91]	2048
$\Delta F \propto (r_G/r_c)^x$ ($r_G/r_c < 1.0$)	3.8 ± 0.1	MC [91]	2048
	3.93	DFT [120]	2000
	3.44	MC [92]	2000

Table 3.2: Changes in the confinement free energy (ΔF) of a self-avoiding chain inside a sphere of radius r_c/a, at different concentrations (η), as reported by several authors.

Osmotic pressure (P) also follows a power-law when expressed as a function of polymer concentration. The osmotic pressure is determined using the relation $P = \frac{\partial F}{\partial V}$ with $V \simeq (2r_c)^3$; hence, $P \sim \eta F$. Using Eq. 1.13, P can be written as $P \sim \eta^{3\nu/(3\nu-1)} \sim \eta^{2.30}$, with $\nu = 0.588$. From our simulations at low concentration ($0.01 \leqslant \eta \leqslant 0.15$) $\Delta F \propto \eta^{0.96\pm0.01}$, and the osmotic pressure P is predicted to be: $P \sim \eta^{1.96\pm0.02}$ - smaller than that given by blob theory ($P \sim \eta^{2.30}$ using $\nu = 0.558$) or observed in other MC [91] ($P \sim \eta^{2.28}$) and MD simulations [93] ($P \sim \eta^{2.27}$). At higher polymer concentrations ($0.15 < \eta < 0.3$), we get $\Delta F \propto \eta^{1.16\pm0.04}$; thus, P is predicted to be: $P \propto \eta^{2.16\pm0.04}$, which is smaller than that predicted by MC simulations [91]: $P \propto \eta^{2.97\pm0.07}$.

The power-law dependence of the free energy on a confining radius r_c becomes evident when the confinement free energy is plotted on logarithmic scale as a function of r_G/r_c, as shown in Fig. 3.7 for a self-avoiding chain confined inside a sphere

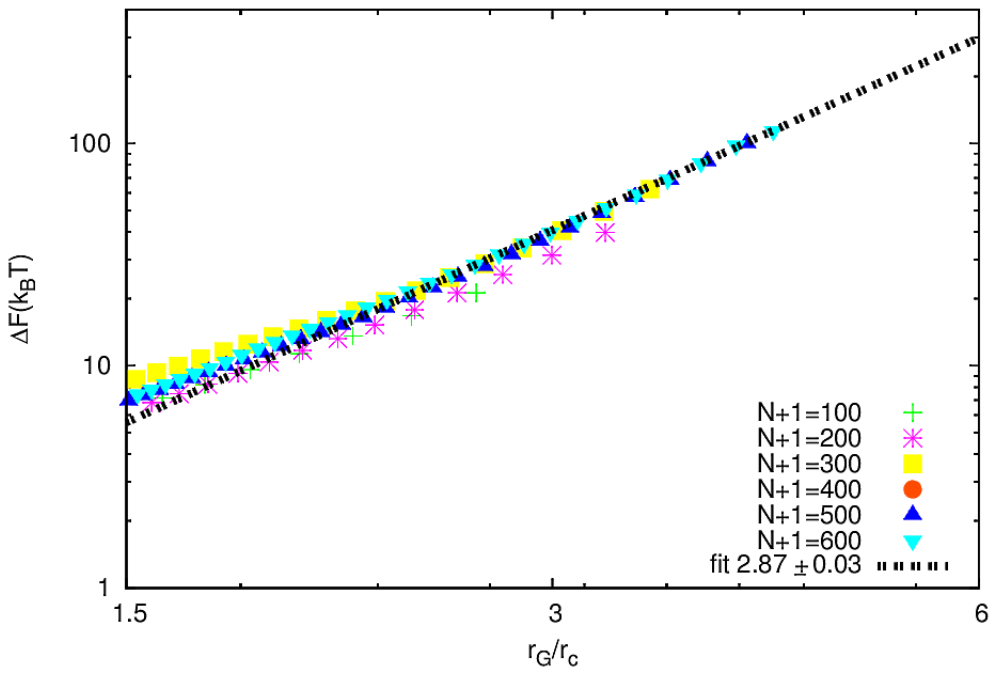

Figure 3.7: Confinement free energy change (ΔF) as a function of the ratio of polymer size to a confining radius (r_G/r_c), for various degrees of polymerization ($N+1$). Dashed line shows a linear fit of data obtained for a self-avoiding chain of 600 monomers, confined inside a sphere of size $r_c < r_G$ (with slope of 2.87 ± 0.03).

of size smaller than r_G. The free energy of the longest chain (600 monomers) was fitted linearly to determine the power law exponent, giving $\Delta F \propto (r_G/r_c)^{2.87\pm0.02}$. The scaling exponent obtained here is smaller than that found by other MC simulations [91, 92] ($\Delta F \propto (r_G/r_c)^{3.8\pm0.1}$ and $\Delta F \propto (1/r_c)^{3.44}$) and blob theory in Eq. 1.13 ($\Delta F \propto (r_G/r_c)^{3.93}$) for $\nu \approx 0.588$. The difference in the obtained exponent value, compared to other simulations, can be attributed to the short length of the chains studied here. Indeed, we observe that as the chain length increases from 100 to 600 monomers, the exponent of r_G/r_c also increases from 2.09 to 2.87. The other works presented in Table 3.2 that give higher exponent value, use much longer chains ($N+1 \geq 2000$).

3.3.2.3 Structural analysis

To investigate the effect that the confining surface has on structure of confined polymers, we determine the monomer number density $n(r/a) * a^3$ as a function of distance from the center of a sphere (r/a). Fig. 3.8 shows the number density of monomers for various chain lengths $N+1$ confined in spheres of different radii. In a weak confinement (Fig. 3.8 a) $r_c/a = 20.5$), where the confining radius is large, most monomers avoid the periphery to maximize the number of accessible conformations. They move towards the center of a sphere for more available volume to orient and translate. Therefore, as we can see in Fig. 3.8 a) the number density for all chains is high near the center of the confining sphere. Longer chains have larger unconfined size; therefore, they have a broader density curve, compared to the shorter chains, as seen in Fig. 3.8 a).

As the degree of confinement increases (Fig. 3.8 b) $r_c/a = 14.5$), the number density of short chains remains almost unchanged ($100 \leq N+1 \leq 300$), compared to that in Fig. 3.8 a). On the other hand, longer chains ($500 \leq N+1 \leq 600$) with $r_G >> r_c$, are pushed inwards by the force created due to overlap of monomers inside the sphere. Thus, more monomers can be found in the vicinity of the confining surface in the chains of length $500 \leq N+1 \leq 600$.

In stronger confinement (Fig. 3.8 c) $r_c/a = 8.5$), all chains are affected by the presence of a confining surface. Due to surface-monomer interaction, we observe sharp drop of the monomer number density close to the surface for all chains, which

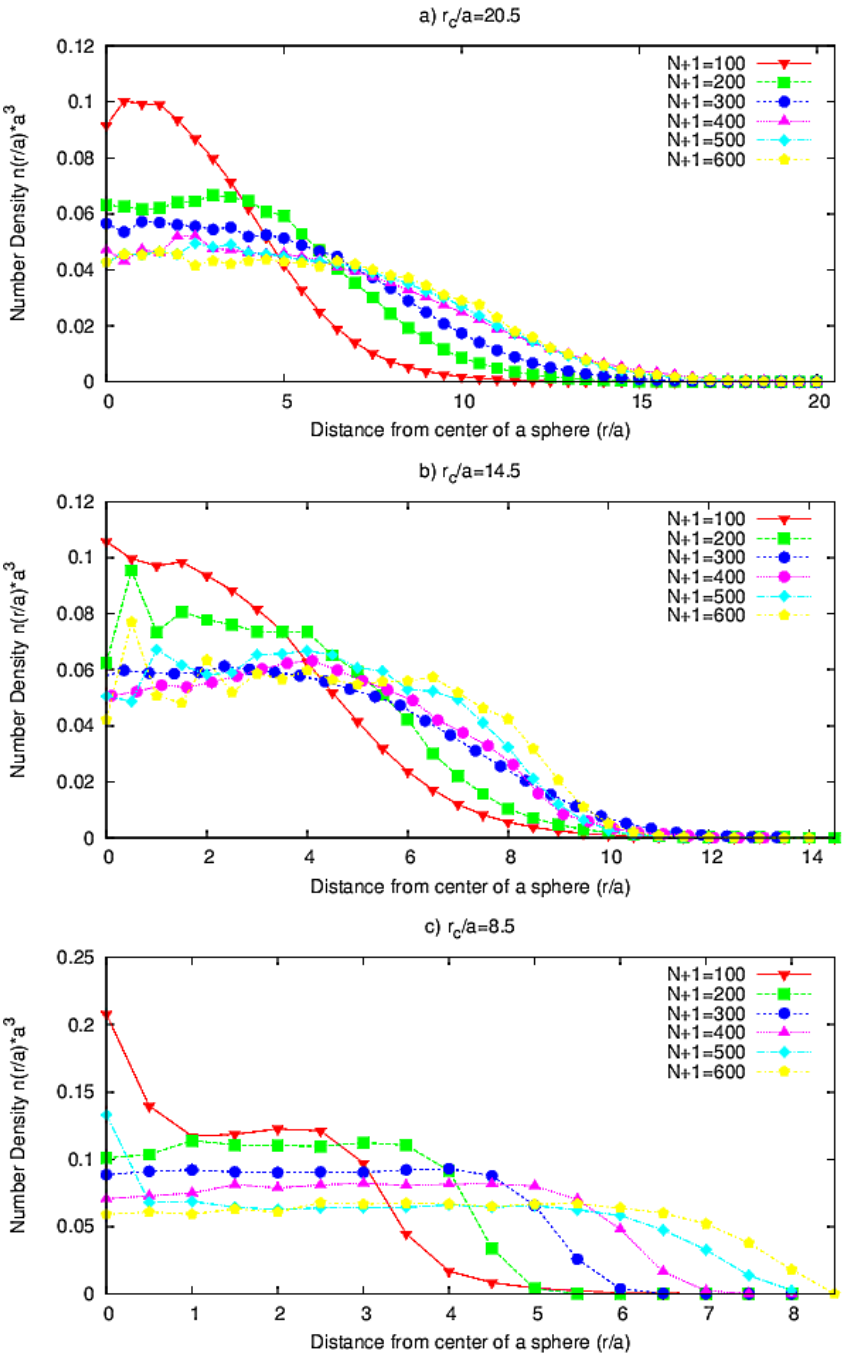

Figure 3.8: The number density $(n(r/a) * a^3)$ as a function of distance from the center of a confining sphere (r/a) for chains with different degrees of polymerization $(N + 1)$, labeled in the legends.

is similar to that observed in other simulations [92, 93]. The force created due to monomer-monomer overlap inside the sphere, is larger for longer chains; thus, larger number of monomers in the longer chain lies close to the surface.

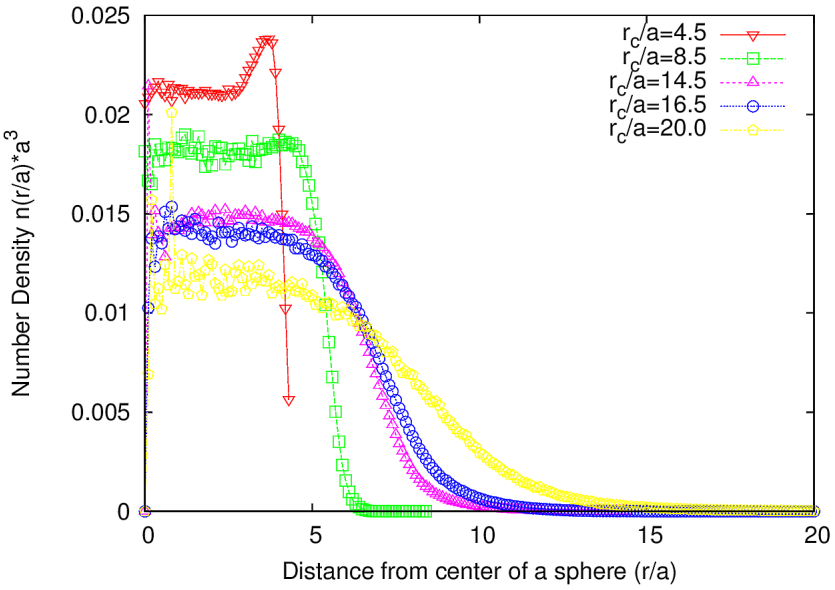

Figure 3.9: The number density $(n(r/a) * a^3)$ as a function of distance from the center of a confining sphere (r/a) for the chain of $N + 1 = 300$. Legends show the reduced radii (r_c/a), where a is a bond length. As concentration decreases (larger r_c/a), the chain tends to reorient and translate through available volume to increase its accessible conformations; thus, the number density profile broadens.

At a very high concentration (Fig. 3.9 $r_c/a = 4.5$), we start seeing the oscillation in the number density profile. It can be easily seen when the chain of length $N + 1 = 300$ is confined inside a sphere of radius $r_c/a = 4.5$, as shown in Fig. 3.9. This oscillation appears due to reorientation of the monomers affected by the confining surface, which was also observed in other studies [93, 120, 132].

3.4 Conclusions

In this chapter, we adopted the EXEDOS algorithm to determine the density of configuration states (DOS) of an ideal and a self-avoiding flexible chain with small and

intermediate chain lengths, confined inside a sphere of various sizes. The density of states was then used for calculation of the confinement free energy, to study the relationship between the confinement free energy of a chain and other parameters such as volume fraction, osmotic pressure and confining radius .

For an ideal chain, confined inside a sphere of radius $r_c \leq r_G$, calculated free energy for the chain of $N+1 = 600$ depends on the confining radius as: $\Delta F \propto (r_c/a)^{-2.03\pm0.05}$, which is in a good agreement with theoretical predictions [61, 118]. With this good agreement, we expanded the EXEDOS algorithm to systems of self-avoiding chains confined inside a sphere.

For short self-avoiding chains ($100 \leq N+1 \leq 400$), the EXEDOS method allowed us to determine the density-of-states (DOS) in both weak and relatively strong confinement. Efficiency of the EXEDOS method, when applied to shorter chains, is limited at very strong confinement (polymer volume fraction $\eta \sim 0.3$). For the longer chains $500 \leq N+1 \leq 600$, the EXEDOS method is capable of determining the DOS, when systems are under intermediate and weak confinement (polymer volume fraction $\eta \sim 0.15$). When used on a single CPU, the algorithm becomes less efficient for polymer models of sizes larger than 500 monomers. This limitation was also reported when the original Wang-Landau (WL) algorithm was used for a flexible polymer system of size larger than 512, to determine its density of energy states [125].

The density of configuration states determined using the EXEDOS method for systems of self-avoiding chains was then used to calculate the confinement free energy. The power law dependence of the confinement free energy on the confining radius (r_c) and polymer concentration η deviates from that predicted by the blob theory, which can be attributed due to the finite size effects. The estimated density of configuration states was used as a weighting factor for acceptance criterion to investigate structures of confined chains. The reduced number density ($n(r/a) * a^3$) was used to predict distribution of monomers along the radius of a confining sphere. In a very small cavity with $r_c << r_G$, the confined chain is forced towards the surface by a force appearing due to the monomer overlap, and the number density becomes high at the surface leading to oscillations. In a bigger cavity with $r_c < r_G$, the competition between the force due to monomer-monomer interactions and monomer-surface interactions, produce a monotonous decline of the number density with radius. As the confining radius becomes even larger, the polymer chain

concentrates near the center of the sphere to maximize its accessible conformations, which it will otherwise lose due to non-bonded interactions with the confining surface.

Chapter 4

Formation of ordered structures in a spherically confined semi-flexible chain

4.1 Introduction

It is well known that confinement of rigid molecules in a cavity at high density can cause distortion in their orientations. This leads to formation of structures that are not observed in bulk, and areas, in which the directors are undefined and order parameter is close to zero, called defects. Theoretical study by Lubensky *et al.* had predicted that three possible configurations, namely spool-like or helicoildal (Fig. 1.7 a)), splay (Fig. 1.7 b)) and tetrahedral (Fig. 1.7 c)), would form when orientationally ordered molecules are confined on a spherical surface [85]. The total topological charge of configurations formed on spherical surface is limited to +2, following Poincare-Hopf theorem (see section 1.2.3.2 for more details) [79, 80]. The first two configurations both have defects positioned at the north and the south poles of the sphere, each defect having a charge of +1, but orientation of the director field in the helicoildal configuration is different from that in the splay configuration. The director field of the helicoildal configuration (Fig. 1.7 a)), spirals around the poles, parallel to the equator line, while in the splay configuration (Fig. 1.7 b)) it radiates away from the poles, along the meridians. The third theoretically possible configuration is a tetrahedron. It has four defects, each with charge of +1/2, located

at vertices of the tetrahedron. The results from this theoretical study [85] were later supported by simulations [34] and experiments [32, 33]. Both experiments [33] and simulations [34] show that the tetrahedral structure usually forms in a thin and uniform nematic shell. In a non-uniform nematic shell, locations of the four +1/2 defects no longer coincide with the vertices of tetrahedron [32, 34] and this structure is called tennis ball structure [95]. In a droplet or thick nematic shell, a structure with two +1 defects usually forms [32, 34]. Recent theoretical studies of rigid linear particles densely packed on a spherical surface [126] also show the presence of these three spatial configurations. For a self-avoiding semi-flexible polymer chain densely confined on a sphere, Zhang and Chen reported formation of a tennis ball structure when using MC simulations [95]. Other MC simulations by Angelescu *et al.* have found that the tennis ball structure was observed in a polyelectrolyte semi-flexible chain with excluded volume interaction, confined on a spherical surface at low density [127].

For a semi-flexible chain confined inside a cavity, Spakowiz and Wang observed imperfect spool-like structures when using MD simulations on a DNA chain released into a capsid with twisting [24]. The same structure was also observed in the study of Oskolkov *et al.*, who employed the density functional theory to model a semi-flexible chain confined in the inner surface of a spherical cavity [26]. These configurations were confirmed in experimental studies of bacteriophage T7 DNA packing by cryo-electron microscopy [23]. As reported by Fathizadeh *et al.*, this structure was found in particle dynamics simulations only when the DNA was injected into a sphere (without twisting) [25]. However, the tennis ball structure was observed instead, when the DNA was compressed into a sphere [25]. Recent MD simulations of multiple stiff chains, each having 32 monomers and $l_p/a = 32$, confined in a sphere showed that the tennis ball structure was obviously observed only when system density $\rho \geqslant 0.4$ [35].

The studies above did not attempt to identify a tennis ball structure with a tennis ball order parameter [25, 35]. Instead, the tennis ball structure was visualized [25, 35] and defect locations were identified as the areas where the nematic order parameter was low [25]. In crystallization problems [36–38], the presence of tetrahedral phase was successfully identified using a tetrahedral order parameter (q_3) developed by Steinhart *et al.* [75]. Among a series of q_l values, in a structure with tetrahedral symmetry, q_3 was the only strong and non-zero component but q_3 was

near zero when used with non-tetrahedral structures [36–38]. Due to the similarity between tetrahedral and tennis ball configurations which both have 4 defects, q_3 should be applicable to the tennis ball structure as well.

In the current study, we perform MC simulations to investigate spatial arrangement of a semi-flexible chain of varying bending rigidities, confined inside a sphere of various sizes. As described in Chapter 1, the traditional MC simulations suffer from insufficient sampling when used on systems at low temperatures and with rugged energy landscape [40, 41]. To remedy the problem, we applied the expanded density-of-states (EXEDOS), which has the acceptance probability inversely proportional to the density of configuration states, making a state with lower number of visits more likely to be visited and ensuring more efficient sampling of coordinate space. The EXEDOS algorithm is described in detail in Chapter 2.

In this chapter, we look at the conformations of a semi-flexible polymer inside a spherical cavity by applying the EXEDOS algorithm to a simple polymer model. Nematic order parameter (Q) is used to investigate isotropic-nematic (I-N) transition, while, helicoildal (P_2), and tetrahedral order parameters (q_3) are used to test the confined polymer for the presence of helicoildal and tetrahedral conformations, respectively.

The rest of this chapter addresses the model and simulation methodology, followed by discussion of simulation results and summary of findings.

4.2 Simulation description

4.2.1 Determination of density-of-states

Simulations here follow the approach described in Chapter 2. Only specific details related to this chapter are described here. Five pearl-necklace chains with 300 identical monomers were randomly generated ($N + 1 = 300$). Each monomer was separated from adjacent monomers by a distance a, and was described with a hard sphere potential having diameter of 0.8 ($d/a = 0.8$). Bending rigidity was introduced to the chains through the bending constant (k_b), and bending energy was calculated using Eq. 2.21 given in Chapter 2. In elasticity theory, the persistence length (l_p) is related to bending constant (k_b) as $k_b/k_BT = l_p/a$ [35, 56, 58, 59]. The five generated chains were assigned persistence lengths l_p/a of 8, 12, 16, 24, and 32.

A range of confining radii (r_c) was then selected for each chain. Since we are interested in the investigation of a confined chain at high packing density ($\rho > 0.15$, where $\rho = 3 * (N+1) * d^3/4\pi r_c^3$); the smallest confining radius should be much smaller than a polymer length. The biggest radius was selected at $r_c/a = 13$, corresponding to $\rho = \rho_{min} = 0.017$, so that the confining range was not too wide and the simulation would converge in a reasonable time. For the chains with $l_p/a = 8$, 12 and 16, the confining radius (r_c/a) varied from 5.6 to 13. For stiffer chains with $l_p/a = 24$ and 32, spherical radii r_c/a varied from 6 to 13 and 6.6 to 13, respectively. This range was then divided into a number of bins, each bin having width $dr/a = 0.2$. It is worth noting that if the bin width is made larger, the systems tend to get stuck in the same bin and convergence is slow. On the other hand, if the bin width is too narrow, significant computational time is required to obtain reasonable data. Each polymer chain was confined by placing its center of mass at the center of the confining sphere as depicted in Fig. 2.3 a). The simulations then followed the method described in Chapter 2. The convergence was considered satisfied when the modification factor was $ln\ c = 10^{-8}$, and the flatness constant was 0.85.

4.2.2 Production runs

After achieving convergence of the simulations described above, free energy difference (ΔF) of the each chain was determined using Eq. 2.25. Here, $\Delta F = F(r_c/a) - F(r_c/a = 13)$.

In production runs, the estimated density of states for each system was used as a weighting factor for all measurements. The total bending energy (E_b) was calculated using Eq. 2.21. Radial number density (n) of a chain was also investigated. To find n, the confining radius was first divided into a number of concentric shells, of width $dr/a = 0.125$. A number of monomers was counted in each shell, occupying the radius between r/a and $(r + dr)/a$, divided by the shell volume $V = 4\pi \cdot ((r+dr)^3 - r^3))/3a^3$, and normalized to give $\int n(r/a) * 4\pi (r/a)^2 (dr/a) = 1$. In order to investigate isotropic-nematic (I-N) transition of a system, nematic order parameter (Q) was defined as the largest eigenvalue of the tensor $\mathbf{T}$ (expressed in Eq. 2.26), constructed from all surface bonds of each chain. The surface bonds were defined as the bonds laying between the confining surface and the shell holding the

largest number of monomers. I-N transition density (ρ) was determined by finding the mean square fluctuation of Q from Eq. 2.27. Error bars in all plots were calculated from standard error of the mean value using Eq. 2.28. To test whether the structures formed helicoildal or tetrahedral arrangements, helicoildal (P_2) and tetrahedral (q_3) order parameters were calculated for surface bonds using Eq. 2.29, and Eq. 2.30, respectively. It is worth noting that measurements in the production runs were made once every 2500 MC steps, and averaged over 8000 conformations, while 1 MC step is defined by $N + 1 = 300$ trial moves [40].

4.3 Results and discussion

4.3.1 Energy calculations

The confinement free energy difference of a polymer chain (ΔF) calculated in this study represents the work required to compress the confined chain from $r_c/a = 13$ to r_c/a. Fig. 4.1 shows free energy change (ΔF) as a function of a confining radius (r_c/a). The confinement free energy of all polymer chains increases with the decreasing confining radius. The rate of change of the free energy gradually increases for all chains as the volume of confining sphere decreases, as shown in Fig. 4.1. The stiffer chains have larger sizes, and lose more accessible conformations, when confined inside a sphere of the same size. Therefore, the rate of change of their free energy is higher, compared to that of more flexible chains.

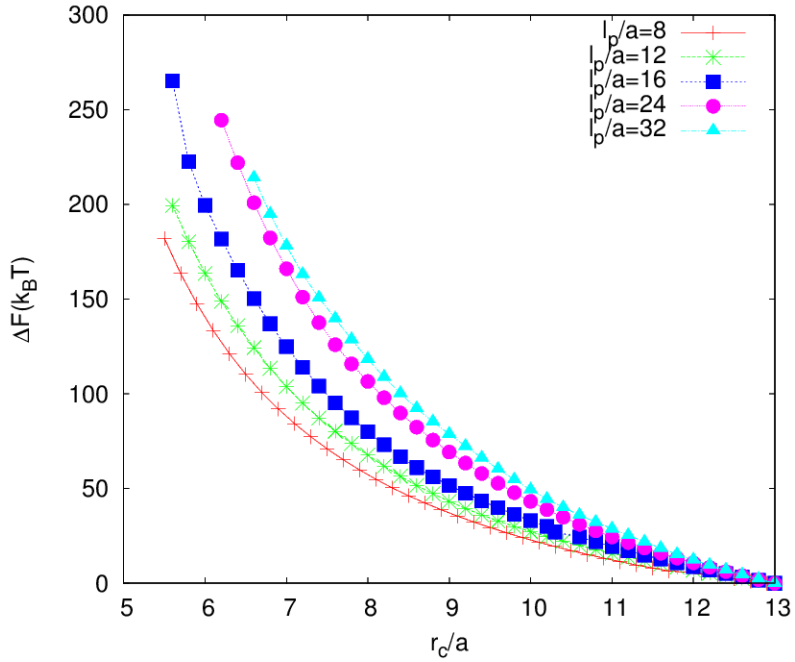

Figure 4.1: Free energy change ($\Delta F = F(r_c/a) - F(r_c/a = 13)$) as a function of a confining radius (r_c/a). Legends denote persistence length (l_p/a).

In the production runs, bending energy (E_b) was calculated and plotted as a function of polymer density (ρ), as shown in Fig. 4.2. At densities $\rho \leq 0.1$, the capsid 's environment is not significantly crowded, and the confined chains tend to occupy less ordered state with higher entropy, leading to a steady increase in bending energy, similar to that observed in the simulations of DNA confined inside a sphere at low density ($\rho < 0.2$) [21, 129, 130].

At $\rho > 0.1$, the confined chains quickly lose their accessible conformations. They minimize their free energy change by wrapping around the confining surface, which leads to a decline of the growth rate of ΔE_b, as shown in Fig. 4.2. The decrease in the rate of bending energy growth at relatively high densities was also observed in the simulations of densely packed DNA inside a cavity ($\rho > 0.3$) [24, 131].

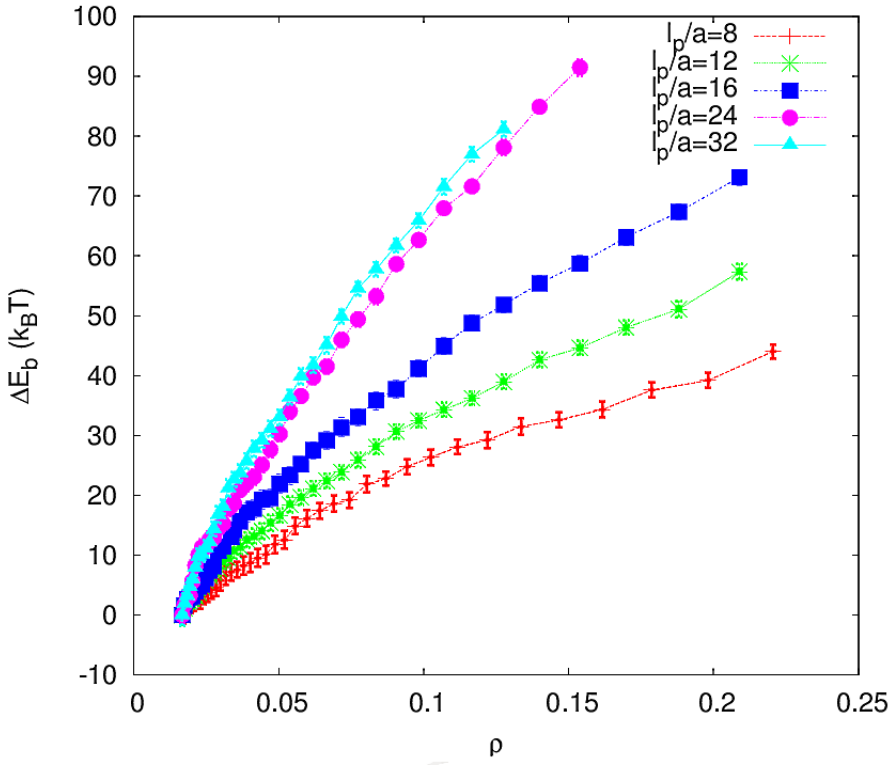

Figure 4.2: Bending energy change ($\Delta E_b = E_b(\rho) - E(\rho = 0.017)$) as a function of polymer packing density (ρ). Legends display persistence length (l_p/a).

4.3.2 Conformational analysis

4.3.2.1 Radial number density (n)

To investigate the effects of confinement on monomer packing, we look at the number density of chain monomers (n) as a function of distance from the center of a confining sphere (r/a).

Fig. 4.3 shows the reduced number density of monomers ($n(r/a) * a^3$) for various densities (ρ). At low polymer density (ρ=0.017, $r_c/a = 13$), most monomers stay away from the confining surface (Fig.4.3 a)), to maximize their entropy. We can see in Fig. 4.3 a) that in stiffer chains the density peaks closer to the confining surface to minimize their bending energy cost. For the chain with small bending rigidity ($l_p/a = 8$), most monomers are located near center of the confining sphere, similar to what was observed in simulations of a flexible chain, confined inside a sphere of size bigger than its unconfined size ($r_c > r_G$) [93, 120]. This indicates that the chain of this stiffness ($l_p/a = 8, r_c/a = 13$) is not strongly affected by the presence of the confining surface and behaves similar to an unconfined chain, forming disordered configuration at this density, as seen in the inset of Fig. 4.3 a). This observation is supported by findings of several authors, who have shown that confined semi-flexible chains start wrapping a confining surface at $l_p \sim r_c$ and at $r_c >> l_p$ they form a disordered structure [28, 29].

As polymer density increases up to $\rho = 0.083, r_c/a = 7.6$, the environment inside the cavity becomes significantly crowded and entropic penalty due to monomer-monomer interactions increases in all the confined chains. This results in a large force driving each confined chain to position closer to the periphery, where it wraps around the sphere (inset of Fig. 4.3 b)) to minimize its bending energy cost. Therefore, the number density peak shifts towards the surface, as seen in Fig. 4.3 b).

As the degree of confinement increases up to $\rho = 0.127, r_c/a = 6.6$, the maximum of the number density for all chains shifts even further (Fig. 4.3 c)). At this density, the molecule forms thin ordered polymer layer, close to the confining surface (i.e. inset of Fig.4.3 c)), similar to that observed in simulations [96, 118, 132] and experimental [23, 133] studies of bacteriophage DNA at relatively high density.

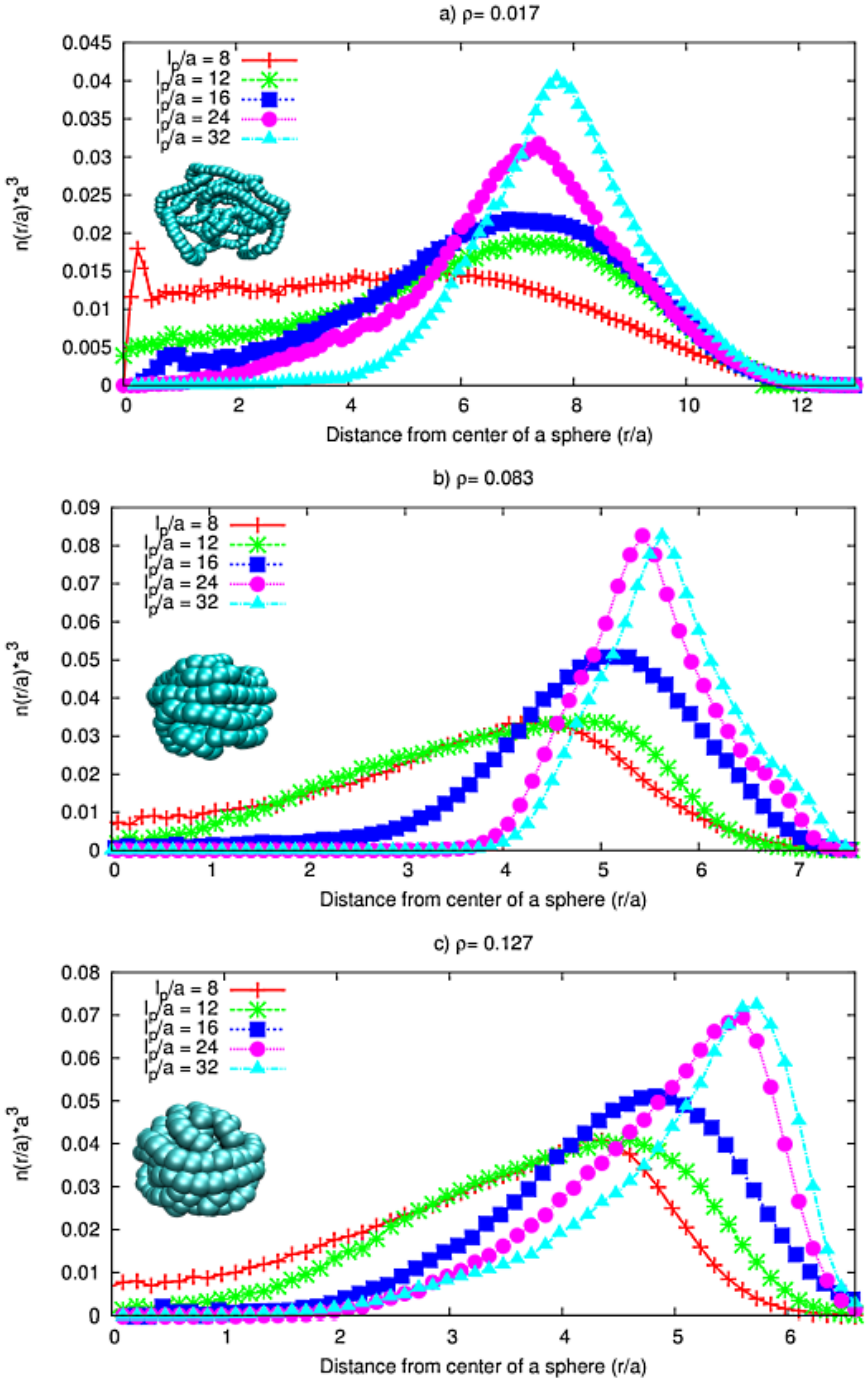

Figure 4.3: Radial number density of confined chains ($n(r/a) * a^3$) as a function of distance from center of a sphere (r/a) at a) $\rho = 0.017$, $r_c/a = 13$, b) $\rho = 0.083$, $r_c/a = 7.6$, and c) $\rho = 0.127$, $r_c/a = 6.6$. Legends represent persistence length (l_p/a). The inset pictures are conformations of the chain with $l_p/a = 8$ confined inside a sphere of $r_c/a = 13$ a), $r_c/a = 7.6$ b) and $r_c/a = 6.6$ c).

4.3.2.2 Order parameters

To examine how the density of confined polymer affects the transition from disordered to ordered states (I-N), nematic order parameter (Q) was calculated using surface bonds lying in the shells between the confining surface and the shell, where the number of monomers was the highest. To obtain good statistics, the value of Q was averaged over 8000 polymer conformations. $\langle Q \rangle$ can have value from -0.5 to 1. $\langle Q \rangle = 1$ indicates a perfectly ordered state, which never happens in practice due to strong frustration caused by spherical confinement. $\langle Q \rangle = 0$ denotes isotropic state, which only happens in an infinitely large system and $\langle Q \rangle = -0.5$ implies a state with long-axis of a polymer lying perpendicular to its director ($\hat{n}$). The director $\hat{n}$ is a unit vector in the direction of an eigenvector, corresponding to the largest eigenvalue of the tensor $\mathbf{T}$, constructed from all surface bonds, in Eq. 2.26. We chose $\langle Q \rangle < 0$ as a nematic parameter, since its $|\langle Q \rangle|$ is the highest among other eigenvalues of the tensor $\mathbf{T}$ (see Chapter 2 for details).

Fig. 4.4 shows the absolute value of nematic order parameter ($|\langle Q \rangle|$) and its mean square fluctuation (Σ_Q^2), for all chains, as a function of polymer packing density (ρ). At the same density, the nematic order parameter tends to be larger for chains with larger persistence length. This happens because stiffer chains are affected by a larger entropic penalty due to their larger size, forcing them to reorient and wrap around the interior of the confining surface (Fig. 4.2). As a result, stiff chains become more ordered.

As the polymer density increases, confined chains are driven by an increasing of entropic penalty, and wrap tighter around the confining surface, forming an ordered layer with larger values of $|\langle Q \rangle|$, compared to those found at lower densities.

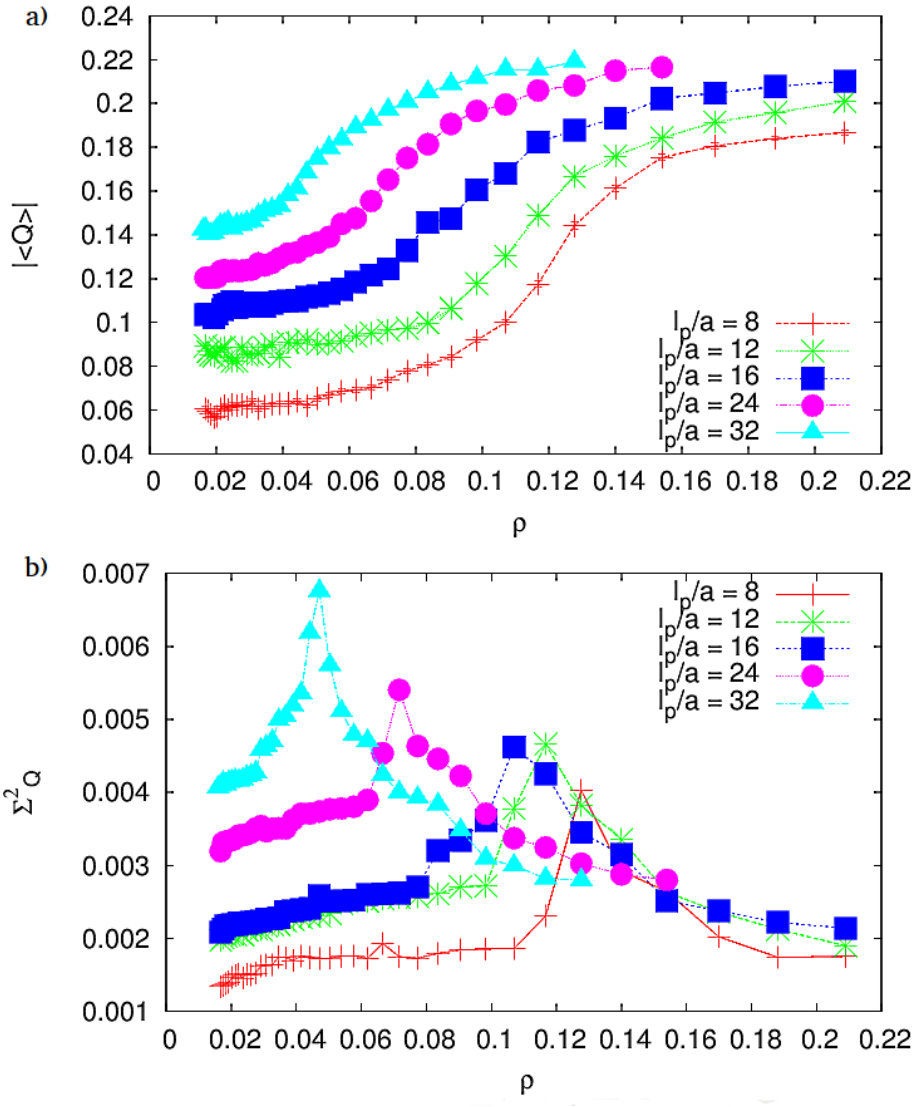

Figure 4.4: a) Absolute value of the nematic order parameter ($|\langle Q \rangle|$) and b) mean square fluctuation of Q (Σ_Q^2). Legends display persistence length (l_p/a). Lines are drown to guide the eye.

Fig. 4.4 a) shows that the $|\langle Q \rangle|$ increases continuously, forming S-shaped curve, as a degree of confinement increases. The S-shaped change of the order parameter can be interpreted as a transition from disordered to ordered state [26, 71] and the point of steepest change can be considered to be phase transition density. This point shifts to higher polymer densities as l_p/a decreases, which means that less stiff chains need to be compressed into a smaller sphere to form a structure with the same degree of order, as observed in stiff chains. Shift of the I-N transition density to higher values as the chain stiffness decreases was also observed in simulations of semi-flexible chains confined in a slit [134] and rod-like molecules confined in a cubic cell [135].

Theoretically, the transition between phases is only defined in thermodynamic limit i.e. a system with infinite number of particles and volume, but finite density, and only in equilibrium [70]. Such a system exists only in theory but it can be a good approximation for the bulk system if the system's surface to volume ratio is negligibly small. We use this approximation when describing phase transition. The transition from disordered to ordered state (I-N) is usually accompanied by a spike in the mean square fluctuation (Σ^2) of the nematic order parameter (Q), which can be seen in Fig.4.4 b). The values Σ_Q^2 of all semi-flexible chains slightly increase at low polymer densities, and spike as the system transitions into ordered state. The location of Σ_Q^2 corresponds to the transition density of a semi-flexible chain. In theoretical study, the I-N transition of ϕ29 DNA of length 7 μm, $l_p = 50$ nm, $d = 2.5$ nm was also predicted to happen at low volume density ($r_c = 100$ $nm, \rho = 0.01$) or as soon as the DNA phage are injected inside a sphere for only few percents of its length [27]. However, results from MD simulations of multiple semi-flexible chains, each having 32 monomers and $l_p/a = 32$ predicted that I-N transition happens at a high density $\rho \approx 0.27$ [35]. It is worth noting that $\langle Q \rangle$ in our study was calculated using all surface bonds located between the confining sphere and the shell holding the largest number of monomers. Thus, there was a significant number of internal bonds with uncorrelated directions involved in the calculations, making $\langle Q \rangle$ relatively small. In some other studies [136, 137], $\langle Q \rangle$ was averaged over Q_{local}, calculated using only a bond and its nearest neighbors located within the distance of a bond length, compared to the selected bond, giving higher $\langle Q \rangle$ at I-N transition ($\langle Q \rangle \sim 0.49$).

To determine the type of ordered structure formed by the confined chain, we calculated P_2 and q_3 order parameters. Parameter P_2 identifies the presence of

helicoildal structure. It is calculated in the same manner as Q, using the second rank Legendre polynomial, with an angle between a bond vector ($\vec{u_i}$) and $\hat{e}_\phi$ used in the calculation of P_2 (instead of the angle between the surface bond $\vec{u_i}$ and director $\hat{n}$, used in the calculation of Q). Thus, P_2 probes how well the surface bonds align along the latitude of the confining sphere, while Q probes the alignment of the surface bonds relative to their average director $\hat{n}$.

Fig.4.5 shows average value of P_2 order parameter ($\langle P_2 \rangle$) for confined chains of

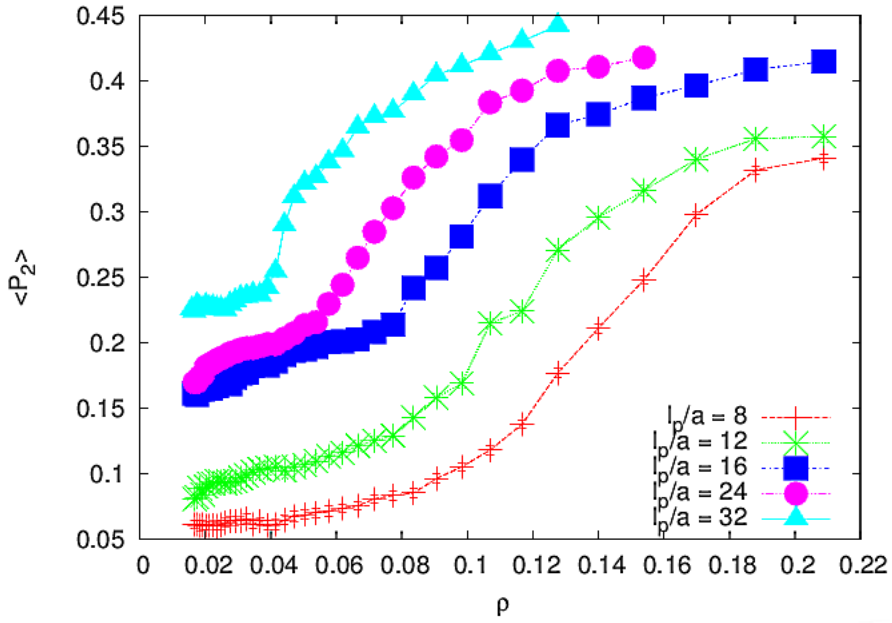

Figure 4.5: Helicoildal order parameter ($\langle P_2 \rangle$). Legends represent persistence length (l_p/a).

various persistence length. The values of $\langle P_2 \rangle$ increase gradually with ρ, with highest rate of change happening at density of phase transition from disordered to ordered state. The values of $\langle P_2 \rangle$ computed for the chains with lower persistence length rise later than those calculated for the chains with higher persistence length, in agreement with the later transition of these chains into ordered state. Although we observed the increase in the values of $\langle P_2 \rangle$, as the polymers transitioned into ordered state, they remained far from unity even for the chain with the largest persistence length. This indicates the structures observed in the current study are not even close to a perfectly helicoildal structure. This finding is similar to that observed

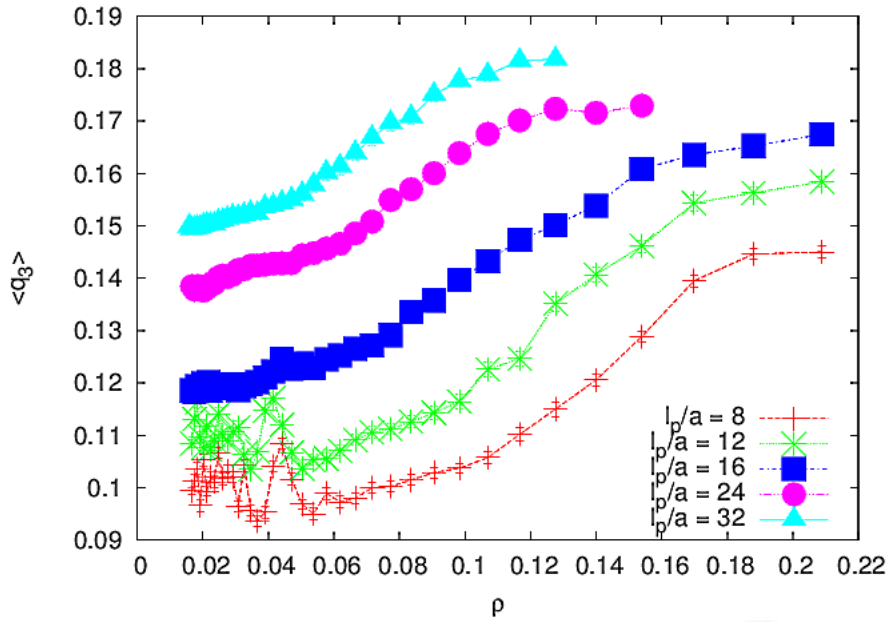

Figure 4.6: Average value of the tetrahedral order parameter ($\langle q_3 \rangle$). Legends display persistence length (l_p/a).

when rod like molecules were confined inside a sphere in the work of Trukhina *et al.* [71].

The tetrahedral order parameter (q_3) was also calculated, and its average value is shown in Fig. 4.6. The values of $\langle q_3 \rangle$ increase in the same manner as the $\langle P_2 \rangle$ values; however, the values of $\langle q_3 \rangle$ are noticeably lower than $\langle P_2 \rangle$. At low polymer density, the $\langle q_3 \rangle$ values for the chains with $l_p/a = 8$ and $l_p/a = 12$ fluctuate; therefore, it is possible that the surface bonds temporarily approach the tetrahedral structure while the chains are compressed into a cavity with radius of approximately l_p. The presence of a tennis ball texture at low polymer concentration was also reported by Angelescu *et al.* [127], when a polyelectrolyte chain with $l_p/a = 45$ was confined on a spherical surface with confining radius of approximately l_p/a. In Fig. 4.6, the values of $\langle q_3 \rangle$ calculated for the chains with larger persistence length rise earlier, compared to those calculated for less rigid chains, with an increase in polymer density. This indicates that tetrahedral structures become more pronounced, when chain persistence length increases.

4.3.2.3 Defects in nematic liquid crystals

In this section, we will investigate surface distribution of defects in a chain confined to the density above I-N transition. To identify the defect locations, local nematic order parameter (Q_{local}) was calculated for all surface bonds. Calculation of each Q_{local} value involved a selected bond and its neighboring bonds located within the distance $r/a = 1.2$ from the selected bond. Location of Q_{local} was then defined with respect to polar angles (θ, ϕ), and averaged over 8000 conformations. Here the angle θ is the angle between a selected bond and $\hat{e}_z$, ϕ is the angle between projection of midpoint of the selected bond and $\hat{e}_x$. Vector $\hat{e}_z$ is defined as a unit vector in the direction, corresponding to the largest eigenvalue of the $\mathbf{T}$ tensor, constructed from all bonds in a considered chain. $\hat{e}_x$ and $\hat{e}_y$ are unit vectors in the directions the eigenvalues are second largest and smallest, respectively.

At density $\rho \approx 0.2$, and l_p/a ranging from 8 to 16, the chains form structures with four regions of low Q_{local} ($Q_{local} < 0.375$ for the chain with $l_p/a = 12 - 16$ and $Q_{local} < 0.25$ for the chain with $l_p/a = 8$). These regions appear in places where the strands that aligned in vertical and horizontal directions intersect with each other, as shown in Fig. 4.7 b), c), and f). The surface arrangement of polymer in this case is similar to the field observed around a charge of +1/2 on a spherical surface in Fig. 1.6 b), and that observed earlier in MC simulations of a thin uniform nematic shell with planar anchoring [34, 138], as well as that found in MD simulations of a system of multiple semi-flexible chains, each having 32 monomers and $l_p/a = 32$, confined inside a sphere at $\rho = 0.7$ [35]. Surface distribution of Q_{local} plotted in polar coordinates is shown in Fig. 4.7 a), c), and e) for the chains with $l_p/a = 16$, 12, and 8, respectively. Spots with low Q_{local} values (black and red colors) nearly merge at $\theta \sim 1.57$ $(\pi/2)$ radians for the chain with $l_p/a = 12$ (Fig. 4.7 c)), and distribute randomly in the chain with $l_p/a = 8$ (Fig. 4.7 e)). For the stiffer chain with $l_p/a = 16$, defect locations (low Q_{local}) are better localized, compared to the other two chains (Fig. 4.7 a)). Defects in this chain ($l_p/a = 16$) appear at $\theta \geq 1.5$ and $\theta \leq 1$ radians and defect in each pair is separated by the angle of $\phi \approx 3$ radians.

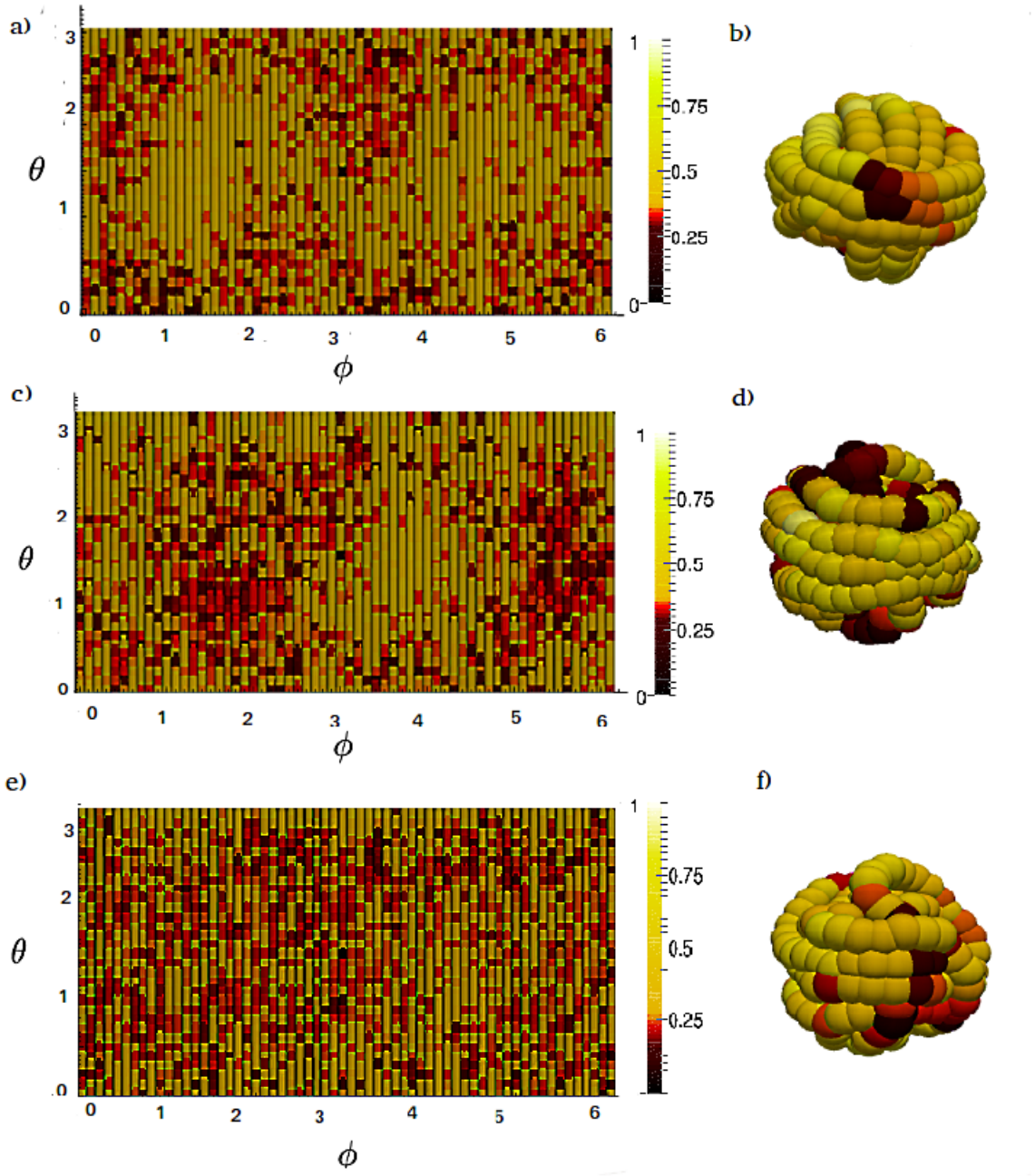

Figure 4.7: Distribution of Q_{local} on a spherical surface and snapshots of the chains with $l_p/a = 16$ (a) and b)), $l_p/a = 12$ (c) and d)), and $l_p/a = 8$ (e) and f)), confined at density $\rho \approx 0.2$.

For the chains with $l_p/a = 24$ and 32, confined in a sphere at density $\rho \approx 0.127$, we also observe areas with low value of Q_{local} at intersections of horizontal and vertical strands, and the chains locally form structures with four $+1/2$ defects. At this density, the chains still have available volume to move around, and most defect locations are not well defined, appearing in the range of $1 < \theta < 3$ radians, as seen in Fig. 4.8.

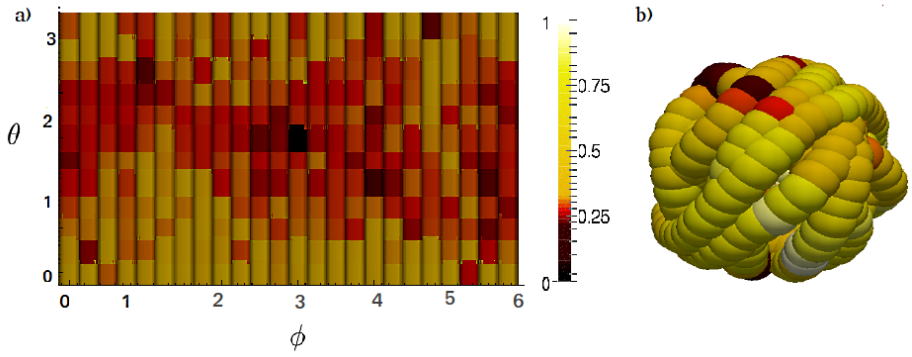

Figure 4.8: Distribution of Q_{local} on a spherical surface (a)) and snapshots (b)) for the chains with $l_p/a = 24$, confined at density $\rho \approx 0.127$.

Although, we observe the structure with four $+1/2$ defects in all considered polymer conformations at densities higher than their I-N transition density. The defects are better localized in stiffer chains with sufficiently high density such as in a chain with $l_p/a = 16$ at $\rho \approx 0.2$ in Fig. 4.7 a). The defect arrangement we see is far from the ideal tennis ball structure, which should satisfy the following symmetries:

$$\varrho(\theta, \phi; \varphi) = \varrho(\theta, 2\pi - \phi; \pi - \varphi) = \varrho(\theta, \phi; \pi + \varphi) = \varrho(\theta, \pi + \phi; \varphi) = \varrho(\pi - \theta, \phi + \pi/2; -\varphi), \tag{4.1}$$

where $\varrho(\theta, \phi; \varphi)$ is the positional and orientational distribution function of a structure, θ and ϕ are polar and azimuthal angles, and φ is the angle between a bond and latitude. Fig. 4.9 shows 3 - and 2- dimensional plots of an ideal tennis ball structure. In the 3-dimensional plot of the tennis ball structure type B (Fig. 4.9 (a)), at angle $\phi = 0$, $\pi/2$ and π radians, the segments align along the meridians ($\varphi = 0$), while in the region $\phi = (0, \pi/2)$ radians the segments have azimuthal angle higher than zero, as seen in Fig. 4.9. In equatorial plane $\theta \sim \pi/2$ radians, the

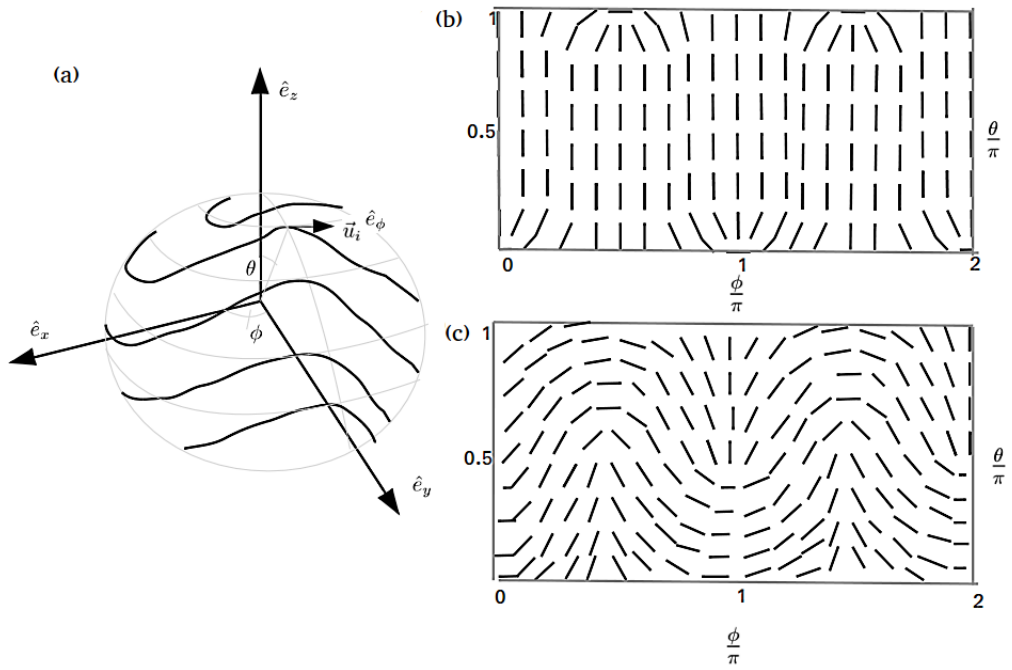

Figure 4.9: 3-dimensional plot of an ideal tennis ball type B (a). 2-dimensional plot of segment arrangement of an ideal tennis ball type A (b) and type B (c).

segments form most anisotropic structure. Tennis ball structure type A satisfies the same symmetries (Eq. 4.1); but most segments in type A structure arrange along the meridians of a sphere. In 2-dimensional plot of the tennis ball A (Fig. 4.9 (b)) and B (Fig. 4.9 (c)), each defect is separated by the angle of $\phi = \pi/2$ radians from the defects, but each defect pair in the type A move closer to poles, compared to that in type B structure.

Defect locations in our 2-dimensional plot for the chain with $l_p/a = 16$ at $\rho \approx 0.2$ in Fig. 4.8 (a) are similar to those in the tennis ball A in Fig. 4.9 (a), but in our study we see them separated by the angle $\phi > \pi/2$ radians. Also, most monomers in a chain are neither parallel nor perpendicular to equatorial plane. In the chain with $l_p/a = 24$ at $\rho \approx 0.127$ (having Q_{local} values less than 0.375) the bonds form a strip between $1 < \theta < 3$ radians, similar to that observed in ideal tennis ball B (Fig. 4.9 (c)), but the defect locations are not well localized. Thus, our observed structure (Fig. 4.8 (b)) is far from the ideal tennis ball B, shown in Fig. 4.9 (a). As reported in recent MD simulations of multiple semi-flexible chains, each having

32 monomers and $l_p/a = 32$, the tennis ball structure was obviously observed only at $\rho \geq 0.4$ [35].

4.4 Conclusions

We applied the EXEDOS algorithm to investigate isotropic-nematic (I-N) transition and spatial conformations of a semi-flexible chain with persistence length (l_p/a) ranging from 8 to 32, confined inside a cavity of various radii. Tetrahedral order parameter (q_3) was used to identify structures with four +1/2 defects and helicoildal order parameter (P_2) was used to test the presence of helicoildal configuration.

The confinement free energy increases with an increase in degree of confinement. Bending energy increases linearly at low monomer densities, as the confined chains tend to be in a rather disordered state with high entropy. At higher densities, bending energy grows slower as the confined chain wraps around a confining boundary to minimize bending energy cost.

I-N transition was observed for all the confined chains. In chains with higher persistence length, the transition occurs at lower densities, compared to less rigid chains. As polymer density increases, the values of both $\langle P_2 \rangle$ and $\langle q_3 \rangle$ increase monotonously, forming S-shaped curve characteristic to I-N transition. However, $\langle P_2 \rangle$ and $\langle q_3 \rangle$ remain far from the values of 1 and 0.745 [38], expected for a perfect helicoildal and tetrahedral structure, respectively.

Local nematic (Q_{local}) order parameter was calculated for each chain configuration. It was found that above the density of I-N transition, the layer of polymer adjacent to the confining sphere forms a structure with four +1/2 surface defects, and each defect is located at intersection between horizontal and vertical monomer strands. At high density ($\rho \approx 0.2$), the defect locations are better localized in a stiffer chain ($l_p/a = 16$), which exhibits two defects close to the poles. In less stiff chain ($l_p/a = 12$), defect locations move closer to equator, while in the more flexible chain ($l_p/a = 8$) the defects distribute randomly. For the chains with $l_p/a = 24 - 32$ at intermediate density ($\rho \approx 0.127$), most defects appear between $1 < \theta < 3$ radians. It is worth noting that structures with four +1/2 defects that we observed, are far from both ideal tennis ball and tetrahedral structure.

Chapter 5

Effects of chain flexibility on molecular arrangement of a spherically confined semi-flexible chain

5.1 Introduction

Molecular flexibility plays a crucial role in various biological processes, such as gene replication and expression, DNA-protein recognition, and gene regulation, which require structural, functional and biochemical specificity between coupling molecules. Flexibility of biomolecules can be affected by several factors such as temperature, salt concentration, and steric hindrance. The latter effect restricts molecular movement and is believed to play a role in formation of ordered structures in sufficiently dense systems.

One of the key parameters that determines steric interaction between particles in a system is a ratio of diameter (d) to a bond length (a) of particles. Depending on the value of this ratio, molecules can form different spatial arrangements. For example, MC simulations have shown that a system of dumbbell shaped particles with $d/a \geqslant 2.5$ forms a face-centered cubic structure at a density of around 0.57, while particles with $d/a < 2.5$ form a fluid phase at the same density [139].

When particles are confined to a curved surface, the interplay between surface

curvature and particle-particle interaction leads to a structure different from that formed by particles confined on a flat surface. For example, hexagonal lattice was observed when spherical particles were packed on a flat surface, while the structure with a minimum of 12 point defects was found, when these particles were confined on a spherical substrate [140]. The d/a ratio further affects the particle distribution on the curved surface. Rods of length a, and diameter d confined on a spherical surface have been found to form a global baseball-like structure at $d/a \geqslant 0.5$, but stayed disordered at smaller d/a when confined on the same sphere [141].

For a semi-flexible polymer, local stiffness or persistence length (l_p) is believed to be the main factor determining the chain ordering. A well known example of the effect that the l_p has on the chain structure is formation of double stranded (ds) RNA, and ds DNA. The ds RNA typically have $l_p \sim 60 \pm 10$ nm [142, 143], approximately $20-30\%$ larger than that of the ds DNA, forms right-handed duplex in 'A-form' with shorter distance between base-pairs [144], while ds DNA forms duplex in 'B form' [145]. MC simulations of a semi-flexible chain show that altering the d/a ratio changes correlation in its bond-orientation, consequently l_p is affected. With $d/a = 0$ or without excluded volume, the bond-orientation correlation of the chain falls off exponentially, but that of the chain with excluded volume decays algebraically [146, 147].

In the current chapter, we investigate the effects that d/a ratio has on the molecular organization of a semi-flexible chain, confined inside a sphere, using the EXEDOS algorithm. In simulations for this chapter, we varied the d/a ratio of a pearl-necklace chain of 300 monomers with stiffness $l_p/a = 16$ between $d/a = 1$ to $d/a = 0$. In addition, a fused-sphere chain was modeled by adding overlap monomers at the middle point of each bond in a pearl-necklace chain of diameter $d/a = 1$ (see Fig. 5.1 d)). The ratio d/a in the fused-sphere chain was equal to 2.0, and number of monomers was 599. Allowing overlap between monomers adds flexibility to the chain, since the chain can now bend and twist more, compared to the pearl-necklace chain of the same diameter, where overlap is forbidden. The bending stiffness of $l_p/a = 16$ was used to test if excluded volume interaction will affect the distribution of monomers and formation of an ordered structure.

The methodology applied in the study and major findings will be described in the following sections.

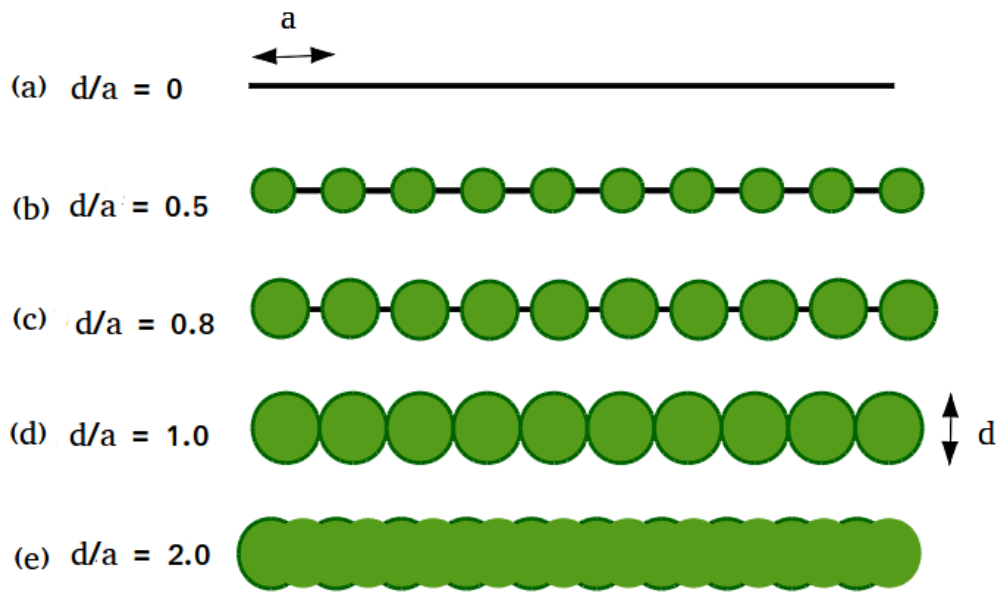

Figure 5.1: Illustration of pearl-necklace chains ((a)-(d)) and a fused-sphere chain ((e)) with various monomer diameters.

5.2 Simulation procedure

Four pearl-necklace chains were composed of 300 monomers with diameter d/a of 0, 0.5, 0.8 and 1, and one fused-hard sphere chain was made of 599 monomers of diameter $d/a = 2$, where a is a bond length, as shown in Fig. 5.1. The Metropolis MC simulations of 10^8 MC steps were conducted on all chains while keeping bending rigidity in Eq. 2.21 constant for all chains with $l_p/a = 16$. The radius of a confining sphere varied from $r_c/a = 4.8$ to $r_c/a = 9$ for the ideal chain ($d/a = 0$), and from $r_c/a = 5.6$ to $r_c/a = 9$ for self-avoiding chains of $d/a = 0.5$, and $d/a = 0.8$. For the pearl-necklace chain with $d/a = 1$ and the fused-sphere chain with $d/a = 2$, radius of a confining sphere varied from $r_c/a = 6$ to $r_c/a = 9$ and $r_c/a = 6.6$ to $r_c/a = 9$, respectively. Each spherical volume was divided into a number of shells, each shell having thickness dr/a =0.2. The chains were initially confined inside a sphere by placing their center of mass at the center of the sphere. In order to change a chain configuration, a randomly selected monomer was rotated about the bond connecting its adjacent neighbors. The chain was then tested for monomer-monomer overlap using the listed cell method [39]. No overlap was allowed for the pearl-necklace chains

with excluded volume. For the self-avoiding fused-sphere chain, overlap of consecutive monomers was allowed (but was not allowed for non-consecutive monomers), giving it more bending freedom, compared to the self-avoiding pearl-necklace chains. After checking for excluded volume interaction, acceptance criterion of the EXEDOS method was applied. Simulations then followed the procedure described in Chapter 2. The convergence was considered acceptable when modification factor ($ln(c)$) was 10^{-6}, and flatness constant was 0.65. The density of states obtained for each chain in this manner was then used as an acceptance probability in production runs.

Radial distribution of monomers (p), nematic order parameter (Q), helicoildal order parameter (P_2) and tetrahedral order parameter (q_3) were calculated from 8000 equilibrium configurations of each chain located in the shell. To calculate the radial distribution of monomers in each chain, the number of monomers was counted in each shell, and divided by the total number of monomers in a sphere, and then normalized to give $\int p(r/a)dr/a = 1$. In order to calculate all order parameters (except the Q_{local}), we used the surface bonds, which are the bonds located between the confining surface and the shell having the highest number of monomers. The local nematic order parameter (Q_{local}) and Q were calculated using the same equation (Eq. 2.26), but here we used surface bonds plus their nearest neighbors located within a distance of $r/a = 1.2$ from any considered bond.

Each considered chain had different d/a ratio; thus, when confined inside a sphere of the same size, they occupied different volumes. Volume fraction η is calculated as: $\eta = \rho v_n$, where $\rho = 1.0/V$ is density of the chain, $V = 4\pi r_c^3/3$ is volume of the system, r_c is radius of a confining sphere and v_n is volume of the chain. For a fused-sphere chain, v_n was calculated as [148] : $\pi d^3(1+(N/2)*(3(a/d)-(a/d)^3)/6$, where a is the bond length, d is monomer diameter and N is the number of bonds. For self-avoiding pearl-necklace chains, v_n was calculated as: $(N+1)*\pi d^3/6$. Table 5.1 shows volume fractions for the chains with different d/a ratios, confined in spheres of various diameters.

r_c/a	η $(d/a = 0.8)$	η $(d/a = 1.0)$	η $(d/a = 2.0)$
6.60	0.07	0.13	0.18
7.00	0.06	0.11	0.15
7.40	0.05	0.09	0.13
7.80	0.04	0.08	0.11
8.00	0.03	0.07	0.10

Table 5.1: Volume fraction (η) of the pearl-necklace chains with $d/a = 1$ and $d/a = 0.8$, made of 300 monomers, and a fused-sphere chain with $d/a = 2$ made of 599 monomers. All chains are confined inside a sphere of various radii (r_c/a).

5.3 Results and discussion

5.3.1 Radial positioning of chains

In this section, we investigate how chain flexibility due to steric hindrance affects spatial distribution of monomers inside a confining sphere. When confined inside a large sphere ($r_c/a = 8.0$), chains with smaller d/a are affected by weaker force due to smaller monomer-monomer overlap, compared to the chains with bigger d/a. Therefore, we observe that chains with smaller d/a position themselves further away from the surface, compared to the pearl-necklace chains with larger d/a (Fig. 5.2 a)). Similar behavior was also observed in simulations of a polymer confined at low system concentration: the chain of smaller diameter stayed further away from the surface, compared to the chain with bigger diameter, at the same degree of polymerization [96].

As the confining radius decreases to $r_c/a = 6.6$, all chains are affected by force due to monomer-monomer overlap, and the monomer radial distribution peaks closer to the surface, as shown in Fig. 5.2 b). As expected, probability that monomers will be in a shell of given radius, for the chains of larger d/a peaks closer to the surface, since their size is bigger. Monomers push each other away from the center of confining sphere, and position closer to the surface. Since, overall density is not large, and the force resulting from the interaction between monomers and surface is weaker than force due to monomer-monomer interaction. Thus, we observe the monomer radial distribution of bigger chains peaks closer to the surface.

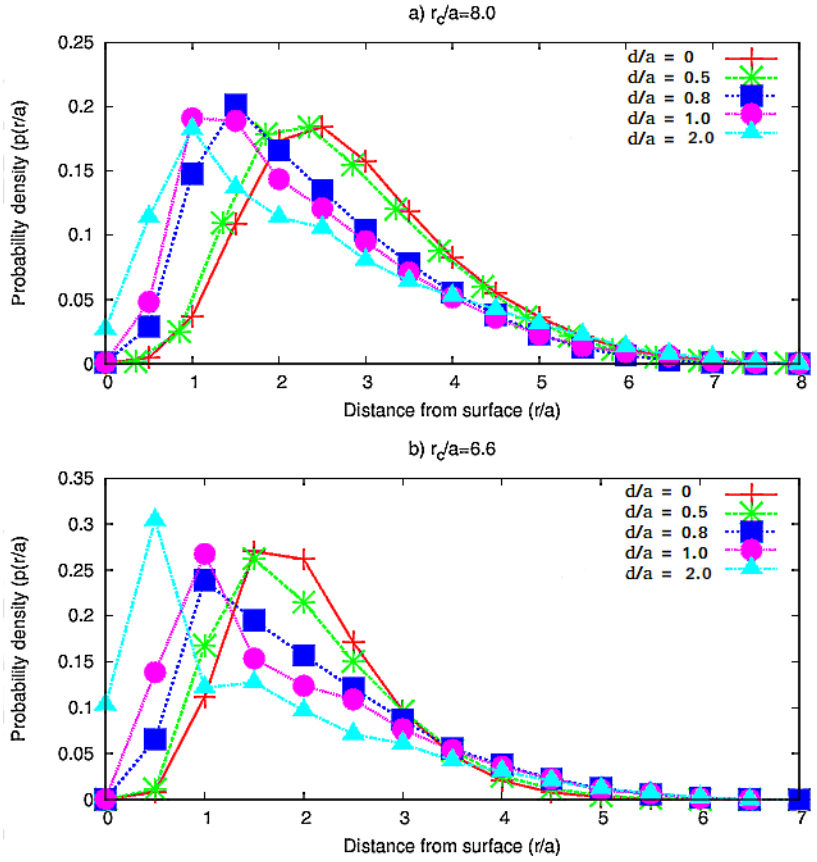

Figure 5.2: Probability distribution $p(r/a)$ for chains of different diameters, confined inside a sphere of a) $r_c/a = 8.0$ and b) $r_c/a = 6.6$, as a function of the distance from a confining surface.

5.3.2 Order parameters

5.3.2.1 Nematic order parameter (Q)

To investigate the effects of steric hindrance on spatial organization of a confined chain, the nematic order parameter (Q) was calculated as a function of a confining radius (r_c/a), as shown in Fig. 5.3. Q is the largest eigenvalue of the tensor $\mathbf{T}$ defined in Chapter 2 and constructed from surface bonds of a confined polymer. $\langle Q \rangle$ is an average value of the nematic order parameter and the larger $|\langle Q \rangle|$ value implies that the structure is more ordered. In Fig. 5.3 a), we can see that the ideal chain and the chain with $d/a = 0.5$ are highly disordered, with a small value of order parameter $\langle Q \rangle$. It remains nearly constant regardless of the diameter of a

confining sphere and approximately the same as $\langle Q \rangle$ obtained for disordered chain (large r_c/a). For the chains with $d/a \geqslant 0.8$, the order parameter $\langle Q \rangle$ follows an S-shaped curve, showing the transition of structure from disordered to ordered state, similar to that observed in other numerical studies of a semi-flexible chain and rods confined inside a sphere [26, 130]. The chain with a larger d/a forms a more ordered structure, compared to those with smaller d/a. This observation is in a good agreement with a study of dumbbell shaped particles confined inside a cubic cell, where a system of bigger dumbbells formed more ordered structure [139]. In our model, as a confining radius decreases just below $r_c/a = 8.0$, the fused-sphere chain becomes highly ordered. At this concentration, its $\langle Q \rangle$ value increases sharply before leveling off at $r_c/a < 7.4$, and the monomers form thin ordered layer close to the surface, similar to that observed in theoretical studies of a long semi-flexible chain confined inside a sphere of size much smaller than persistence length [118]. Similar transition from disordered to ordered state was observed in a self-consistent mean field study of a semi-flexible fused-sphere chain (20 monomers, $l_p/a = 50$, and $d/a = 1$), when the monomer concentration increased from $\eta = 0.14$ to 0.18 [149]. The $\langle Q \rangle$ value of the fused-sphere chain is significantly larger than that of other chains, confined inside spheres of sizes smaller than $r_c/a = 7.4$ because the fused-sphere chain system has better ability to bend and wrap around the confining surface.

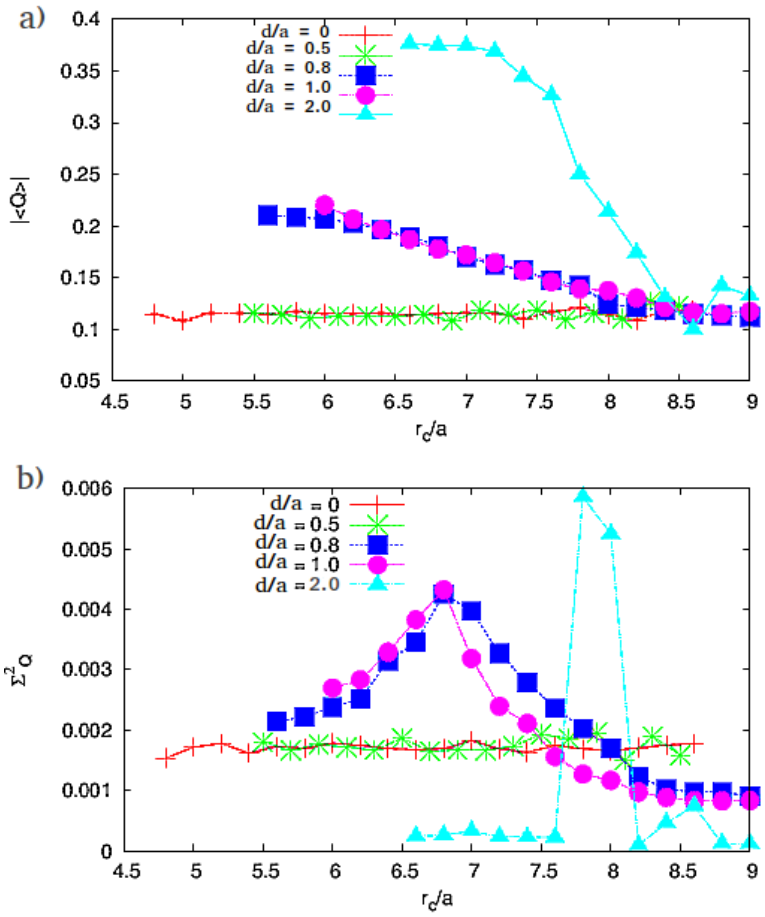

Figure 5.3: a) Average value of nematic order parameter ($\langle Q \rangle$) for chains ($l_p/a = 16$) with different d/a. b) Mean square fluctuation of Q (Σ_Q^2). Degree of polymerization of the pearl-necklace chains is 300, and that of the fused-sphere chain is 599; a is a bond length and d is monomer diameter.

Together with S-shaped profile of $\langle Q \rangle$, isotropic-nematic (I-N) transition is reflected in a peak of mean square fluctuation of the Q values (Σ_Q^2). At transition concentration, Σ_Q^2 spikes as the system reorients itself to the more ordered state. Fig. 5.3 shows Σ_Q^2 for all chains, as a function of a confining radius. For the fused-sphere chain, the mean square fluctuation Σ_Q^2 peaks at higher radius ($r_c/a = 7.8, \eta = 0.11$), than for the other chains, indicating earlier transition of this chain to ordered state. Indeed, the chain with larger d/a is more concentrated, when confined inside the same sphere, compared to the other two pearl-necklace chains. Transition from disordered to ordered state at similar concentrations ($\eta = 0.15$) was also observed

in theoretical study of a semi-flexible fused-sphere chain of 20 monomers, $d/a > 1$ and $l_p/a = 50$ [149]. Chain flexibility and size of the pearl-necklace chains with $d/a = 0.8$ and $d/a = 1$ do not considerably affect their I-N transition: the Σ^2_Q of these two chains peaks at nearly the same concentration (at $r_c/a = 6.6$, η=1.4 and 1.3 for the chain of $d/a = 1$ and $d/a = 0.8$, respectively.)

5.3.2.2 Helicoidal order parameter (P_2)

To test confined chains for the presence of helicoidal arrangement, helicoidal order parameter (P_2) was calculated for all considered chains using Eq. 2.29. Fig. 5.4 a) shows $\langle P_2 \rangle$ curves for four pearl-necklace chains and a fused-sphere chain. For chains with $d/a \geqslant 0.8$, $\langle P_2 \rangle$ follows S-shaped curve, indicating the transition to helicoidal ordered states, as r_c/a decreases. As long as excluded volume is small (the ideal chain and the chain with $d/a = 0.5$), $\langle P_2 \rangle$ remains small and constant at all confining radii, implying that these two chains remain disordered at all considered concentrations. The rate of change of $\langle P_2 \rangle$ in the fused-sphere chain is large, compared to that of the pearl-necklace chains with $d/a = 0.8$ and $d/a = 1$. At larger confining radii, $\langle P_2 \rangle$ of the fused-sphere chain decreases faster. On the other hand, at smaller radii, the fused-sphere chain better conforms the surface because of its high flexibility, and larger size, making the surface layer more dense and more ordered.

The transition concentration from disordered to helicoidal state can also be seen in the mean square fluctuation of P_2 ($\Sigma^2_{P_2}$). Although, we observe a slight increase in $\langle P_2 \rangle$ for the chain with $d/a = 0.5$, $\Sigma^2_{P_2}$ does not present a peak, implying that the chain remains disordered at all considered concentrations. For the chains with $d/a = 0.8$ and 1.0, $\Sigma^2_{P_2}$ peaks at approximately the same concentrations as Σ^2_Q.

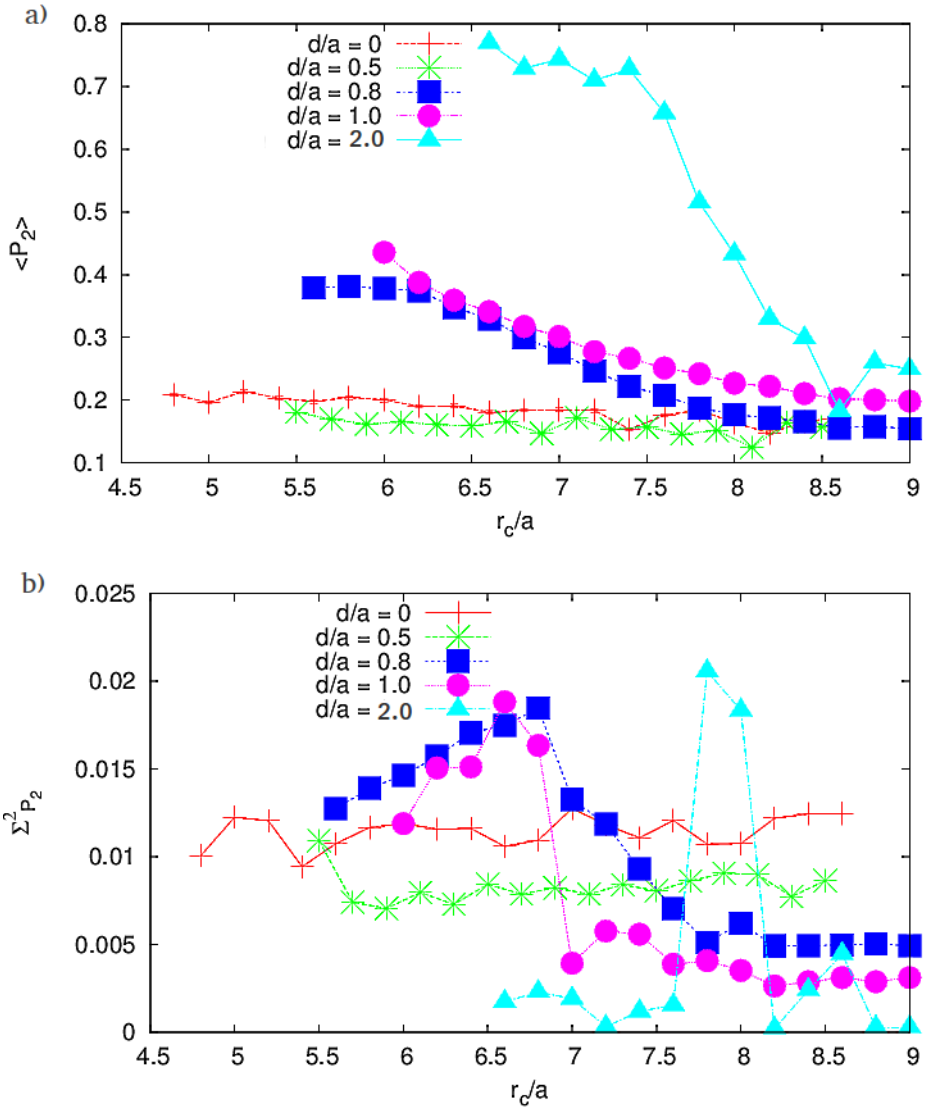

Figure 5.4: a) Average value of the helicoildal order parameter ($\langle P_2 \rangle$) for chains with various d/a and $l_p/a = 16$, b) Mean square fluctuation of P_2 ($\Sigma^2_{P_2}$).

5.3.2.3 Tetrahedral order parameters (q_3)

In this section, we look for the presence of tetrahedral ordering in the chains of various flexibilities, by calculating the tetrahedral order parameter (q_3). The $\langle q_3 \rangle$ is calculated using Eq. 2.30, averaged over 8000 configurations, and shown in Fig. 5.5. The values of $\langle q_3 \rangle$ for an ideal chain and the chain with $d/a = 0.5$ remain constant at all considered radii, since the structures remain disordered at all considered con-

centrations. For the chains with $d/a \leq 0.8$, $\langle q_3 \rangle$ curves are similar to those obtained for $\langle Q \rangle$ and $\langle P_2 \rangle$. The mean square fluctuation of q_3 ($\Sigma^2_{q_3}$) for the pearl-necklace chains of $d/a = 0.8$ and $d/a = 1$ peaks slightly earlier, compared to that of $\langle Q \rangle$ and $\langle P_2 \rangle$. This may imply that while forming an ordered arrangement, these two chains first approach a tetrahedral configuration, but later reorder themselves into a shape closer to helicoildal. Similar findings were also reported in theoretical studies of a rods confined on a spherical surface [150]. For the fused-sphere chain, the $\Sigma^2_{q_3}$ peaks at the same concentration (same r_c/a) as $\langle Q \rangle$ shown in Fig. 5.3 and $\langle P_2 \rangle$ in Fig. 5.4. This can be expected, since the fused-sphere chain has high flexibility, and the ordered state here forms relatively fast, making it hard to distinguish the fine interplay between tetrahedral and helicoildal configurations.

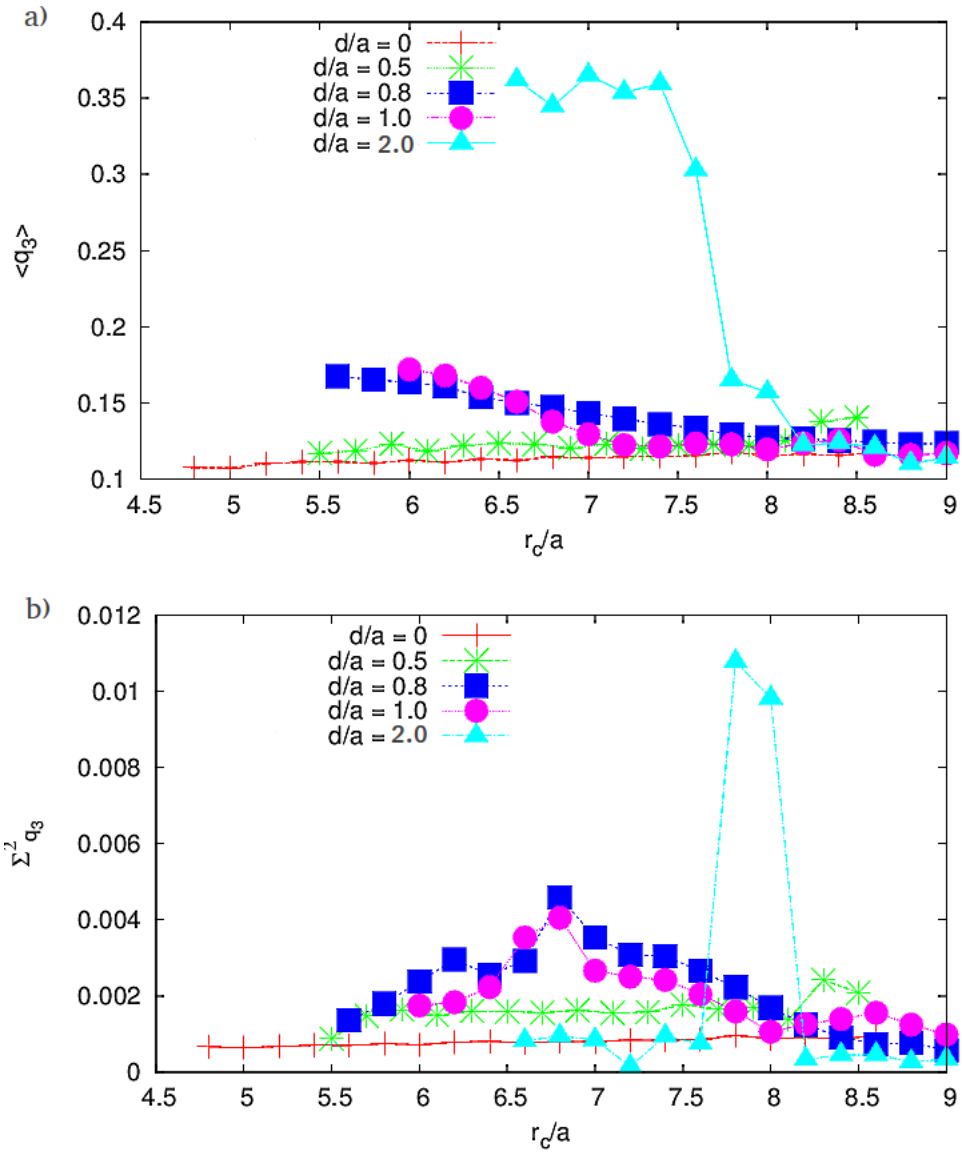

Figure 5.5: a) Average value of the tetrahedral order parameter ($\langle q_3 \rangle$) for chains with various d/a, as labeled in the legends, and $l_p/a = 16$. b) Mean square fluctuation of q_3 ($\Sigma^2_{q_3}$).

The value of tetrahedral order parameter is small but not zero for the chains of $d/a \geqslant 0.8$; therefore, we further investigate defect locations on the confining sphere of radius $r/a = 6.6$. The local nematic order parameter Q_{local} was calculated for all surface bonds, located between the confining surface and the shell where the

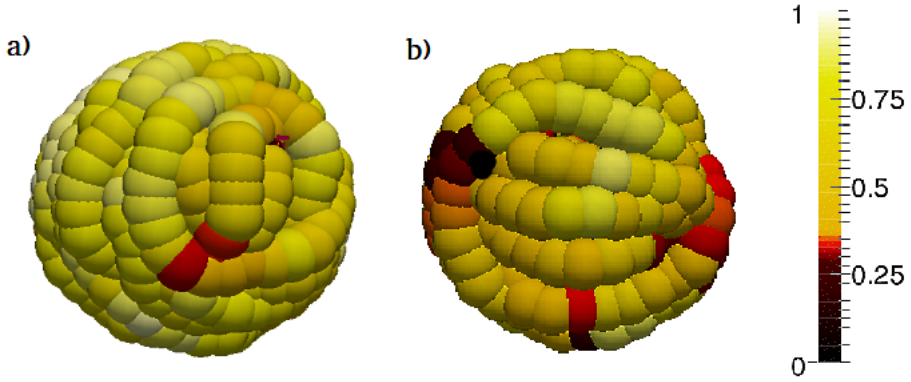

Figure 5.6: Snapshots of a fused-sphere chain with $d/a = 2$ (a)) and the pearl-necklace chain with $d/a = 1$ (b)), confined inside a sphere of $r_c/a = 6.6$. Colors present Q_{local} values.

number of monomers was the highest. Bonds located within the distance $r/a = 1.2$ from any selected surface bond were used to calculate Q_{local}. Using a threshold of $Q_{local} \leqslant 0.32$, we see four defects forming where the strands of polymer sink under the helically wound surface strands at the sharp angles, as seen in Fig. 5.6 a) and b). Snapshots in the Fig. 5.6 show the fused-sphere chain with $d/a = 2$ a) and the pearl-necklace chain with $d/a = 1.0$ b). Red dots in Fig. 5.6 a) and b) show the locations of two $+1/2$ defects. Other two defects are located on the opposite side of the sphere. We can see that the local bond vectors at each defect location form a vector field similar to that shown in Fig. 1.6 b), which has topological charge of $+1/2$. In the fused-sphere chain, the two defects located at the same pole nearly merge to form a single point defect with a charge of $+1$. In the pearl-necklace chain with $d/a = 0.8$ and $d/a = 1$, we do not observe any obvious difference in the defect locations. The $+1/2$ defects in this case are also more separated, compared to the fused-sphere chain. The proximity of $+1/2$ defects in the pairs, explains why the helicoildal order parameter is much higher than tetrahedral order parameter, especially in the fused-sphere chain. While, locally the surface monomers form four $+1/2$ defects, on the larger scale, the structure is very close to a helicoildal arrangement with two $+1$ defects positioned at the poles.

5.4 Conclusions

In this chapter, we have applied the EXEDOS algorithm to the pearl-necklace and the fused-sphere polymer chains with $l_p/a = 16$ confined inside a sphere, to investigate the effects of steric hindrance (the ratio of diameter d to bond length a) on chain spatial organization. The d/a ratio was varied from 0 (ideal) to 2. A pearl-necklace chain of 300 monomers was used to model chains with and without overlap monomers. For fused sphere chain, overlap monomers were added at the middle point of each bond of a pear-necklace chain with $d/a = 1$. By allowing overlap, we provide the chain with larger flexibility, compared to the pearl-necklace chain of the same diameter.

Simulations show that the pearl-necklace chains with $d/a \leqslant 0.5$ do not form ordered state, at any levels of confinement. For the chains with larger d/a (0.8 and 1), however, we observe isotropic-nematic (I-N) transition. The fused-sphere chain, having the highest concentration, compared to the pearl-necklace chains, transitions into nematic state earlier (at larger confining radius). We also find that the chain with $d/a = 0.8$ and 1 transition into nematic state, when confined inside a sphere of about the same size.

Molecular arrangement of confined chains was studied, using helicoidal (P_2) and tetrahedral (q_3) order parameters. Only chains with $d/a \geqslant 0.8$ transition from disordered to imperfect helicoidal arrangement. The defect studies confirm that the structure forms four surface defects, each having a charge of $+1/2$. The defects form at the intersection of the bond vectors, where the polymer strands sink under the helically wound surface. For the fused-sphere chain, two $+1/2$ defects at each pole nearly merge to form a single point defect with a charge of $+1$.

Chapter 6

Conclusions and future work

The work presented above applies the expanded density-of-states (EXEDOS) method to simulate coarse-grained flexible and semi-flexible polymers, confined inside a sphere. The model can be related to DNA packed in a nucleus of eukaryotes [18] or in cytoplasm of prokaryotes [23], and liquid crystalline polymer inside a droplet [15] to name few examples. In the current study, we only considered a specific case, where the sphere contains a single molecule. However, the model can also be expanded to problems related to confinement of several molecules inside a sphere such as proteins in a membrane-bound vesicles inside the cell where physiological environment greatly affects protein geometry and function. In addition to biological systems, this model can also be related to liquid crystals confined inside a micro/nano droplet [11]. Defects in polymer alignment that appear on a surface of such droplets can form a base for several technological and medical applications [11, 32].

Before applying the method to a more sophisticated model of a semi-flexible chain, the EXEDOS method was tested on both ideal and self-avoiding flexible chains of lengths 100 to 600 monomers, confined inside a sphere to test the applicability of the algorithm to a coarse-grained polymer model. Self-avoiding chain was modeled as a pearl-necklace chain of monomers with diameter $d/a = 0.8$ (where a is a bond length), while in an ideal chain monomers were represented by points without any volume. The density of configuration states and confinement free energy were calculated, and the power-law dependence of free energy on parameters such as polymer concentration and confining radius was compared to that obtained in other studies [91, 120]. For ideal-flexible chains, the power-law dependence of

free energy on the confining radius is in a good agreement with earlier theoretical predictions [61, 118]. In self-avoiding flexible chains, finite size effects cause deviation of the power-law dependence from theoretical [120] as well as other simulation studies [91, 92] performed for longer chains. The performance of EXEDOS algorithm was limited when it was applied to self-avoiding chains of sizes larger than 600 monomers, as simulations became excessively long when performed on a single CPU. The radial density distribution of confined chains was also investigated and was found consistent with that obtained in other simulations [28, 120]. At low concentrations, polymer chains avoid the confining surface, to minimize the loss of accessible configurations. At higher concentrations, we observe oscillation in the number density near the surface, which points to a formation of ordered polymer layer adjacent to it. This effect was also observed in other studies, where self-avoiding flexible polymer chains were confined inside spheres of sizes much smaller than their unconfined size [93, 132]. We could not observe chain behavior at very high concentration ($\eta > 0.2$) though, since the efficiency of the EXEDOS algorithm decreases, when it is applied to very dense systems.

Simulation of a spherically confined semi-flexible chain is more complicated than that of a flexible chain, since added stiffness adversely impacts the algorithm convergence, limiting simulations to densities lower than $\rho \leq 0.25$ and maximum chain lengths to 300 monomers. Polymer chain was modeled as a pearl-necklace chain with diameter $d/a = 0.8$ and persistence length l_p/a ranging from 8 to 32. Even though, chemical structure of the polymer was neglected in this model, it still presents important features of semi-flexible polymer. For example, this semi-flexible polymer chain can be mapped to a double stranded (ds) DNA with diameter $d \approx 2$ nanometers (nm), and persistence length $l_p \approx 50\ nm$ or $l_p/a \approx 25$ [151], where a is a bond length and $a = d$. One turn of the DNA strand is composed of approximately 10 base pairs (bp) with length of 3.4 nm [145]. In this case, semi-flexible chain of 300 monomers, each having size of 2 nm, will have a contour length of approximately 600 nm, corresponding to roughly 1765 bp. Persistence length of our simulated chains falls in the range of $8 \leq l_p/a \leq 32$, which includes DNA persistence length ($l_p/a \approx 25$). At the densities allowed by the EXEDOS algorithm ($\rho \leq 0.25$), we were able to observe changes in the free energy, bending energy, and order parameters, that were characteristic to isotropic-nematic (I-N) transition, similar to that observed in other studies [24, 71, 130]. The free energy increases with a decrease in

confining radius since more work is required to pack a chain in a denser environment. The bending energy grows linearly at low densities, but this growth slows down at higher densities and tends to level off, as the polymer forms ordered layer conforming to the surrounding surface to minimize the bending energy cost. In experiments, it was also found that T7 DNA forms a thin DNA layer near a capsid's surface [61]. These effects become even more pronounced as chain persistence length increases *i.e.* more work is required to pack a stiffer chain inside a sphere of the same size, and the stiffer chain forms tighter polymer layer near the surface, compared to a more flexible chain.

The I-N transition was observed through the changes in nematic order parameter (Q). The order parameter follows S-shaped curve, when plotted as a function of a system density, and its transition from low to high value corresponds to a transition from disordered to ordered state (I-N transition). The transition also presents itself as a spike of mean square fluctuation of Q. The I-N transition of more flexible chains occurs at higher densities, compared to stiffer chains. Due to the finite size effects the transition densities of all chains were slightly lower than those observed in other studies [130]. Helicoildal order parameter P_2 was calculated to probe for the presence of helicoildal arrangement in simulated chains. The values of P_2 increase with density and persistence length. We also probed for the presence of a tennis ball configuration by calculating the tetrahedral order parameter (q_3), at different densities. In all considered chains, the value of q_3 is relatively low and increases in S-shaped manner as the polymer density increases. At the same density, the value of q_3 is higher in stiffer chains. Even though, both P_2 and q_3 are non-zero, they are far from those obtained in systems with ideal helicoidal and tetrahedral arrangements. This indicates that structures we observe in our studies to some extent present characteristics of both helicoidal and tetrahedral arrangements. This was confirmed by calculation of local nematic order parameter (Q_{local}) and studying the molecular snapshots of chains. Molecular snapshots show that all chains form an imperfect helicoidal arrangement at densities higher that I-N transition density. By comparison with a structure observed in other simulations [24, 25], and that observed experimentally in a bacteriophage DNA [23] at high density, our study reproduces the correct equilibrium geometry of the DNA chain. Q_{local} values plotted for a selected polymer configuration show that locally we can have a structure with four $+1/2$ defects located where bundles of the polymer strands sink under

the helically wound surface. However, this structure is far from an ideal tennis ball structure. Also, it was observed that the polymer structure at high density was strongly influenced by the initial configuration of simulated chain.

Molecular organization in a confined cavity can be determined not only by its bending stiffness, but also by a chain geometry, resulting from steric interactions or monomer-monomer overlap [153]. To study the influence of chain geometry on its spatial organization in confining sphere, four pearl-necklace chains with 300 monomers were formed, each having different ratio of chain diameter d to a bond length a, ranging from 0 to 1. The fifth chain was formed by adding overlap monomer at the middle point of each bond of the chain with $d/a = 1$, to form a fused-sphere chain with $d/a = 2$. All considered chains were assigned persistence length $l_p/a = 16$. The persistence length and number of monomers were limited to these values so that the simulation could finish in a reasonable time. Nematic order parameter Q was calculated to investigate effect of steric interaction on I-N transition. It was found that the ideal chain and the self-avoiding pearl-necklace chain with $d/a = 0.5$ did not form ordered structure at any considered concentration. However, chains with $d/a \geqslant 0.8$, undergo transition to ordered state. This happens at higher concentration in the chain with smaller d/a since inside a sphere of the same size this chain has more available volume to reorient and translate, and forms a less ordered structure to maximize its entropy. At concentrations above I-N transition, the chains with $d/a = 1$ and 0.8 form an imperfect helicoidal structure. The fused-sphere chain forms a distinct helicoildal structure, similar to that observed in T7 DNA [23], with high value of the helicoildal order parameter (P_2). The presence of a tennis ball structure was also tested with tetrahedral order parameter (q_3). It looks like during the transition into the ordered state the chains with $d/a \geqslant 0.8$ approach tetrahedral configuration, but as the system becomes denser, they assume the helicoildal state. Allowing monomer overlap in the pearl-necklace chain increases chain ability to bend around a confining surface, and such a chain (called the fused-sphere chain) shows higher value of order parameters, compared to a similar chain without overlap. The local nematic order parameter Q_{local} was calculated and plotted for selected polymer configurations confined at concentrations higher that I-N transition concentration. It was found that chains with $d/a \geqslant 0.8$ form a structure with four $+1/2$ defects. In the fused-sphere chain, two $+1/2$ defects at each pole nearly merge to form a single point defect with a charge of $+1$.

Polymer models studied in this work were limited to a pearl-necklace chain [58] composed of spherical beads and fused-sphere [153] chain, which works as a little better approximation of polymer chain-like DNA. Even though, these models are rather simple, they show that local geometry of the polymer chain can have the same effects on spatial conformation of confined molecules, as the presence of explicit bending energy in the chain [153]. In proteins and nucleic acids, this local geometry would typically vary along the strand of the molecule, and often may not be described by a bead-shaped monomers. For instance, double stranded DNA molecule poses internal twisting, and may be better described by a stack of ellipsoidal particles with twisting energy, than by a chain of simple beads [152]. Expanding the EXEDOS algorithm to more detailed models of DNA will likely bring significant improvement into the predictive abilities of the simulation.

Taking into account hydrophobic and hydrophilic interactions can further bring the EXEDOS method to a more practical level, when used to study protein folding [154, 155]. The models considered here can be treated as uniformly hydrophilic chains suspended in implicit water. As a result, spatial organization was determined only by excluded monomer volume, monomer diameter and aspect ratio, and relation to the confining surface. If hydrophobic - hydrophilic interactions are taken into account, the model can be expanded to the vital problem of protein folding and exchange in confined cellular structures [154, 155]. Both proteins and nucleic acids typically consist of hundreds to hundreds of thousands building blocks (amino acids and nucleotides) [18]; therefore, the algorithm employed in this thesis is limited only to the shortest examples. While, this limitation can be partially resolved by grouping series of amino acids (or nucleotides) into a single monomer, the usability of the EXEDOS method would still be limited to at most thousands amino acids/nucleotides per chain, if the algorithm is used in a current form. Further research is therefore needed to speed-up the algorithm convergence. The mobility of densely confined polymer chain is intrinsically limited; therefore, improving the geometry update [156] procedure may be crucial in speeding up the algorithm. The geometry update method may affect not only speed of the convergence, but can also affect accuracy, and the equilibrium geometry of the chain. More research should be done, therefore, to adapt different relaxation algorithms [157] that lead the polymer to a global equilibrium. Simulation methods, such as umbrella sampling [158], incorporating the external force to compress the chain can be also considered to

adequately reflect the real-world processes.

While the work presented above have shown that EXEDOS method is capable of producing meaningful and interesting results for short coarse-grained models of flexible and semi-flexible polymers. The performance is still limited at higher densities, and convergence is excessively slow for chains of realistic lengths. As new features are introduced into the model, new challenges are likely to arise, which will call for further algorithm improvements, and verification against other simulation methods and experiments. Through this conjunction with experiments and other simulation techniques we hope this method will evolve into a practical tool of protein modeling and analysis.

References

[1] S. M. Simon and G. Blobel, "A protein-conducting channel in the endoplasmic reticulum," *Cell*, vol. 65, pp. 371-380, May 1991.

[2] J. Helenius, D. T. Ng, C. L. Marolda, P. Walter, M. A. Valvano, and M. Aebi, "Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein," *Nature*, vol. 415, pp. 447-450, January 2002.

[3] Y. Higuchi, K. Yoshikawa, and T. Iwaki, "Confinement causes opposite effects on the folding transition of a single polymer chain depending on its stiffness," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 84, pp. 021924-5, August 2011.

[4] Y. Taniguchi, M. Kondo, Y. Haruyama, S. Matsui, and N. Kawatsuki, "Influence of film thickness on the reorientation structure of photoalignable liquid crystalline polymer films," *Polymers*, vol. 90, pp. 290-294, May 2016.

[5] H. Zhao, J-J. Wie, D. Copic, R. Oliver, A. O. White, S. Kim, and A. J. Hart, "High-fidelity replica molding of glassy liquid crystalline polymer microstructures," *Acs Applied Materials and Interfaces*, vol. 8, pp. 8110-8117, March 2016.

[6] A. Winkler, A. Stat, P. Virnau, and K. Binder, "Phase transitions and phase equilibria in spherical confinement," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 87, pp. 032307-13, March 2013.

[7] M. Nakata, G. Zanchetta, M. Buscaglia, T. Bellini, and N. A. Clark, "Liquid crystal alignment on a chiral surface: interfacial interaction with sheared DNA films," *Langmuir*, vol. 24, pp. 10390-4, September 2008.

[8] Y. J. Cha, M-J. Gim, K. Oh, and D. K. Yoon, "In-plane switching mode for liquid crystal displays using a DNA alignment layer," *Acs Applied Materials and Interfaces*, vol. 7, pp. 13627-13632, June 2015.

[9] Y. J. Cha, M-J. Gim, K. Oh, and D. K. Yoon, "Twisted-nematic-mode liquid crystal display with a DNA alignment layer," *Journal of Information Display*, vol. 16, pp. 129-135, July 2015.

[10] M. Jaffe, "Applications of Liquid Crystal Polymers," *Journal of Statistical Physics*, vol. 62, pp. 985-995 , March 1991.

[11] G. P. Crawford and S. Zumer, *Liquid Crystals in Complex Geometries*. London: Francis and Taylor, 1996.

[12] J. R. C. van der Maarel, *Introduction to Biopolymer Physics*. Singapore: World Scientific, 2007.

[13] A. R. Guzdar, R. W. Lusignea, K. G. Blizard, and L. S. Rubin, *Commercial Uses of Liquid Crystalline Polymers in Frontiers of Polymer and Advanced Materials*. New York: State University of New York, 1994.

[14] S. H. Ryu and D. K. Yoon, "Liquid crystal phases in confined geometries," *Liquid Crystals*, vol. 43, pp. 1951-1972, December 2016.

[15] A. Fernndez-Nieves, D. R. Link, and D. A. Weitz, "Polarization dependent Bragg diffraction and electro-optic switching of three-dimensional assemblies of nematic liquid crystal droplets," *Applied Physics Letters*, vol. 88, pp. 121911-3, March 2006.

[16] D. R. Nelson, "Toward a Tetravalent Chemistry of Colloids," *Nano Letters*, vol. 2, pp. 1125-1129, October 2002.

[17] Z. Zhang, A. S. Keys, T. Chen, and S. C. Glotzer, "Self-Assembly of patchy particles into diamond structures through molecular mimicry," *Langmuir*, vol. 21, pp. 11547-51, December 2005.

[18] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell,* 4^{th} ed. New York: Garland Science, 2002.

[19] H. Shin and G. M. Grayson, "Filling the void in confined polymer nematics: Phase transitions in a minimal model of dsDNA packing," *Europhysics Letters*, vol. 96, pp. 36007-6, October 2011.

[20] D. Svensek and R. Podgornik, "Confined chiral polymer nematics: Ordering and spontaneous condensation," *Europhysics Letters*, vol. 100, pp. 66005-6, 2012.

[21] C. R. Locker, S. D. Fuller, and S.C. Harvey, "DNA Organization and Thermodynamics during Viral Packing," *Biophysical Journal*, vol. 93, pp. 2861-2869 , January 2007.

[22] U. Qimron, S. Tabor, and C. C. Richardson, "New Details about Bacteriophage T7-Host Interactions," *MICROBE MAGAZINE*, vol. 5, pp. 117-122, 2010.

[23] M. E. Cerritelli, N. Cheng, A. H. Resenberg, C. E. McPherson, F. P. Booy, and A. C. Steven, "Encapsulated conformation of bacteriophage T7 DNA," *Cell*, vol. 91, pp. 271-280, October 1997.

[24] A. J. Spakowitz and Z-G. Wang, "DNA Packing in Bacteriophage: Is Twist Important?," *Biophysical Journal*, vol. 88, pp. 3912-3923, June 2005.

[25] A. Fathizadeh, M. Heidari, B. Eslami-Mossallam, and M. R. Ejtehadi. "Confinement dynamics of a semiflexible chain inside nano-spheres," *The Journal of Chemical Physics*, vol. 139, pp. 044912-8, July 2013.

[26] N. N. Oskolkov, P. Linse, I. I. Potemkin, and A. R. Khokhlov, "Nematic ordering of polymers in confined geometry applied to DNA pack-

aging in viral capsids," *The Journal of Physical Chemistry B*, vol. 115, pp. 422-432, January 2011.

[27] E. Katzav, M. Adda-Bedia, and A. Boudaoud, "A statistical approach to close packing of elastic rods and to DNA packaging in viral capsids," *Proceedings of the National Academy of Sciences of the United States*, vol. 103, pp. 18900-5, December 2006.

[28] P. Crifa and T. Bleha, "Simulation of Chain Organization in Encapsulated Polymer," *Macromolecular Symposia*, vol. 296, pp. 336-341, October 2010.

[29] P. Crifa and T. Bleha, "Shape transition of semi-flexible macromolecules confined in channel and cavity," *The European Physical Journal E*, vol. 32, pp. 273-279, July 2010.

[30] C. R. Locker and S. C. Harvey, "A Model for Viral Genome Packing," *Multiscale Modeling and Simulation*, vol. 5, pp. 1264-1279, December 2006.

[31] P. K. Purohit, M. M. Inamdar, and P. D. Grayson, "Forces during Bacteriophage DNA Packaging and Ejection," *Biophysical Journal*, vol. 88, pp. 851-866, February 2005.

[32] A. Fernandez-Nieves, V. Vitelli, A. S. Utada, D. R. Link, M. Marquez, D. R. Nelson, and D. A. Weitz, "Novel Defect Structures in Nematic Liquid Crystal Shells," *Physical Review Letters*, vol. 99, pp. 157801-4, October 2007.

[33] T. Lopez-Leon, V. Koning, K. B. S. Devaiah, V. Vitelli, and A. Fernandez-Nieves, "Frustrated nematic order in spherical geometries," *Nature Physics*, vol. 7, pp. 391-394, February 2011.

[34] C. R. Wand and M. R. Bates, "Monte Carlo simulations of nematic and chiral nematic shells," *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, vol. 92, pp. 012502-4, January 2015.

[35] A. Nikoubashman, A. D. Vega, K. Binder, and A. Milchev, "Semiflexible Polymers in Spherical Confinement: Bipolar Orientational Order Versus Tennis Ball States," *Physical Review Letters*, vol. 118, pp. 217803-5, May 2017.

[36] A. Reinhardt, J. P. K. Doye, E. G. Noya, and C. Vega, "Local order parameters for use in driving homogeneous ice nucleation with all-atom models of water," *The Journal of Chemical Physics*, vol. 137, pp. 194504-9, November 2012.

[37] C. Balbuena, C. Brito, and D. A. Stario, "Structural signatures of (two) characteristic dynamical temperatures in lithium metasilicate," *Journal of Physics: Condensed Matter*, vol. 26, pp. 155104-8, March 2014.

[38] O. L. Chau, T. R. Forester, and W. Smith, "Curvature effects on hidrophobic solvation," *Molecular Physics*, vol. 89, pp. 1033-1055, November 1996.

[39] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquid*. Oxford: Oxford Science, 1987.

[40] D. P. Landau and K. Binder, *A Guide to Monte Carlo Methods in Statistical Physics,* 3^{rd} ed. Cambridge: Cambridge University Press, 2000.

[41] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics*. Oxford: Oxford University, 1999.

[42] F. Wang and D. P. Landau, "Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States," *Physical Review Letters*, vol. 86, pp. 2050-3, March 2001.

[43] D. P. Landau, S-H. Tsai, and M. Exler, "A new approach to Monte Carlo simulations in statistical physics: Wang-Landau sampling," *American Journal of Physics*, vol. 72, pp. 1294-1302, October 2004.

[44] N. Rathore and J. J. de Pablo, "Monte Carlo simulation of proteins through a random walk in energy space," *The Journal of Chemical Physics*, vol. 116, pp. 7225-7230, April 2002.

[45] P. Poulain, F. Calvo, R. Antoine, M. Broyer, and P. H. Dugourd, "Performances of Wang-Landau algorithms for continuous systems," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 73, pp. 056704-11, May 2006.

[46] Y. Sliozberg and C-F Abrams, "Effects of Confinement on the Thermodynamics of a Collapsing Heteropolymer: An Off-Lattice Wang-Landau Monte Carlo Simulation Study," *Macromolecules*, vol. 38, pp. 5321-5329, June 2005.

[47] D. Jayasri, T. Sairamb, K. P. N. Murthy, and V. S. S. Sastry, "Liquid crystal films on curved surfaces: An entropic sampling study, " *Physica A: Statistical Mechanics and its Applications*, vol. 390, pp. 4549-4554, November 2011.

[48] B. Kim, R. Faller, Q. Yan, N. L. Abbott, and J. J. de Pablo, "Potential of mean force between a spherical particle suspended in a nematic liquid crystal and a substrate," *The Journal of Chemical Physics*, vol. 117, pp. 7781-7787, October 2002.

[49] M. Chopra, M. Muller, and J. J. de Pablo, "Order-parameter-based Monte Carlo simulation of crystallization," *The Journal of Chemical Physics*, vol. 124, pp. 134102-8, April 2006.

[50] J. Zelko, A. Iglic, V. Kralj-Igli, and P. B. S. Kumar, "Effects of counterion size on the attraction between similarly charged surfaces," *The Journal of Chemical Physics*, vol. 133, pp. 204901-8, November 2010.

[51] J. Z. Y. Chen, "Coil-Bridge transition and Monte Carlo simulation of a stretched polymer," *Physical Review E*, vol. 84, pp. 041809-9, October 2011.

[52] M. Rubinstein and R. Colby, *Polymer Physics*. Oxford: Oxford University, 2003.

[53] P. J. Flory, *Statistical mechanics of chain molecules,* Vol. 8. New York: Wiley InterScience, 1969.

[54] M. Rawiso, R. Duplessix, and C. Picot, "Scattering function of polystyrene," *Macromolecules*, vol. 20, pp. 630-648, May 1987.

[55] M. Muthukumar and B. G. Nickel, "Perturbation theory for a polymer chain with excluded volume interaction," *The Journal of Chemical Physics*, vol. 80, pp. 5839-5850, June 1984.

[56] J. Kierfeld, O. Niamploy, V. Sa-yakanit, and R. Lipowsky, "Stretching of semiflexible polymers with elastic bonds," *The European Physical Journal E*, vol. 14, pp. 17-34, May 2004.

[57] C. Hakansson and C. Elvingson, "Semi-flexible Chain Molecules with Nonuniform Curvature. 1. Structural Properties," *Macromolecules*, vol. 27, pp. 3843-3849, July 1994.

[58] R. A. Harris and J. E. Hearst, "On Polymer Dynamics," *The Journal of Chemical Physics* , vol. 44, pp. 2595-2602, April 1966.

[59] J. Wang and H. Gao, "A generalized bead-rod model for Brownian dynamics simulations of wormlike chains under strong confinement," *The Journal of Chemical Physics*, vol. 123, pp. 084906-13, August 2005.

[60] P. Pincus, "Excluded volume effects and stretched polymer chains," *Macromolecules*, vol. 9, pp. 386-388, May 1976.

[61] E. F. Casassa, "Equilibrium distribution of flexible polymer chains between a macroscopic solution phase and small voids," *Journal of Polymer Science Part C: Polymer Letters*, vol. 5, pp. 773-778, September 1967.

[62] M. Daoud and P. G. De Gennes, "Statistics of macromolecular solutions trapped in small pores," *LE JOURNAL DE PHYSIQUE*, vol. 38, pp. 85-93, January 1977.

[63] A. Y. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules*. New York: American Institute of Physics, 1979.

[64] T. Sakaue and E. Raphael, "Polymer Chains in Confined Spaces and Flow-Injection Problems:Some Remarks," *Macromolecules*, vol. 39, pp. 2621-2628, April 2006.

[65] P. G. de Genne, *Scaling Concepts in Polymer Physics*. London: Cornell University, 1979.

[66] T. Sakaue, "Semiflexible Polymer Confined in closed Space," *Macromolecules*, vol. 40, pp. 5206-5211, July 2007.

[67] E. B. Priestly, P. J. Wojtowicz, and P. Sheng, *Introduction to Liquid Crystals*. New York: Plenum, 1946.

[68] P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals*. Oxford: Oxford University, 1993.

[69] S. Chandrasekhar, *Liquid Crystals,* 2^{nd} ed. Cambridge: Cambridge University, 1992.

[70] L. D. Landau and E. M. Lifschitz, Eds., *Statistical Physics: Volume 5 (Course of Theoretical Physics),* 3^{rd} ed. Massachusetts: ELSEVIER, 2008.

[71] Y. Trukhina and T. Schilling, "Computer simulation study of a liquid crystal confined to a spherical cavity," *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, vol. 77. pp. 011701-7, January 2008.

[72] T. S. Yankova, A. Y. Bobrovsky, and A. K. Vorobiev, "Order Parameters P_2, P_4, and P_6 of Aligned Nematic Liquid Crystalline Polymer As Determined by Numerical Simulation of Electron Paramagnetic Resonance Spectra," *The Journal of Physical Chemistry B*, vol. 116, pp. 6010-6016, May 2012.

[73] L. G. P. Dalmolen, S. J. Picken, A. J. De Jong, and W. H. De Jeu, "The order parameters $\langle P_2 \rangle$ and $\langle P_4 \rangle$ in nematic p-alkyl-p-cyanobiphenyls: polarized Raman measurements and the influence of molecular association," *Journal de Physique*, vol. 46, pp. 1443-1449, January 1985.

[74] D. Frenkel and R. Eppenga, "Monte Carlo Study of the Isotropic-Nematic Transition in a Fluid of Thin Hard Disks," *Physical Review Letters*, vol. 49, pp. 1089-1092, October 1982.

[75] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, "Bond-orientational order in liquids and glasses," *Physical Review B*, vol. 28, pp. 784-805, July 1983.

[76] A. Young, *Phase Diagram of The Elements*. Berkeley: University of California, 1991.

[77] I. W. Stewart, *The Static and Dynamic Continuum Theory of Liquid Crystals: A Mathematical Introduction*. New York: Taylor and Francis, 2004.

[78] R. D. Kamien, "The geometry of soft materials: a primer," *Reviews of Modern Physics*, vol. 74, pp. 953-971, October 2002.

[79] J. Poincare, *Journal de Mathmatiques Pures et Appliques*, vol. 4, pp. 167-244, 1885.

[80] H. Hopf, "Vectorfelder in n-dimensionalen Mannigfaltigkeiten," *Mathematische Annalen*, vol. 96, pp. 225-250, 1926.

[81] S. Riniker, J. R. Allison, and W. F. van Gunsteren, "On developing coarse-grained models for biomolecular simulation: a review," *Physical Chemistry Chemical Physics*, vol. 14, pp. 12423-12430, March 2012.

[82] R. H. Swendsen and J-S. Wang, "Nonuniversal critical dynamics in Monte Carlo simulations," *Physical Review Letters*, vol. 58, pp. 86-88, January 1987.

[83] D. F. Parsons and D. R. M. Williams, "An off-lattice Wang-Landau study of the coil-globule and melting transitions of a flexible homopolymer," *The Journal of Chemical Physics*, vol. 124, pp. 221103-4, June 2006.

[84] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, "New approach to Monte Carlo calculation of the free energy: Method of expanded ensembles," *The Journal of Chemical Physics*, vol. 96, pp. 1776-1783, February 1992.

[85] T. C. Lubensky and J. Prost, "Orientational order and vesicle shape," *Journal de Physique II*, vol. 2, pp. 371-382, January 1992.

[86] C. Ohm, M. Brehmer, and R. Zentel, "Applications of Liquid Crystalline Elastomers," *Advances in Polymer Science*, vol. 250, pp. 49-93, March 2012.

[87] M. Alcoutlabi and G. McKenna1, "Effects of confinement on material behavior at the nanometer size scale," *Journal of Physics: Condensed Matter*, vol. 17, pp. R461-R524, March 2005.

[88] S. L. Cravens, J. D. Schonhoft, M. M. Rowland, A. A. Rodriguez, B. G. Anderson, and J. T. Stivers, "Molecular crowding enhances facilitated diffusion of two human DNA glycosylases," *Nucleic Acids Research*, vol. 43, pp. 4087-4097, April 2015.

[89] N. Rathore, T. A. Knotts, and J. J. de Pablo, "Confinement Effects on the Thermodynamics of Protein Folding: Monte Carlo Simulations," *Biophysical Journal*, vol. 90, pp. 1767-1773, March 2006.

[90] J. Kunkel and P. Asuri, "Function, Structure, and Stability of Enzymes Confined in Agarose Gels," *Plos one*, vol. 9, pp. 86785-6, January 2014.

[91] A. Cacciuto and E. Luijten, "Self-avoiding flexible polymers under spherical confinement," *Nano Letters*, vol. 6, pp. 901-905, May 2006.

[92] P. Cifra and T. Bleha, "Free Energy of Polymer Confined in Open and Closed Cavities," *Macromolecular Theory and Simulations*, vol. 21, pp. 15-23, January 2012.

[93] S. Jun, A. Arnold, and B-Y. Ha, "Confined space and effective interactions of multiple self-avoiding chains," *Physical Review Letters*, vol. 98, pp. 128303-4, March 2007.

[94] Y. Jin and H. A. Makse, "A first-order phase transition defines the random close packing of hard spheres," *Physica A: Statistical Mechanics and its Applications*, vol. 389, pp. 5362-5379, December 2010.

[95] W-Y. Zhang and Z. Y. Chen, "Tennis-ball state of a self-avoiding worm-like polymer on a spherical surface," *Europhysics Letters*, vol. 94, pp. 43001-6, May 2011.

[96] P. R. Cook and D. Marenduzzo, "Entropic organization of interphase chromosomes," *The Journal of Cell Biology*, vol. 186, pp. 825-834, September 2009.

[97] T. Takizawa, K. J. Meaburn, and T. Misteli, "The meaning of gene positioning," *Cell*, vol. 135, pp. 9-13, October 2008.

[98] B. A. Berg and T. Neuhaus, "Multicanonical ensemble: A new approach to simulate first-order phase transitions," *Physical Review Letters*, vol. 68, pp. 9-12, January 1992.

[99] B. A. Berg, U. H. E. Hansmann, and Y. Okamoto, "Comment on Monte Carlo Simulation of a First-Order Transition for Protein Folding," *The Journal of Physical Chemistry*, vol. 99, pp. 2236-2237, February 1995.

[100] B. A. Berg and W. Janke, "Multioverlap Simulations of the 3D Edwards-Anderson Ising Spin Glass," *Physical Review Letters*, vol. 80, pp. 4771-4774, May 1998.

[101] W. Janke, "Multicanonical Monte Carlo Simulations," *Physica A: Statistical Mechanics and its Applications*, vol. 254, pp. 164-178, May 1998.

[102] J. Lee, "New Monte Carlo algorithm: Entropic sampling," *Physical Review Letters*, vol. 71, pp. 211-214, July 1993.

[103] A. Mitsutake, Y. Sugita, and Y. Okamoto, "Generalized-ensemble algorithms for molecular simulations of biopolymers," *Biopolymers*, vol. 60, pp. 96-123, January 2001.

[104] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, "The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method," *Journal of Computational Chemistry*, vol. 13, pp. 1011-1021, October 1992.

[105] R. G. Ghulghazaryan, S. Hayryan, and C-K. Hu, "Efficient combination of Wang-Landau and transition matrix Monte Carlo methods for protein simulations," *Journal of Computational Chemistry*, vol. 28, pp. 715-726, January 2007.

[106] Q. Yan, R. Faller, and J. J. de Pablo, "Density-of-states Monte Carlo method for simulation of fluids," *The Journal of Chemical Physics*, vol. 116, pp. 8745-8749, May 2002.

[107] M. S. Shell, P. G. Debendetti, and A. Z. Panagiotopoulos, "Molecular structural order and anomalies in liquid silica," *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, vol. 66, pp. 011202-8, July 2002.

[108] P. N. Vorontsov-Velyaminov, N. A. Volkov, and A. A. Yurchenko, "Entropic sampling of simple polymer models within Wang-Landau algorithm," *Journal of Physics A: Mathematical and General*, vol. 37, pp. 1573-1588, January 2004.

[109] K. A. Maerzke, L. Gai, P. T. Cumming, and C. McCabe, "Incorporating configurational-bias Monte Carlo into the Wang-Landau algorithm for continuous molecular systems," *The Journal of Chemical Physics*, vol. 137, pp. 204105-9, October 2012.

[110] N. Rathore, T. A. Knotts, and J. J. de Plblo, "Configurational Temperature Density of States Simulations of Proteins," *Biophysical Journal*, vol. 118, pp. 3963-3968, December 2003.

[111] L. Janosi and M. Doxastakis, "Accelerating flat-histogram methods for potential of mean force calculations," *The Journal of Chemical Physics*, vol. 131, pp. 054105-12, August 2009.

[112] H. Fynewever and A. Yethiraj, "Phase behavior of semiflexible tangent hard sphere chains," *The Journal of Chemical Physics*, vol. 108, pp. 1636-1644, January 1998.

[113] O. L. Chau and A. J. Hardwick, "A new order parameter for tetrahedral configurations," *Molecular Physics*, vol. 93, pp. 511-518, February 1998.

[114] E. Torino, R. Aruta, T. Sibillano, C. Giannini, and P. A. Netti, "Synthesis of Semicrystalline Nanocapsular Structures Obtained by Thermally Induced Phase Separation in Nanoconfinement," *Scientific Reports*, vol. 6, pp. 32727-12, September 2016.

[115] S. K. Kundu, R. J. Sharma, S-S. Lee, G. Sharma, G. P. Doss, S. Yagihara, D-Y. Kim, J-S. Nam, and C. Chakraborty, "Recent Trends of Polymer Mediated Liposomal Gene Delivery System," *BioMed Research International*, vol. 2014, pp. 1-25, August 2014.

[116] B. H. Zimm and S. D. Levene, "Problems and prospects in the theory of gelelectrophoresis of DNA," *Quarterly Reviews of Biophysics*, vol. 25, pp. 171-204, May 1992.

[117] P. Cifra, P. Linse, and E. Nies, "Energy-driven Asymmetric Partitioning of a Semiflexible Polymer between Interconnected Cavities," *The Journal of Physical Chemistry B*, vol. 112, pp. 8923-7, July 2008.

[118] J. Gao, P. Tang, Y. Yang, and J. Z. Y. Chen, "Free energy of a long semiflexible polymer confined in a spherical cavity," *Soft Matter*, vol. 10, pp. 4674-4685, April 2014.

[119] P. Cifra, "Partitioning of Macromolecules between Two Interconnected Spherical Cavities," *Macromolecules*, vol. 38, pp. 3984-3989, May 2005.

[120] Z. Jin, S. Zhao, and J. Wu, "Entropic forces of single-chain confinement in spherical cavities," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 82, pp. 04185-7, October 2010.

[121] J. Pelletier, K. Halvorsen, B-Y. Ha, R. Paparcone, S. J. Sandler, C. L. Woldringh, W. P. Wong, and S. Jun, "Physical manipulation of the Escherichia coli chromosome reveals its soft nature," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 14, pp. E2649-56, October 2012.

[122] J. Kim, C. Jeon, H. Jeong, Y. Jung, and B-Y. Ha, "Elasticity of flexible polymers under cylindrical confinement: appreciating the blob scaling regime in computer simulations," *Soft Matter*, vol. 9, pp. 6142-6150, May 2013.

[123] D. Marenduzzo and C. Micheletti, "Thermodynamics of DNA Packaging Inside a Viral Capsid: The Role of DNA Intrinsic Thickness," *Journal of Molecular Biology*, vol. 330, pp. 485-492, July 2003.

[124] S. M. Bezrukov, I. Vodyanoy, R. A. Brutyan, and J. J. Kasianowicz, " Dynamics and Free Energy of Polymers Partitioning into a Nanoscale Pore," *Macromolecules*, vol. 29, pp. 8517-8522, January 1996.

[125] D. T. Seaton, "Wang-Landau simulations of thermodynamic behavior in homopolymer systems," Ph. D. dissertation, University of Georgia, Department of Physics and Astronomy, Athens, Georgia, 2010.

[126] Q. Liang, S. Ye, P. Zhang, and J. Z. Y. Chen, "Rigid linear particle confined on a surface: Phase diagram of nematic defect states," *The Journal of Chemical Physics*, vol. 141, pp. 244901-10, December 2014.

[127] D. G. Angelescu, P. Linse, T. T. Nguyen, and R. F. Bruinsma, "Structural transitions of encapsidated polyelectrolytes," *The European Physical Journal E*, vol. 25, pp. 323-334, March 2008.

[128] J. Arsuaga, R. Tan, M. Vazquez, D. W. Sumners, and S. C. Harvey, "Investigation of DNA packing using molecular mechanics model," *Biophysical Chemistry*, vol. 101, pp. 475-484, December 2002.

[129] A. S. Petrov and S. C. Harvey, "Structural and thermodynamic principles of viral packaging," *Structure*, vol. 15, pp. 21-27, January 2007.

[130] A. S. Petrov and S. C. Harvey, "Packaging Double-Helical DNA into Viral Capsids: Structures, Forces, and Energetics," *Biophysical Journal*, vol. 95, pp. 497-502, July 2008.

[131] C. Forrey and M. Muthukumar, "Lagevine Dynamics Simulation of genome Packing in Bacteriophage," *Biophysical Journal*, vol. 91, pp. 25-41, April 2006.

[132] A. Azari and K. K. Muller-Nedebock, "Entropic competition in polymeric systems under geometrical confinement," *Europhysics letters*, vol. 110, pp. 68004-6, June 2015.

[133] I. M. Solovei, S. Kreysing, C. Lanctt, S. Ksem, L. Peichl, T. Cremer, J. Guck, and B. Joffe, "Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution," *Cell*, vol. 137, pp. 356-368, April 2009.

[134] D. A. Luzhbin and Y-L. Chen, "Shifting the Isotropic-Nematic Transition in Very Strongly Confined Semi-flexible Polymer Solutions," *Macromolecules*, vol. 49, pp. 6139-6147, August 2016.

[135] S. Lin, N. Numasawa, T. Nose, and J. Lin, "Coarse-Grained Molecular Dynamic Simulations for Lyotropic Liquid-Crystalline Solutions of Semi-flexible Rod-Like Molecules," *Molecular Crystals and Liquid Crystals*, vol. 466, pp. 53-76, May 2007.

[136] A. S. Petrov, R. Locker, and S. C. Harvey, "Characterization of DNA conformation inside bacterial viruses," *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, vol. 80, pp. 021914-9, August 2009.

[137] D. V. Kuznetsov, "Effect of threebody interactions on the nematic ordering of a polymer solution," *The Journal of Chemical Physics*, vol. 103, pp. 7618-7625, November 1995.

[138] H. Shin, M. J. Bowick, and X. Xing, "Topological Defects in Spherical Nematics," *Physical Review Letters*, vol. 101, pp. 037082-4, July 2008.

[139] C. Vega and P. A. Monson, "Plastic crystal phases of hard dumbbells and hard spherocylinders," *The Journal of Chemical Physics*, vol. 107, pp. 2696-2697, August 1997.

[140] A. R. Bausch, M. J. Bowick, A. Cacciuto, A. Dinsmore, M. F. Hsu, D. R. Nelson, M. G. Nikolaides, A. Travesset, and D. A. Weitz, "Grain boundary scars and spherical crystallography," *Science*, vol. 299, pp. 1716-1718, March 2003.

[141] F. Smallenburg and H. Lwen, "Close packing of rods on spherical surfaces," *The Journal of Chemical Physics*, vol. 144, pp. 164903-10, April 2016.

[142] F. U. Gast and P. J. Hagerman, "Electrophoretic and hydrodynamic properties of duplex ribonucleic acid molecules transcribed in vitro: evidence that A-tracts do not generate curvature in RNA," *Biochemistry*, vol. 30, pp. 4268-4277, April 1991.

[143] P. Kebbekus, D. E. Draper, and P. Hagerman, "Persistence length of RNA (ribonucleic acid)," *Biochemistry*, vol. 34, pp. 4354-4358, April 1995.

[144] D. E. Metzler and C. M. Metzler, *Biochemistry: the chemical reactions of living cells,* 2^{nd} ed. New York: Academic Press, 2003.

[145] J. D. Watsan and F. H. C. Chrick, "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid," *Nature*, vol. 171, pp. 737-738, April 1953.

[146] H-P. Hsu, W. Paul, and K. Binder, "Standard Definitions of Persistence length do not Describe the local intrinsic stiffness of Real

Polymer Chains," *Macromolecules*, vol. 43, pp. 3094-3102, February 2010.

[147] H-P. Hsu, W. Paul, and K. Binder, "Polymer Chain stiffness vs. excluded volume: A Monte Carlo study of the crossover towards the worm-like chain model," *Europhysics Letters*, vol. 92, pp. 28003-6, November 2010.

[148] T. Boublik, C. Vega, and M. Diaz-Pena, "Equation of state of chain molecules," *The Journal of Chemical Physics*, vol. 93, pp. 730-736, July 1990.

[149] K. M. Jaffer, S. B. Opps, D. E. Sullivan, and B. G. Nickel, "The nematic-isotropic phase transition in semi-flexible fused hard-sphere chain fluid," *The Journal of Chemical Physics*, vol. 114, pp. 3314-3324, February 2001.

[150] H. Cheng and P. Zhang, "A tensor model for liquid crystals on a spherical surface," *Science China Mathematics*, vol. 56, pp. 2549-2559, December 2013.

[151] J. Kindt, S. Tzlil, A. Ben-Shaul, and M. Gelbart, "DNA packaging and ejection forces in bacteriophage," *Proceedings of the National Academy of Sciences of the United States*, vol. 98, pp. 13671-4, November 2001.

[152] A. Morriss-Andrews, J. Rottler, and S. S. Plotkin, "A systematically coarse-grained model for DNA and its predictions for persistence length, stacking, twist, and chirality," *The Journal of Chemical Physics*, vol. 132, pp. 035105-3, January 2010.

[153] B. Werlich, M. P. Taylor, T. Shakirov, and Wol. Paul, "On the Pseudo Phase Diagram of Single Semi-Flexible Polymer Chains: A Flat-Histogram Monte Carlo Study," *Polymers*, vol. 9, pp. 1-13, January 2017.

[154] S. R. Durell and A. Ben-Naim, "Hydrophobic-hydrophilic forces in protein folding," *Biopolymers*, vol. 107, pp. 1-14, March 2017.

[155] R. R. Joshi, "Protein folding: interplay of hydrophobic-hydrophilic forces?," *Journal of Biomolecular Structure and Dynamics*, vol. 31, pp. 995-996, January 2013.

[156] A. D. Sokal, "Monte Carlo Methods for the Self-Avoiding Walk," in *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*, K. Binder, Eds. New York: Oxford University Press, 1995, pp. 47-124.

[157] D. Curco, M. Laso, and C. Alema, "Generation-Relaxation Algorithms to Construct Representative Atomistic Models of Amorphous Polymers: Influence of the Relaxation Method," *The Journal of Physical Chemistry B*, vol. 108, pp. 20331-20339, October 2004.

[158] M. Mills and I. Andricioaei, "An experimentally guided umbrella sampling protocol for biomolecules," *The Journal of Chemical Physics*, vol. 129, pp. 114101-13, September 2008.

APPENDICES

Appendix A

Definition of spherical harmonics

Laplace spherical harmonics are used in this work to calculate order parameters of the polymer chain. The mathematical definition of spherical harmonics is given here together with the expression for $Y_{lm}(\theta, \phi)$ at the orders that were of interest in this thesis. Consider a right-handed coordinate system shown in Fig. A.1. A point (x, y, z) in Cartesian coordinate can be expressed in the spherical coordinates ((r, θ, ϕ), where $r \in [0, \infty)$, $\theta \in [0, \pi]$, $\phi \in [0, 2\pi)$) as:

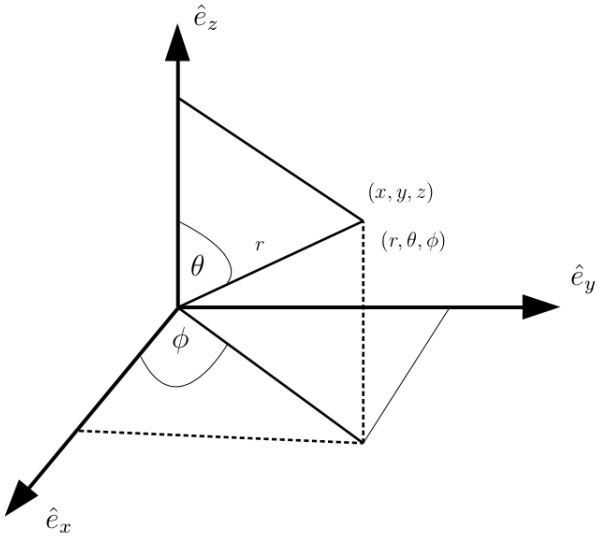

Figure A.1: Polar Coordinates (r, θ, ϕ)

$$x = rsin\theta\ cos\phi, y = rsin\theta\ sin\phi, z = rcos\theta. \tag{A.1}$$

Spherical harmonics $Y_{lm}(\theta, \phi)$, for $l = 0, 1, 2, 3, ..$ and $m \geqslant 0$ are defined as:

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_{lm}(cos\theta) e^{im\phi}, \tag{A.2}$$

where $i \equiv \sqrt{-1}$, $P_{lm}(cos\theta)$ is the Legendre function defined as:

$$P_{lm}(cos\theta) = (-1)^m (1 - cos^2\theta)^{m/2} (\frac{d}{dcos\theta})^m P_l(cos\theta). \tag{A.3}$$

Here $P_l(cos\theta)$ is the Legendre polynomial expressed as:

$$P_l(cos\theta) = \frac{1}{2^l l!} (\frac{d}{dcos\theta})^l (cos^2\theta - 1). \tag{A.4}$$

For $m < 0$, one can use the relation:

$$Y_{l-m}(\theta, \phi) = (-1)^m Y_{lm}(\theta, \phi)^*. \tag{A.5}$$

For $l = 3$ and $m = (0, \pm 1, \pm 2, \pm 3)$ the expressions become:

$$Y_{30}(\theta, \phi) = \frac{1}{4}\sqrt{\frac{7}{\pi}} (5cos^3\theta - 3cos\theta) \tag{A.6}$$

$$Y_{3\pm1}(\theta, \phi) = \frac{\mp 1}{8}\sqrt{\frac{21}{\pi}} \cdot sin\theta (5cos^2\theta - 1) \cdot e^{\pm i\phi} \tag{A.7}$$

$$Y_{3\pm2}(\theta, \phi) = \frac{1}{4}\sqrt{\frac{105}{2\pi}} \cdot sin^2\theta \cdot cos\theta \cdot e^{\pm 2i\phi} \tag{A.8}$$

$$Y_{3\pm3}(\theta, \phi) = \frac{\mp 1}{8}\sqrt{\frac{35}{\pi}} \cdot sin^3\theta \cdot e^{\pm 3i\phi} \tag{A.9}$$

Appendix B

Selected results from Chapter 3

In Chapter 3, the distributions and energy plots are combined for all chains to allow comparison of behavior between chains. Below, the results are separated for each ideal and non-ideal flexible chain showing their distribution of furthermost monomer relative to center of the sphere, and confinement free energy changes. Distribution of the further most monomer position of an ideal and non-ideal flexible chain with lengths 100 - 600 monomers as a function of a confining radius are shown in Fig. B.1 a) - B.6 a) and Fig. B.7 a) - B.12 a), respectively. Changes in the confinement free energy (ΔF) of a flexible chain of lengths 100 - 600 monomers are plotted as a function of a confining radius (r_c/a) and shown in Fig. B.1 b) - B.6 b) for ideal chains and Fig. B.7 b) - B.12 b) for non-ideal chains. Here $\Delta F = F(r_c/a) - F(\infty)$, where $F(\infty)$ is free energy of the chain confined inside a sphere of size $r \approx 4 \cdot r_G/a$, and $r_G/a = N^{\nu}$ is an unconfined size of an ideal chain, a is a bond length, $\nu = 0.5$, and $\nu = 0.588$ for non-ideal chain. These free energy changes (ΔF) are further plotted on logarithmic scale as a function of the ratio of a chain unconfined size to a confining radius (r_G/r_c) as shown in Fig. B.1 c) - B.6 c) for ideal chains, and Fig. B.7 d) - B.12 d) for non-ideal chains. The power law approximation of ΔF dependence on r_G/r_c is shown in dashed line in each plot. For non-ideal chains, changes in specific confinement free energy ($\Delta F/(N+1)$) are also plotted as a function of volume fraction η, where $\eta = (N+1) * d^3/(2r_c)^3$, d is monomer diameter ($d/a = 0.8$), and shown in Fig. B.7 c) - B.12 c).

B.1 Ideal flexible polymer chains

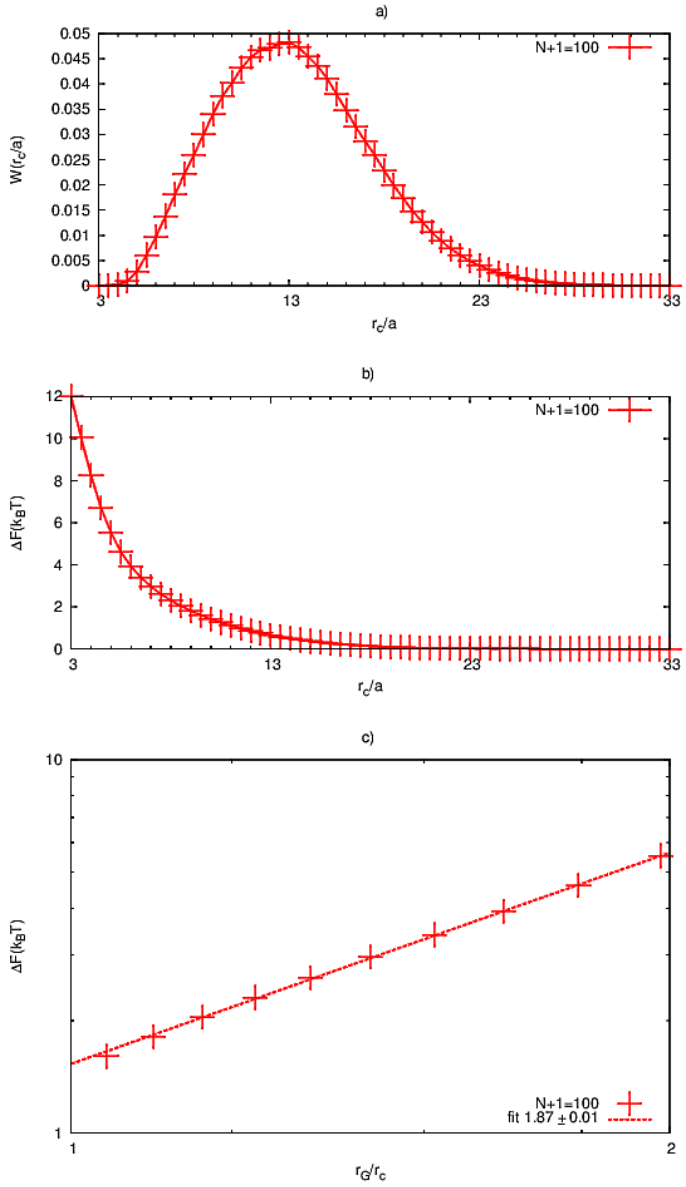

Figure B.1: a) Distribution of the furthermost monomer position ($W(r_c/a)$) for an ideal flexible chain of length 100 monomers, b) confinement free energy changes (ΔF) as a function of a confining radius (r_c/a), and c) confinement free energy change as a function of the ratio of the chain unconfined size to a confining radius (r_G/r_c). Dashed line gives slope of 2.54 ± 0.12.

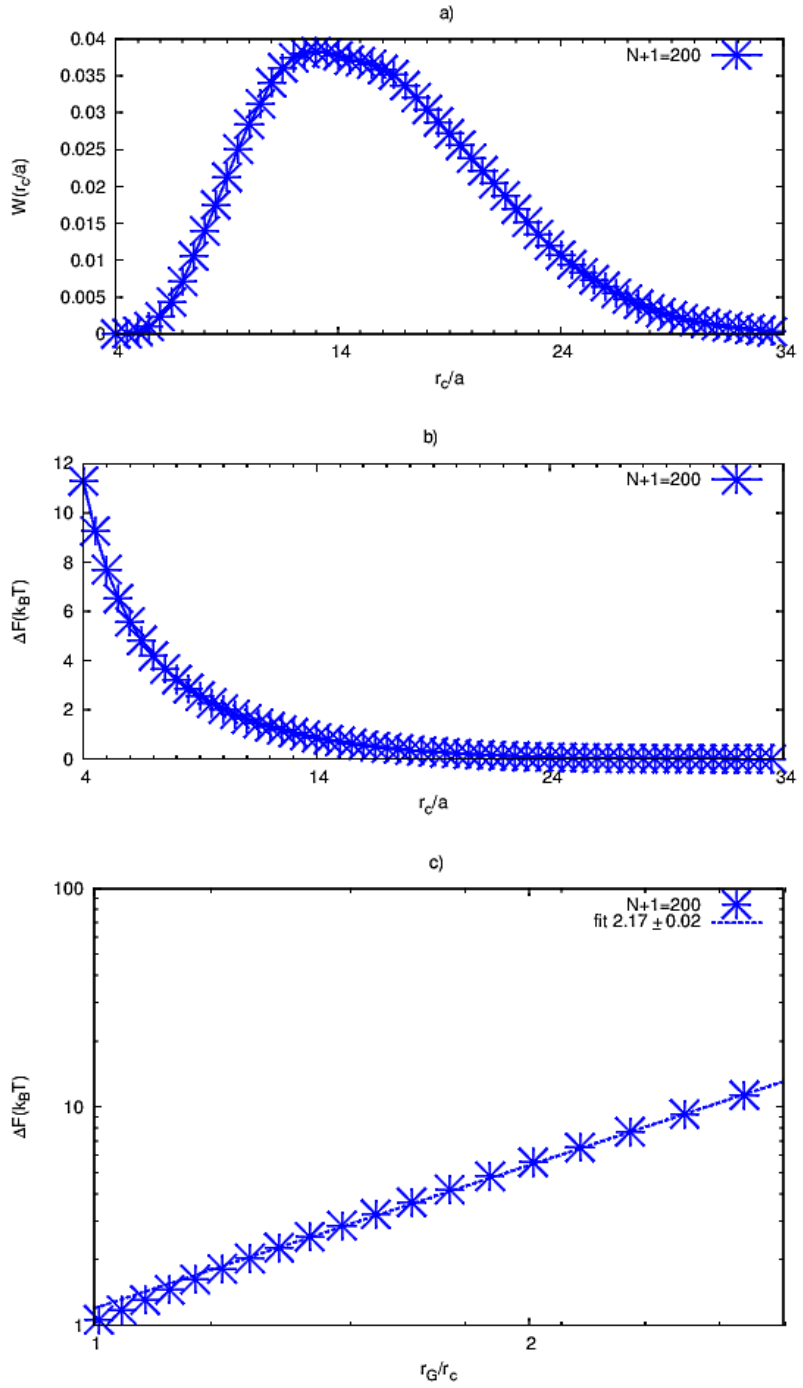

Figure B.2: a) Distribution of the furthermost monomer position ($W(r_c/a)$) for an ideal flexible chain of length 200 monomers, b) free energy change (ΔF) as a function of a confining radius (r_c/a), and c) confinement free energy change as a function of the ratio of the chain unconfined size to a confining radius (r_G/r_c). Dashed line gives slope of 2.17 ± 0.02.

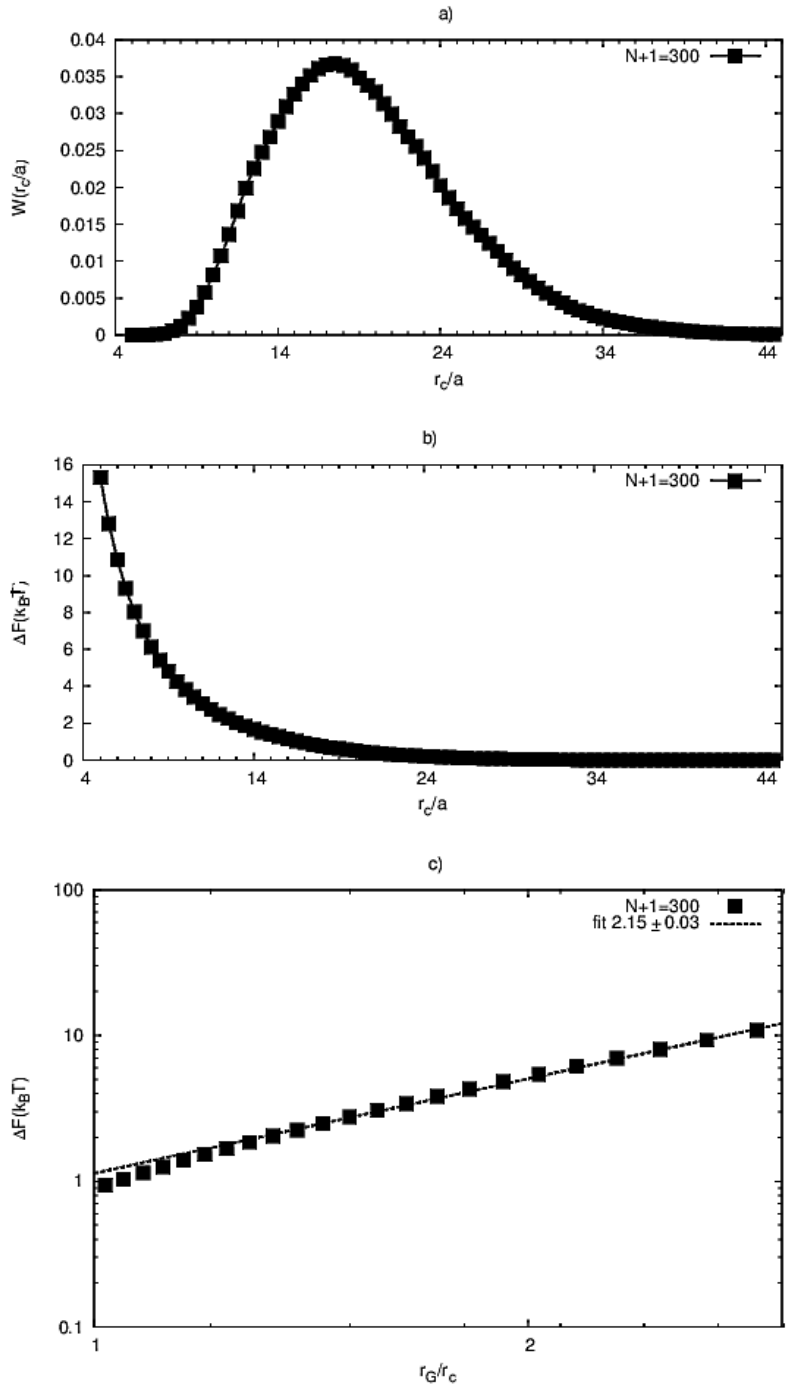

Figure B.3: a) Distribution of the furthermost monomer position ($W(r_c/a)$) for an ideal flexible chain of length 300 monomers, b) free energy change (ΔF) as a function of a confining radius (r_c/a), and c) confinement free energy change as a function of the ratio of the chain unconfined size to a confining radius (r_G/r_c). Dashed line gives slope of 2.15 ± 0.03.

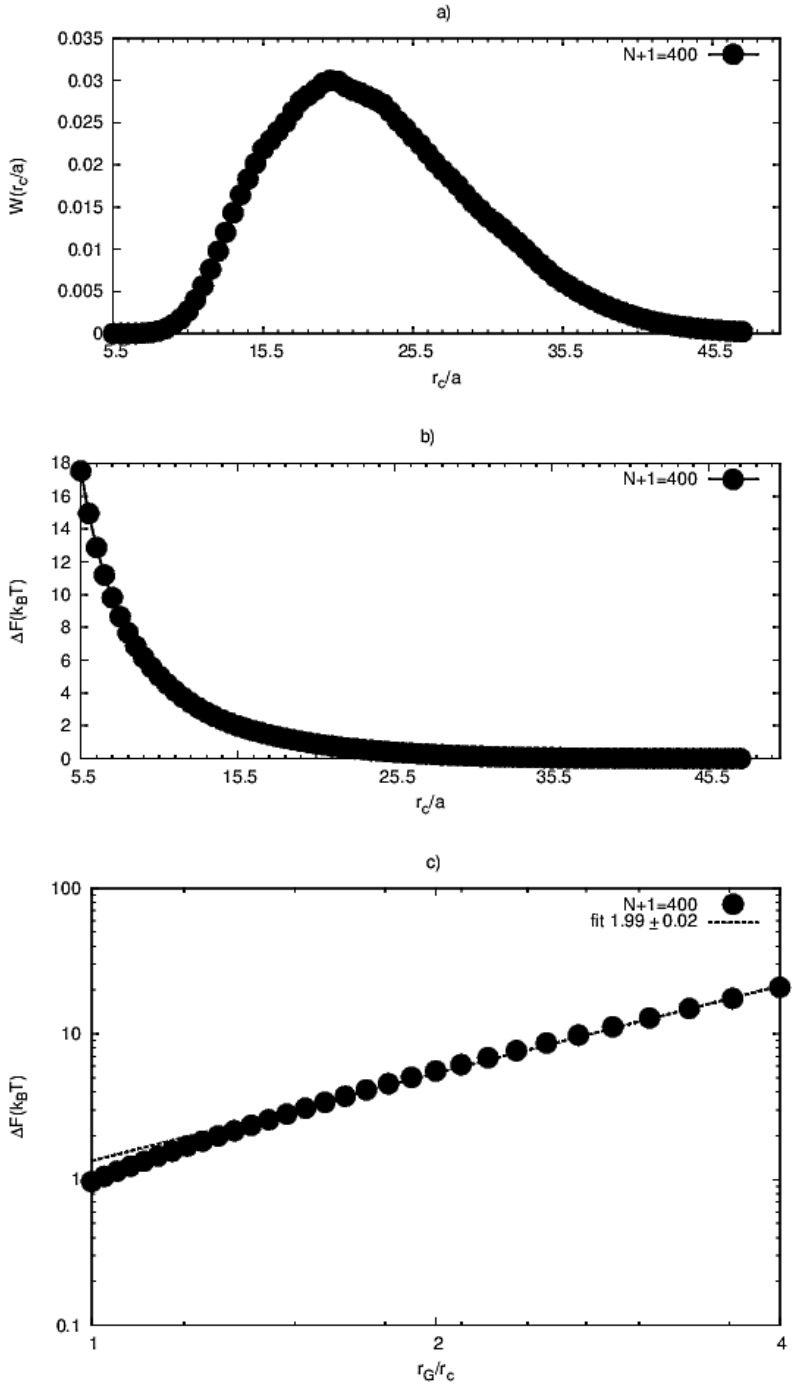

Figure B.4: a) Distribution of the furthermost monomer position ($W(r_c/a)$) for an ideal flexible chain of length 400 monomers, b) free energy change (ΔF) as a function of a confining radius (r_c/a), and c) confinement free energy change as a function of the ratio of the chain unconfined size to a confining radius (r_G/r_c). Dashed line gives slope of 1.99 ± 0.02.

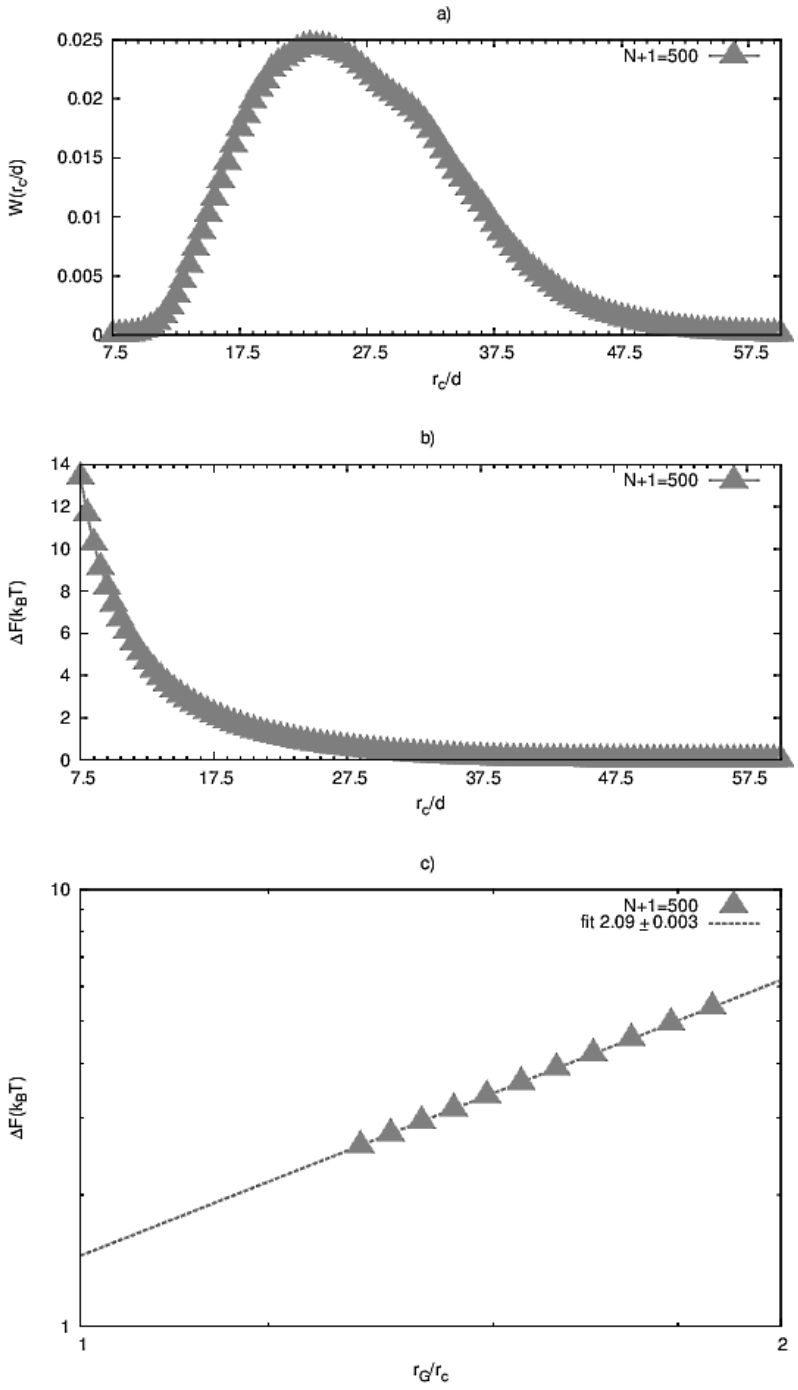

Figure B.5: a) Distribution of the furthermost monomer position ($W(r_c/a)$) for an ideal flexible chain of length 500 monomers, b) confinement free energy change (ΔF) as a function of a confining radius (r_c/a), and c) confinement free energy change as a function of the ratio of the chain unconfined size to a confining radius (r_G/r_c). Dashed line gives slope of 2.09 ± 0.003.

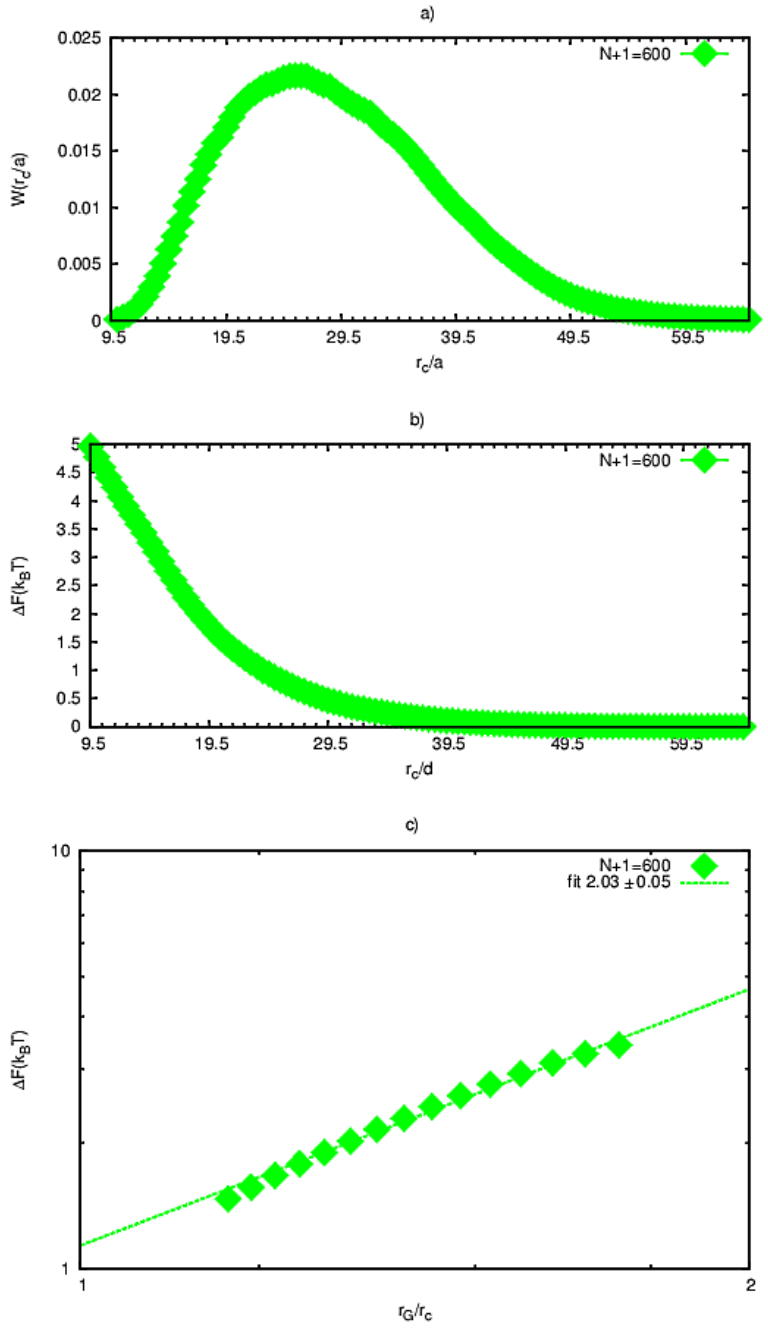

Figure B.6: a) Distribution of the furthermost monomer position ($W(r_c/a)$) of an ideal flexible chain of length 600 monomers, b) confinement free energy change (ΔF) as a function of a confining radius (r_c/a), and c) confinement free energy change as a function of the ratio of the chain unconfined size to a confining radius (r_G/r_c). Dashed line gives slope of 2.03 ± 0.05.

B.2 Self-avoiding flexible polymer chains

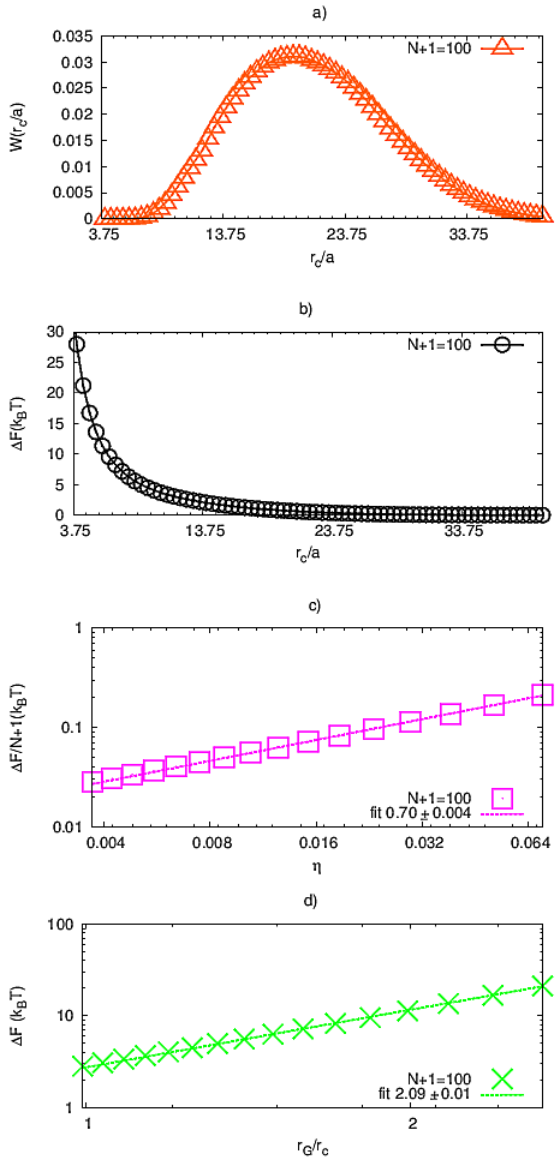

Figure B.7: a) Distribution of the furthermost monomer position ($W(r_c/a)$) for a non-ideal flexible chain of length 100 monomers, b) free energy change (ΔF) as a function of a confining radius (r_c/a), c) confinement free energy change as a function of volume fraction (η), and d) confinement free energy change as a function of the ratio of the chain unconfined size to a confining radius (r_G/r_c).

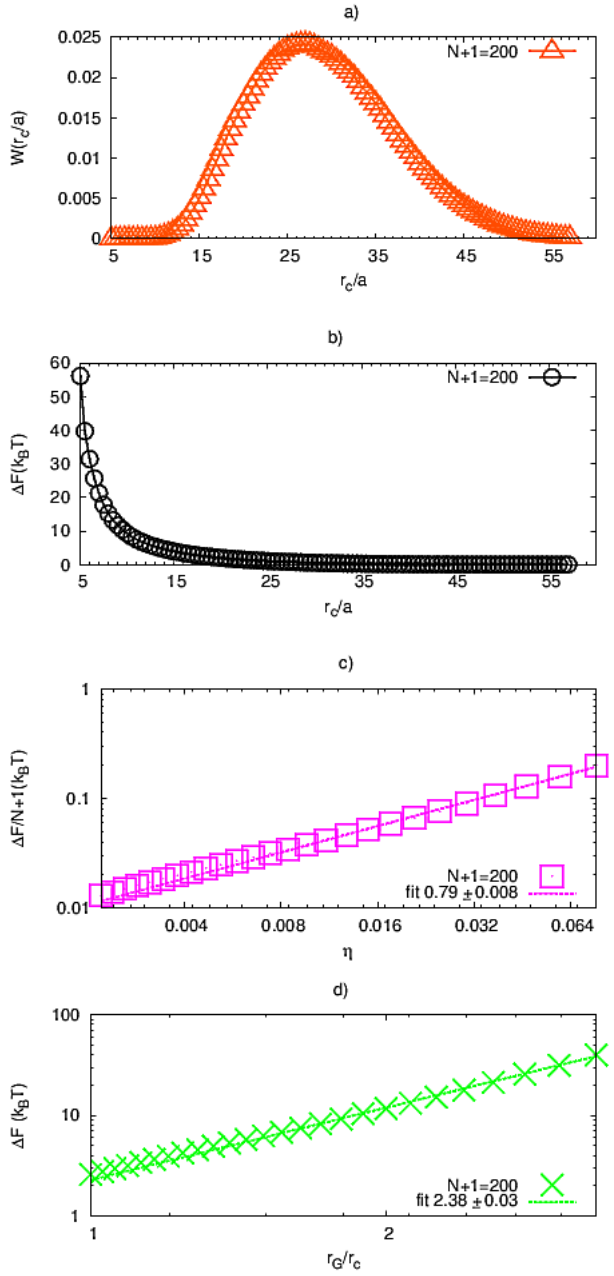

Figure B.8: a) Distribution of the furthermost monomer position ($W(r_c/a)$) for a non-ideal flexible chain of length 200 monomers, b) free energy change (ΔF) as a function of a confining radius (r_c/a), c) confinement free energy change as a function of volume fraction (η), and d) confinement free energy change as a function of the ratio of the chain unconfined size to a confining radius (r_G/r_c).

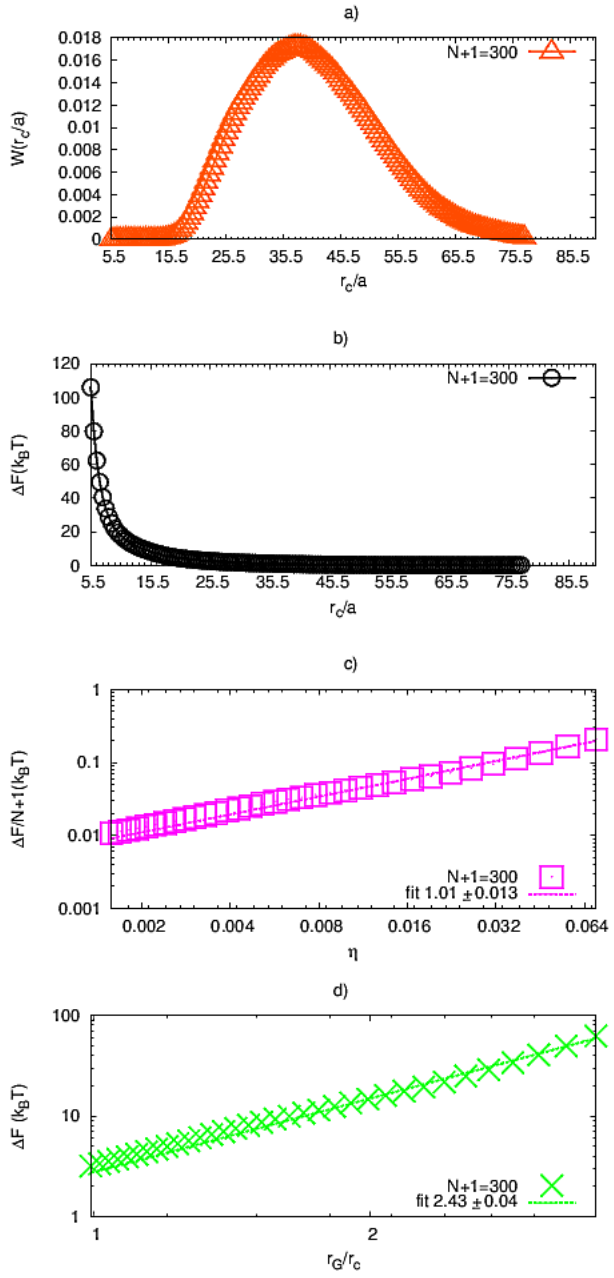

Figure B.9: a) Distribution of the furthermost monomer position ($W(r_c/a)$) for a non-ideal flexible chain of length 300 monomers, b) free energy change (ΔF) as a function of a confining radius (r_c/a), c) confinement free energy change as a function of volume fraction (η), and d) confinement free energy change as a function of the ratio of the chain unconfined size to a confining radius (r_G/r_c).

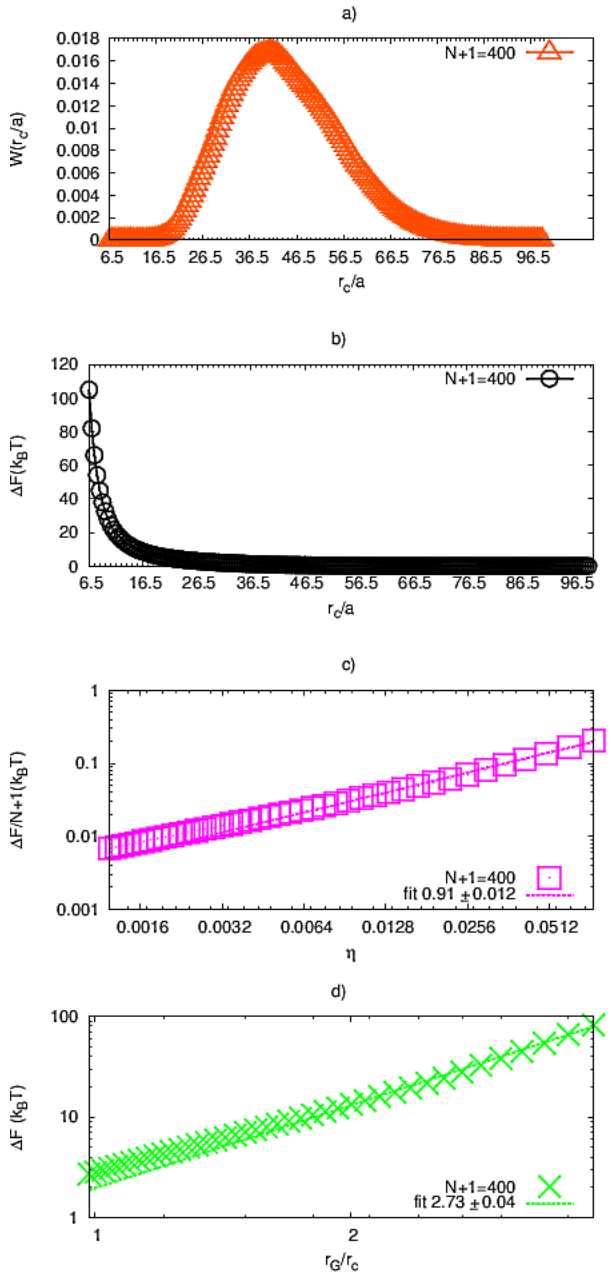

Figure B.10: a) Distribution of the furthermost monomer position ($W(r_c/a)$) for a non-ideal flexible chain of length 400 monomers, b) free energy change (ΔF) as a function of a confining radius (r_c/a), c) confinement free energy change as a function of volume fraction (η), and d) confinement free energy change as a function of the ratio of the chain unconfined size to a confining radius (r_G/r_c).

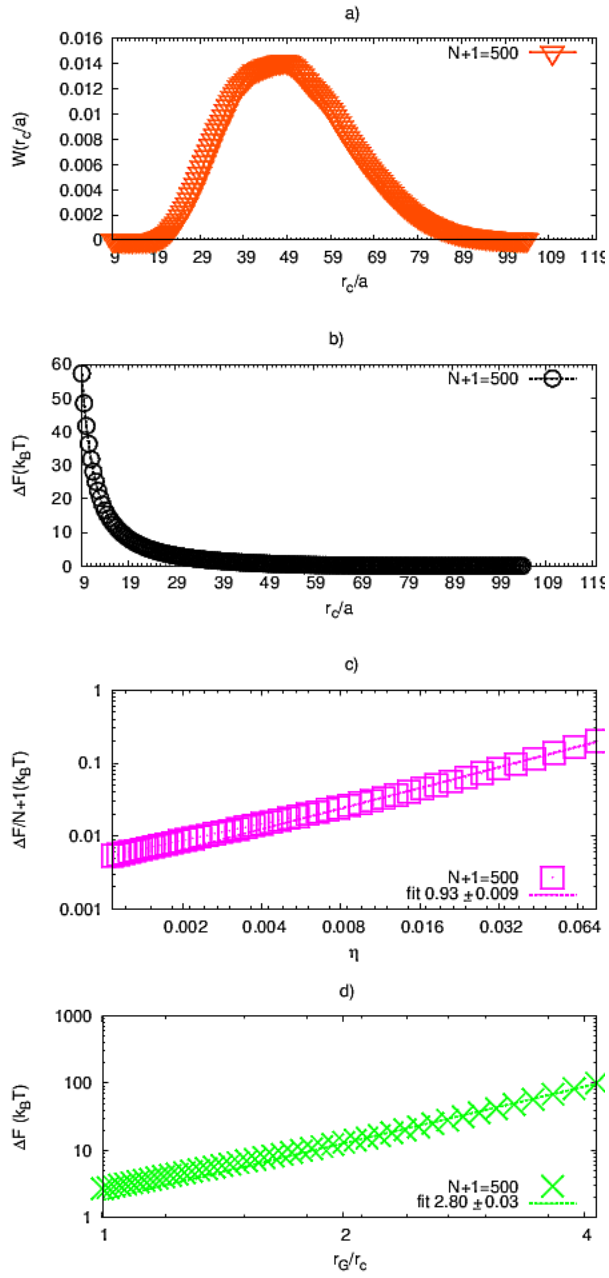

Figure B.11: a) Distribution of the furthermost monomer position ($W(r_c/a)$) for a non-ideal flexible chain of length 500 monomers, b) free energy change (ΔF) as a function of a confining radius (r_c/a), c) confinement free energy change as a function of volume fraction (η), and d) confinement free energy change as a function of the ratio of the chain unconfined size to a confining radius (r_G/r_c).

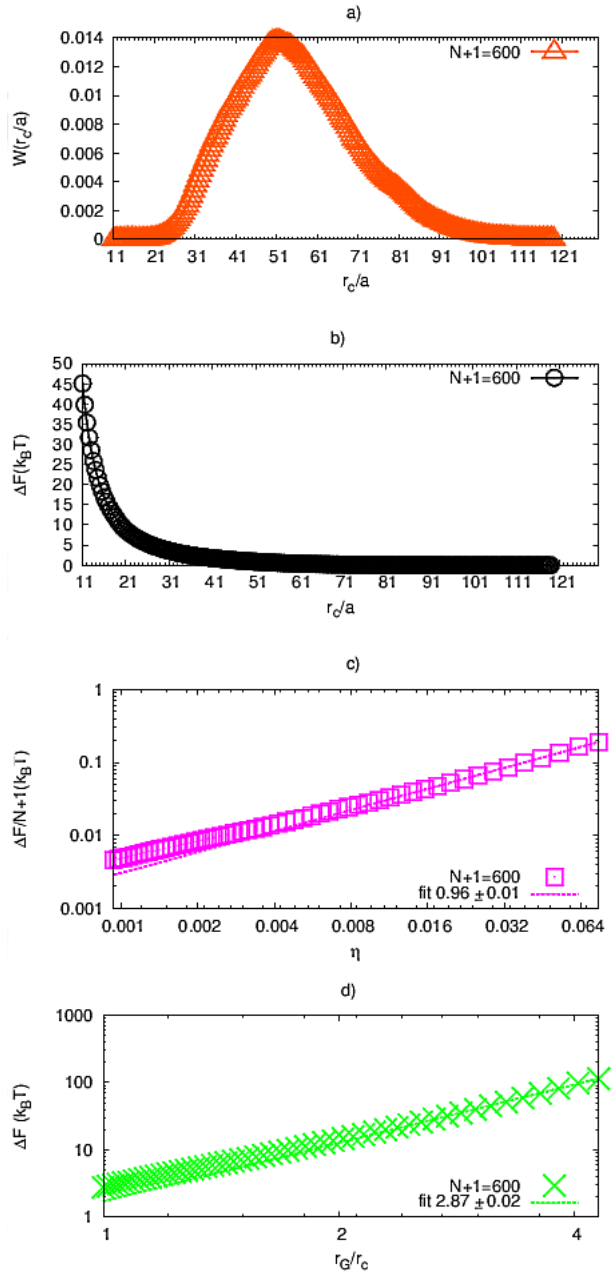

Figure B.12: a) Distribution of the furthermost monomer position ($W(r_c/a)$) for a non-ideal flexible chain of length 600 monomers, b) free energy change (ΔF) as a function of a confining radius (r_c/a), c) confinement free energy change as a function of volume fraction (η), and d) confinement free energy change as a function of the ratio of the chain unconfined size to a confining radius (r_G/r_c).