Molecular dynamics simulation of unzipping of a polymer hairpin by force

Jyoti Rani

A dissertation submitted for the partial fulfilment of BS-MS dual degree in Science

Indian Institute of Science Education and Research Mohali
April 2017

Certificate of Examination

This is to certify that the dissertation titled **Molecular dynamics simulation of unzipping of a polymer hairpin by force** submitted by **Ms. Jyoti Rani** (Reg. No. MS12128) for the partial fulfillment of BS-MS dual degree programme of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr. Abhishek Chaudhuri Dr. Sanjeev Kumar Dr. Rajeev Kapri
(Supervisor)

Dated: April 20, 2017

Declaration

The work presented in this dissertation has been carried out by me under the guidance of Dr. Rajeev Kapri at the Indian Institute of Science Education and Research Mohali.

This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

Jyoti Rani
(Candidate)

Dated: April 20, 2017

In my capacity as the supervisor of the candidates project work, I certify that the above statements by the candidate are true to the best of my knowledge.

Dr. Rajeev Kapri
(Supervisor)

Acknowledgment

I would like to express my sincere gratitude to my supervisor, Dr. Rajeev Kapri, for allowing me to work under his guidance and supporting me throughout the thesis. I doubt that I will ever be able to convey my appreciation fully. He has been a wonderful supervisor and mentor, far above the call of duty. Many thanks to my thesis committee, Dr. Abhishek Chaudhuri and Dr. Sanjeev Kumar for reading my thesis and offering valuable suggestions.

There are many people whose assistance and support have made this work possible, and I am grateful to all of them. I would especially like to thank:

Rajneesh Perhate for helping me to learn: how to work on cluster, some essential softwares and helpful discussions;

Suman Kalyan for the helpful discussions;

My parents, my siblings: who have supported me unconditionally throughout this time.

At last, I would also like to thank IISER Mohali for giving me this opportunity to study in such a research oriented environment and DST for providing me the INSPIRE Scholarship for Higher Education.

List of Figures

Abstract

Polymers have universal behaviour at long time scales and at long length scales. Large scale molecular dynamics simulations were carried out to study the structural properties of polymers. Specifically the dependence of chain length, also the excluded volume effect was introduced. For a linear polymer of chainlength N = 128 a good agreement was found with the size exponent. The model was extended to the hairpin configuration. The polymer hairpin unzipping has been studied via equilibrium force-induced unzipping and temperature-induced unzipping. The force-induced unzipping showed a first order phase transition. While the temperature-induced unzipping is found to be continuous phase transition.

Contents

Chapter 1

Introduction

A polymer is a large molecule containing many repeated subunits called monomers. Polymers can be both natural and synthetic. Natural polymers are important for structure, function and information storage in living cells. For example, in DNA, proteins, microtubules etc. Synthetic polymers are also known as man-made polymers. Many technological innovations have occurred in polymer synthesis in the past century which resulted in making polymers, one of the most commonly used class of materials. It has applications in the food, defence, plastics etc. to name a few. The size and weight differs from polymer to polymer. The degree of polymerisation is a fundamental property of polymers. However, the physical structure of the polymeric chain is also important for determining macroscopic properties. For example, the end-to-end distance is a physical property that estimates the size of a polymer. In the following we will introduce some of the theoretical models used to study polymers.

1.1 Ideal chains

The simplest model to study polymers is an ideal chain model. It models a polymer as a random walk and neglects any kind of interactions between the monomers along the chain (even if they approach each other in space)[1].

Let us evaluate some important quantities for the ideal chains. We can describe the polymer configuration by setting up the positions of each monomer of the chain. For this we can use Cartesian coordinates with respect to a fixed laboratory frame. Consider a linear polymer of $N+1$ backbone monomers a_i (with $0 \leq i \leq N$) as

[1]This condition is, however, never completely accomplished for real chains. The monomer of the real chain can interact with solvent and themselves.

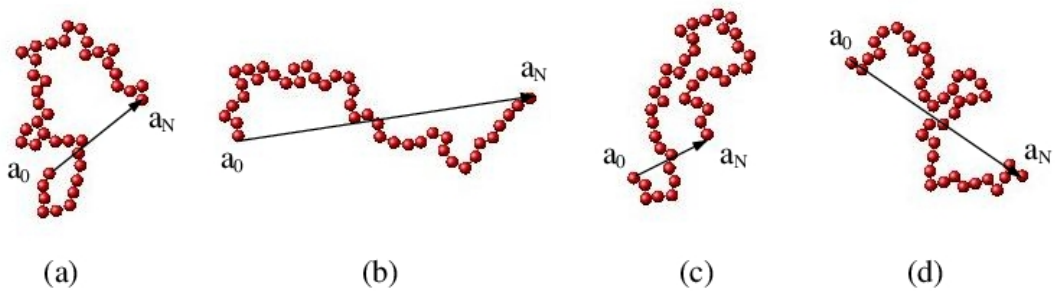

Figure 1.1: Different conformations of a linear polymer

shown in Fig 1.1. The monomers of the polymer can be identical or different. But, without loss of generality we can take them identical. The bond vector $\boldsymbol{r}_i$ goes from monomer a_{i-1} to monomer a_i. The polymer is in ideal state if there are no net interactions between any two monomers separated by many bonds along the chain so that $\mid i - j \mid \gg 1$. The various quantities that estimates the size of a polymer are:

1.1.1 The end-to-end vector

If we define the set of position vector of all the monomers: $\{\mathbf{R}[i]\} = (\mathbf{R}[0], \mathbf{R}[1]....\mathbf{R}[N])$, then the end-to-end vector, $\mathbf{R}_N$ is defined as the distance between the first monomer and last monomer of chain along some direction.

$$\boldsymbol{R}_N = \boldsymbol{R}[N] - \boldsymbol{R}[0] \tag{1.1}$$

where $\mathbf{R}[N]$ is the position vector of last monomer and $\mathbf{R}[0]$ is the position vector of first monomer. It can also be calculated as the sum of all N bond vectors ($r_i = R[i] - R[i-1]$) in the chain,

$$\boldsymbol{R}_N = \sum_{i=1}^{N} \boldsymbol{r}_i. \tag{1.2}$$

The different individual chains will have different conformations and therefore different end-to-end vectors. So the average end-to-end vector of an isotropic collection of chains having N backbone monomers is zero,

$$\langle \boldsymbol{R_N} \rangle = 0, \tag{1.3}$$

where, $\langle ... \rangle$ represents the ensemble average. It is an average over all possible states of the system accessed either by considering many chains or many different conformations of the same chain. The simplest non-zero average is the mean square end-to-end

distance given by,

$$\langle \boldsymbol{R}_N^2 \rangle = \langle \boldsymbol{R}_N . \boldsymbol{R}_N \rangle = \Big\langle \sum_{i=1}^{N} \boldsymbol{r}_i . \sum_{j=1}^{N} \boldsymbol{r}_j \Big\rangle = \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \boldsymbol{r}_i . \boldsymbol{r}_j \rangle. \tag{1.4}$$

If all bond vectors have the same length l, the scalar product can be represented in terms of the angle θ_{ij} between bond vectors $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$

$$\boldsymbol{r}_i . \boldsymbol{r}_j = l^2 \cos\theta_{ij}. \tag{1.5}$$

So, the mean-square end-to-end distance can be written as the double sum of average cosines:

$$\langle R_N^2 \rangle = \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \boldsymbol{r}_i . \boldsymbol{r}_j \rangle = l^2 \sum_{i=1}^{N} \sum_{j=1}^{N} \cos\theta_{ij}. \tag{1.6}$$

To calculate the double sum of cosine we need a model. The **freely-jointed chain model** is one of the basic models for describing the properties of ideal polymers. In this model, the bond vectors have a constant bond length $l = |\boldsymbol{r}_i|$ and the spatial orientation of each bond is independent from the orientations of other bond i.e., $\langle \cos\theta_{ij} \rangle = 0$ for $i \neq j$. So only N non-zero terms in the double sum survives ($\cos\theta_{ij} = 1$ for $i = j$). The mean square end-to-end distance for a freely-jointed chain becomes:

$$\langle R_N^2 \rangle = N l^2. \tag{1.7}$$

The probability distribution function of end-to-end vector $\boldsymbol{R}_N$ for an ideal linear chain of N monomers is given by:

$$P_{3d}(\boldsymbol{R}_N) = \left(\frac{3}{2\pi N b^2} \right)^{3/2} \exp\left(- \frac{3\boldsymbol{R}_N^2}{2Nb^2} \right), \tag{1.8}$$

where b is the Kuhn length. The mean-square end-to-end distance estimates the size of linear polymers. But it fails to do so for other architectures for e.g., branched or ring polymers, as they have too many ends or no ends at all. Therefore, for branched or ring polymers this quantity is not well defined.

1.1.2 Radius of gyration

Since all objects possess a **radius of gyration** R_g, so it can also be used to characterize the size of polymers of any architecture. The square radius of gyration is defined as the average square distance between monomers in a given conformation and the polymer's centre of mass, $\boldsymbol{R}_{cm}$

$$R_g^2 \equiv \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{R}_i - \boldsymbol{R}_{cm})^2. \tag{1.9}$$

This expression is useful only when the position of centre of mass $\boldsymbol{R}_{cm}$ is known for the polymer or is easy to evaluate. Otherwise the following expression is useful

$$R_g^2 = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}(\boldsymbol{R}_i - \boldsymbol{R}_j)^2. \tag{1.10}$$

For polymer chains, this squared radius of gyration is usually averaged over the ensemble of allowed conformations of chains which gives the mean-square radius of gyration:

$$\langle R_g^2\rangle = \frac{1}{N}\sum_{i=1}^{N}\langle(\boldsymbol{R}_i - \boldsymbol{R}_{cm})^2\rangle = \frac{1}{N^2}\sum_{i=1}^{N}\langle(\boldsymbol{R}_i - \boldsymbol{R}_j)^2\rangle. \tag{1.11}$$

For an ideal linear chain,

$$\langle R_g^2\rangle = \frac{b^2 N}{6} = \frac{\langle R_N^2\rangle}{6}, \tag{1.12}$$

which is a fundamental result of polymer physics for ideal chains.

1.1.3 The structure factor

The structure factor is another quantity that is used for estimating the size of polymer chain. It is defined as the ratio of intensity scattered at an angle θ (scattering wavevector $\boldsymbol{q}$) to that extrapolated to zero angle ($\theta \to 0$). The expression for the structure factor for isotropic collections of chains is given by:

$$S(\boldsymbol{q}) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\sin(\boldsymbol{q}\boldsymbol{R}_{ij})}{\boldsymbol{q}\boldsymbol{R}_{ij}}. \tag{1.13}$$

The structure factor at low scattering angles can be rewritten, by using Taylor series expansion, as

$$S(\boldsymbol{q}) = 1 - \frac{1}{3}q^2\langle R_g^2\rangle + \ldots. \tag{1.14}$$

The structure factor can be obtained experimentally by X-ray, electron and neutron diffraction experiments.

We have seen above that in ideal chains there are no interactions between monomers of the chain that are far apart even when they are close enough in space. This behaviour is not completely realized for real chains. However, for many real polymers one can find a suitable solvent known as θ-solvent, in which, at some special temperature called the θ-temperature, the polymer takes nearly ideal chain conformations. For example, polystyrene in cyclohexane at $\theta \cong 34.5°\text{C}$ [1]. This can be understood as follows: In real chains the relative strength of interactions determines whether the monomer effectively attract or repel one another. At high temperatures, the repulsive interactions are dominant that results in chain swelling. However, if we go to lower temperatures, the attractive interactions are dominant over the repulsive interactions and the chain can be found in a collapsed conformation. At some intermediate temperature (θ-temperature) the attractive and repulsive part of monomer-monomer interactions cancel each other, which gives a nearly ideal conformation of chains. In concentrated solutions and linear polymer melts, interactions among the monomers are completely screened by the surrounding chain and they have practically ideal chain conformation.

1.2 Real chains

For the sake of simplicity, the ideal chains neglects an important physical ingredient i.e., the interaction experienced by two distinct monomers that are located far distance along the chain. These two monomers can not occupy the same volume. This is called excluded volume effect. It is very important in dilute solutions, where it leads to overall polymer swelling in comparison to the corresponding ideal situation. More importantly, excluded volume effects change the scaling of polymer size with the degree of polymerisation N (i.e., the total number of monomers in a polymer molecule). For real chains

$$\langle R_N^2 \rangle \sim N^{2\nu}, \tag{1.15}$$

where ν is the size exponent. Let us introduce the excluded volume interaction in the polymer models and try to find out how this repulsion of the monomers affects the size of the polymer.

1.2.1 Flory theory for the polymer in a good solvent

A simple way of accounting for the fact that non-consecutive monomers cannot interpenetrate, is provided by a hard-sphere repulsion. The total free energy $F(N, R)$ of the system can then be estimated as follows:

1. In 3-dimensions, the hard sphere repulsion is proportional to the excluded volume v_{ex} of each pair of monomers, times the number of monomer pairs (N^2) per unit of available volume (R^3), that is,

$$\text{repulsive energy} \sim v_{ex}\frac{N^2}{R^3}. \tag{1.16}$$

2. The entropy of the chain having N monomers with end-to-end distance R is given by

$$S(N, R) = k_B \log P(R, N) \sim -\frac{R^2}{Nb^2}, \tag{1.17}$$

where we have used Eq. (1.8) for the probability distribution.

So, the total free energy $F(N, R)$ of the system is given by

$$F(N, R) = F(N, 0) + a\frac{R^2}{Nb^2} + bv_{ex}\frac{N^2}{R^3}, \tag{1.18}$$

where a and b are temperature dependent constants and $F(N, 0)$ is an additive constant to the free energy. The optimal size of a real chain is obtained by minimizing the total free energy. This argument can be generalized to an arbitrary d-dimension. The entropy term is independent of d, but the excluded volume term is replaced by N^2/R^d, where R^d is the volume occupied by the polymer. Writing $F(N, R)$ in d-dimensions and minimizing $F(N, R)$ we obtain

$$\langle R\rangle \sim N^{3/(d+2)}. \tag{1.19}$$

Therefore, the Flory theory predicts that size of the polymer depends on the spatial dimension with size exponent, $\nu = 3/(d + 2)$. Therefore,

- In 1-dimension $\nu = 1$,
- In 2-dimensions $\nu = 3/4$,
- In 3-dimensions $\nu = 3/5$, which is very close to the value $\nu \approx 0.588..$ obtained in experiments [2].

The Flory theory predicts that in 4-dimensions $\nu = 1/2$, which is same as the ideal polymer. This tells that in 4-dimensions the excluded volume interactions are not at all important. In $d > 4$, Equation 1.19 gives the size exponent $\nu < 1/2$, which is not possible, because a repulsion cannot make a polymer chain more compact than a free chain.

The ideal chain model can be solved analytically but too few analytical results are available for the real chains. To get a better understanding of real chains, computer simulations have played a major role [3]. Two important classes of simulation techniques which can be used for this purpose are Monte Carlo (MC) and Molecular Dynamics (MD) simulations.

MC simulations are more suited to simulate a model polymer in discrete space, i.e., on a lattice, where as, in MD simulations the underlying space is continuous. In MD simulations, given the initial positions and velocities of the monomers of the polymer, the Newton's equations of motion for them are integrated at each time step to obtain their positions and velocities as a function of time. From these trajectories, one can calculate different system properties. Monte Carlo techniques are usually orders of magnitude faster than MD simulations in term of computer time. We will talk more about MD simulations in Chapter 2.

1.3 Polymer-hairpin

In biomolecules like ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) the complementary bases are bound together by hydrogen bonds. In the case of DNA, the adenine-thymine (A-T) has two- and the guanine-cytosine (G-C) has three-hydrogen bonds. Therefore, the later are more stable. In the case of RNA, the base thymine is replaced by uracil, and the adenine-uracil (A-U) again has two hydrogen bonds. In a single stranded DNA or, more commonly, in RNA, for suitable sequences that can fold back on itself, stem-loop intramolecular base pairing can occur which forms a pattern known as hairpin or a hairpin loop. The hairpin structures are important in many biological processes such as prokaryotic transcription termination. The hairpin loop forms in an mRNA strand during transcription and causes the RNA polymerase to become dissociated from the DNA template strand. For the sake of simplicity, let us consider a hairpin loop having only one type of base pairs.

The hairpin structure can be destabilized in different ways:

- By heating the solution containing polymer hairpins. This is known as thermal denaturation,
- By changing the pH of the solution, know as chemical denaturation, and
- By applying an external force on the ends of the hairpin. This is known as force induced unzipping.

Thermal denaturation is a subject of studies since the early sixties. At lower temperatures, the polymer remains in a zipped configuration (hairpin state) to minimize its internal energy. As the temperature of the solution is increased, the polymer tries to maximize its entropy. This is done by breaking the bonds between the complementary base pairs. At some temperature, which we call as melting temperature, T_M, the entropy of the polymer wins over the energy and the polymer hairpin opens up. The thermal denaturation of polymer hairpin is a continuous phase transition.

The hairpin can be unzipped below its melting temperatures by applying an external pulling force g at one end while keeping the other end anchored. It was found that for lower forces it remains in a zipped state. If g exceeds a critical value $g_c(T)$, the polymer unzips from the hairpin state to an open configuration. This is know as unzipping transition. Unlike thermal denaturation, the unzipping transition is a first order phase transition. The force induced unzipping can be studied experimentally by using single molecule manipulation techniques such as optical tweezers and atomic microscopy. For example, optical tweezers have been used to apply force and pull apart the two strands at one end of the molecule [4].

In this project our aim is to study the equilibrium properties of a linear polymer, and the unzipping transition of a polymer-hairpin by a pulling force using molecular dynamics (MD) simulations. We will not include hydrodynamic effects in our simulations.

The thesis is organized as follows: In Chapter 2 we presents a description of MD simulation of a linear polymer in the presence of a heat bath. In Chapter 3 we report our results on thermal denaturation and force induced unzipping of a polymer hairpin. Finally, in Chapter 4 we summarize our results.

Chapter 2

Molecular Dynamics Simulation of Polymer

In Chapter 1, we saw that real chains interact with the solvent and with themselves. This leads to a change in the size exponent which depends on spatial dimension. In this chapter we define our model for a real polymer and study its dynamics in the presence of a heat bath using MD simulations.

2.1 Model

The self-avoiding polymer is modeled by the coarse-grained bead-spring chain [5]. The beads represent the monomers of the polymer and the spring mimics the bond that connects the consecutive monomers. The beads of the polymer experience an excluded volume interaction modeled by Weeks-Chandler-Andersen (WCA) potential [6] of the form

$$U_{ij}^{WCA}(r) = \begin{cases} 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} + \frac{1}{4} \right] & \text{for } r \leq 2^{1/6}\sigma \\ 0 & \text{for } r > 2^{1/6}\sigma, \end{cases} \tag{2.1}$$

where, ϵ is the strength of the potential and σ is the diameter of the bead. The consecutive monomers of the chain are bounded by finite extension nonlinear elastic (FENE) potential of the form

$$U_{ch}^{FENE}(r) = -\frac{1}{2} k R_0^2 \ln\left(1 - \frac{r^2}{R_0^2}\right), \tag{2.2}$$

where k is the spring constant and R_0 is the maximum allowed separation between consecutive monomers of the chain. It is an attractive potential modelling interaction between the bonds for a distance less than the cut-off distance (R_0).

The motion of monomers is governed by the following **Langevin equation**

$$\frac{d^2\boldsymbol{r}_i}{dt^2} = -\boldsymbol{\nabla} U_i - \zeta \frac{d\boldsymbol{r}_i}{dt} + \boldsymbol{\Gamma}_i(t), \tag{2.3}$$

where $\boldsymbol{r}_i$ is the position of the ith monomer, ζ is the bead friction, and $\boldsymbol{\Gamma}(t)$ is a time dependent random force acting on the monomer which satisfies the following fluctuation-dissipation relation

$$\langle \boldsymbol{\Gamma}_i(t) \cdot \boldsymbol{\Gamma}_j(t') \rangle = 6 k_B T \zeta \delta_{ij} \delta(t - t'). \tag{2.4}$$

In our simulation, the units of length, energy and mass are set by σ, ϵ, and m, respectively. This sets the unit of time as $\tau = \sigma(m/\epsilon)^{1/2}$. We have used the following parameters in our simulations: $A = 30$, $r_c = 1.5$, $R_0 = 1.5\sigma$, $k = 30.0\epsilon/\sigma^2$, $T = 3\epsilon$, $\Delta t = 0.008\tau$, and $\zeta = 0.5$.

2.2 Algorithm

We use Large Scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) package [6] for simulating the model defined in the previous section.

2.2.1 Creating initial configuration files

As we have seen in Chapter 1 that in MD simulations one needs to integrate the newton's equation of motion for each monomer of the chain. To start the simulation one needs to provide input for the initial configuration of the polymer in a cubical box. In the present simulation we have used **chain.f** tool provided by the LAMMPS package [6], which generates self-avoiding walk using MC method, to create initial configuration of the polymer.

2.2.2 Equilibration of system

Once the initial configuration is created, our aim now is to equilibrate it. For this purpose, we first replace the WCA potential by a soft-pair cosine potential, which

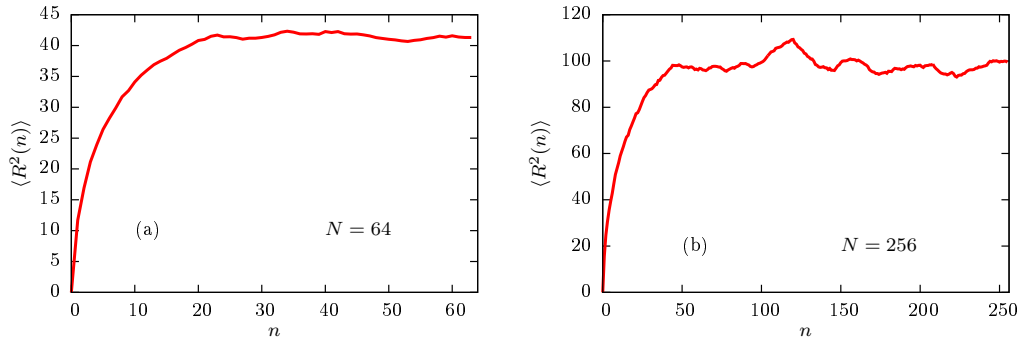

Figure 2.1: Mean square internal distances $\langle R^2(n)\rangle$ as a function of n. MSID for chain length (a) $N = 64$, (b) $N = 256$

allows the overlapping of monomers, of the form [6]

$$U^{soft} = A\left[1 + \cos\frac{\pi r}{r_c}\right], \tag{2.5}$$

where A and r_c are in units of energy and length. One important property of this potential is that this does not blow up in the limit $r \to 0$. We start with a very soft potential (low A) and put a restriction on the maximum displacement of monomer allowed in one time step. We run the simulation for few time steps and increase the rigidity of the potential (by increasing value of A) in steps. This is needed to relax the FENE bond style. Once the FENE bond relaxes we replace the soft-pair cosine potential by the original WCA potential and run the simulation for equilibration. A LAMMPS script to achieve this is given in appendix A.1.

To check if the system is equilibrated or not, we calculate the mean square internal distance (MSID), which involves average squared distance between monomers j and $j + n$ of the same chain [7]

$$\langle R^2(n)\rangle = \frac{1}{N-n}\sum_{j=1}^{N}(\boldsymbol{r}_j - r_{j+n})^2. \tag{2.6}$$

In Fig. 2.1, we have plotted MSID $\langle R^2(n)\rangle$ as a function of n for two different chain lengths $N = 64$ and 256. We see that for chain length $N = 64$, it saturates at the correct end-to-end distance and hence the chain attains equilibrium. But, it has not saturated for $N = 256$ showing that the chain has not attained the equilibrium.

Once the chain is equilibrated, we evolve the system and obtain the positions of each monomer as a function of time, $\boldsymbol{r}_i(t)$, i.e., the chain trajectories, for various chain

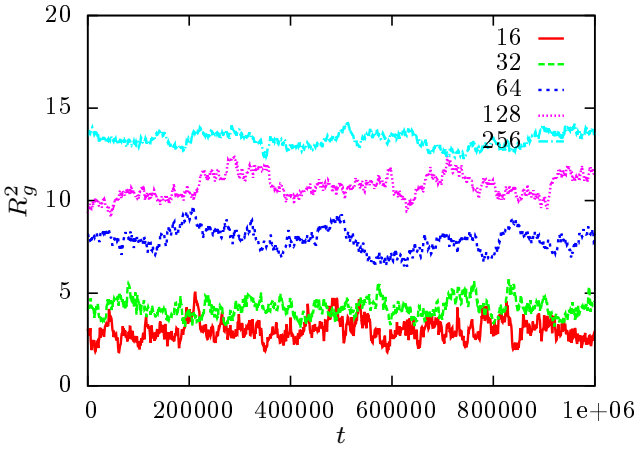

Figure 2.2: The variation of squared radius of gyration, R_g^2 as a function of time t.

lengths and analyse them.

2.3 Results and Discussions

We have studied polymers of length $N = 16$, 32, 64, 128 and 256 and analysed the trajectories to obtain the following quantities.

2.3.1 Mean square end-to-end distance and radius of gyration

In Fig. 2.2, we have plotted the squared radius of gyration, R_g^2 as a function of time t for various chain lengths. It is evident from the plot that this quantity fluctuate about a particular value. Similar behaviour has been observed for the squared end-to-end distance also.

From these trajectories, we obtain the mean square end-to-end distance, $\langle R_N^2 \rangle$, and mean square radius of gyration $\langle R_g^2 \rangle$ as a function of chain length N. The averaging is done over 10^6 time steps. These are shown in Fig. 2.3. These plots show that both $\langle R_N^2 \rangle$ and $\langle R_g^2 \rangle$ increase as we increase the number of monomers in the chain according to a power law $\langle R_N^2 \rangle \sim N^{2\nu}$, and $\langle R_g^2 \rangle \sim N^{2\nu}$, where ν is the size exponent. On fitting the simulation data we obtained $\nu = 0.63 \pm 0.02$. This value is slightly greater than $\nu \approx 0.59$ which is the obtained in experiments. The difference is due to the finite size effects as we have worked with shorter chains. The simulations with longer chains have got results that are very close to experimentally obtained value [8].

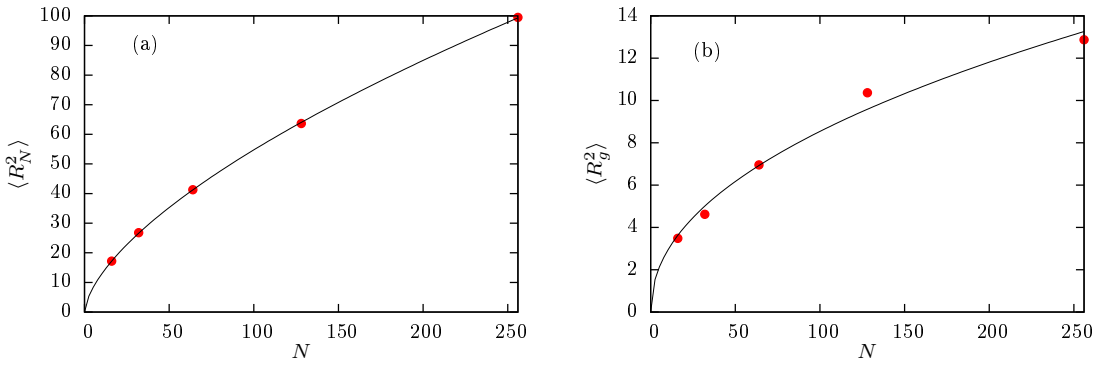

Figure 2.3: The mean-square end-to-end distance $\langle R_N^2 \rangle$ and mean-square radius of gyration $\langle R_g^2 \rangle$ as a function of chain length N are shown in (a) and (b), respectively.

In Chapter 1, we have seen that for ideal chains the ratio of $\langle R_N^2 \rangle$ to the $\langle R_g^2 \rangle$ is independent of chain length (Eq. (1.12)). This relation is obeyed in the asymptotic limit and is true even for real chains. In Fig. 2.4, we have plotted $\langle R_N^2 \rangle / \langle R_g^2 \rangle$ for various chain lengths. We see that except of small chain lengths, $N = 16$, which shows a marked deviation, all other chain lengths that we simulated are very close to actual value (i.e., $\langle R_N^2 \rangle / \langle R_g^2 \rangle = 6$) which is shown by the straight line.

2.3.2 Structure factor

In Chapter 1 we have seen that for low scattering angles the structure factor is proportional to the mean square radius of gyration, $\langle R_g^2 \rangle$ (see Eq. (1.14)). Since $\langle R_g^2 \rangle \sim N^{2\nu}$,

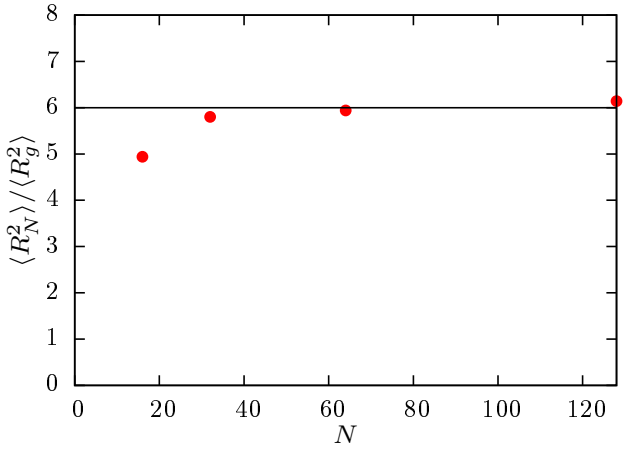

Figure 2.4: The ratio of $\langle R_N^2 \rangle$ to the $\langle R_g^2 \rangle$ for various chain lengths N.

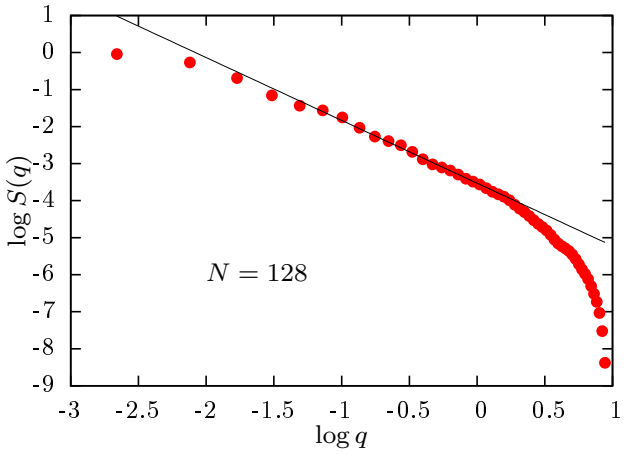

Figure 2.5: The structure factor, $S(q)$, as a function of scattering wavevector q. The line represents the best fit for the curve with slope $m = -1.69$.

a little algebra shows that for small scattering angles the structure factor obeys

$$S(q) \propto q^{-1/\nu} \qquad \text{for} \qquad \frac{2\pi}{\langle R^2 \rangle^{1/2}} \ll q \ll \frac{2\pi}{\sigma}. \tag{2.7}$$

This can be used to estimate the size of the polymer experimentally and hence the size exponent ν. In Fig. 2.5, we have shown a log-log plot of $S(q)$ and q for a chain of length $N = 128$. By fitting the simulation data for lower q values we estimate the value for size exponent, $\nu = 0.59 \pm 0.01$, which is in good agreement with the experimentally obtained value.

2.4 Conclusions

In this chapter, we studied the dynamics of a linear polymer using a coarse-grained bead-spring model. We found that the quantities obtained by using this simple model matches excellently with the experimental results. In this section we summarise the results:

- The mean square end-to-end distance $\langle R_N^2 \rangle$ as well as the radius of gyration $\langle R_g^2 \rangle$ varies with the chain length N as $\sim N^{2\nu}$, where ν is the size exponent. These quantities can be used to estimate the size of the polymer. In three-dimensions, the size exponent is $\nu \approx 0.59$. One needs longer chain lengths to extract this value from these quantities.

- For small angle scattering, the structure factor behaves with wavevector as $S(q) \sim q^{-1/\nu}$. This quantity can be obtained experimentally by using X-ray, electron, or neutron diffraction, and hence the size exponent ν. We found that $\nu = 0.59 \pm 0.01$ from the simulation for the chain of length $N = 128$. This is very close to experimentally obtained value $\nu = 0.588$.
- The ratio $\dfrac{\langle R_N^2 \rangle}{\langle R_g^2 \rangle} \approx 6$ for higher chain lengths.

Chapter 3

Unzipping of a Polymer hairpin

In Chapter 2 we have obtained some important quantities of interest for a bead-spring model of polymer using MD simulations at some temperature. In this chapter, we develop this model to study the thermal denaturation and the force-induced unzipping of a polymer hairpin. In doing so, we will obtain the unzipping phase diagram in the temperature-force plane.

3.1 Model

We model the polymer-hairpin by bead-spring chain of length N as shown in Fig. 3.1(a). The ith $(1 \leqslant i \leqslant N/2)$ monomer is complementary to $(N+1-i)$th monomer of the chain and, at lower temperatures, they can form a base pair if they are close enough. This can be a simple model for studying RNA hairpins. The configurational energy of the system under consideration is given by

$$U = \sum_{i=1}^{N-1} -\frac{1}{2} k R_0^2 \ln\left(1 - \frac{r_{i,i+1}^2}{R_0^2}\right) + \sum_{i=1}^{N-2} \sum_{j>i+1}^{N} 4\epsilon \left[C \left(\frac{\sigma}{r_{i,j}}\right)^{12} - A_{ij} \left(\frac{\sigma}{r_{i,j}}\right)^{6} \right]. \quad (3.1)$$

The first term in the above expression is the finite extension nonlinear elastic (FENE) potential, discussed in the previous chapter, which models the bond between the consecutive monomers of the chain, and the second term represents the interactions between the monomers. Most pairs of the chain interact only through short-range repulsion (excluded volume). This has been taken care by assigning $C = 1$ in Eq. (3.1). The correct base pairing is taken care by assigning $A_{ij} = 1$ to the complementary monomers, and $A_{ij} = 0$ for all other pairs.

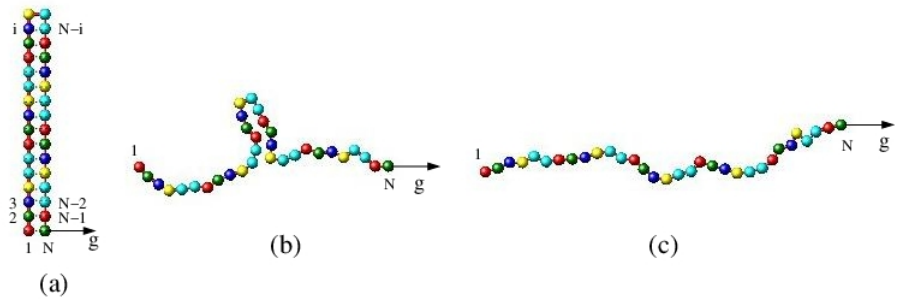

Figure 3.1: Model: Force-induced unzipping (a) Hairpin configuration (bound state) (b)Semi-hairpin state (c) Open configuration (unbound state)

The dynamics of the system is again given by the Langevin equation (Eq. (2.3)) with random force Γ having zero mean and satisfying the fluctuation-dissipation relation

$$\langle \mathbf{\Gamma}_i(t) \cdot \mathbf{\Gamma}_j(t') \rangle = 4\zeta k_B T \delta_{ij}(t - t'), \tag{3.2}$$

since we are working in two dimensions. The following parameters are used in our simulations: $R_0 = 1.5\sigma$, $k = 100.0\epsilon/\sigma^2$, $\Delta t = 0.01\tau$, and $\zeta = 0.4$.

We use LAMMPS package to simulate the model. We take the initial configuration of the polymer in the form of a hairpin. A java code used for this purpose is given in Appendix A.2.

3.2 Results and Discussions

3.2.1 Temperature denaturation

Firstly, we study the effect of temperature T on the polymer hairpin. The first monomer is anchored at the origin and the rest of the monomers are allowed to fluctuate. We start with a low temperature and hairpin is allowed to equilibrate with a Langevin heat bath. Once the system gets equilibrated we monitor the end-to-end separation of the hairpin which serves as an order parameter. We increment the temperature in steps and repeat the above process.

In Fig. 3.2(a), we have shown the scaled average end-to-end separation $\langle x \rangle / \sqrt{N}$ as a function of temperature for three different chain lengths $N = 32, 48$, and 64. We see that below some temperature T_M the average separation between the end monomers of the chain is zero and above T_M it jumps to some finite value which scales as $N^{1/2}$. The figure also shows that as the length of the hairpin increases the jump becomes

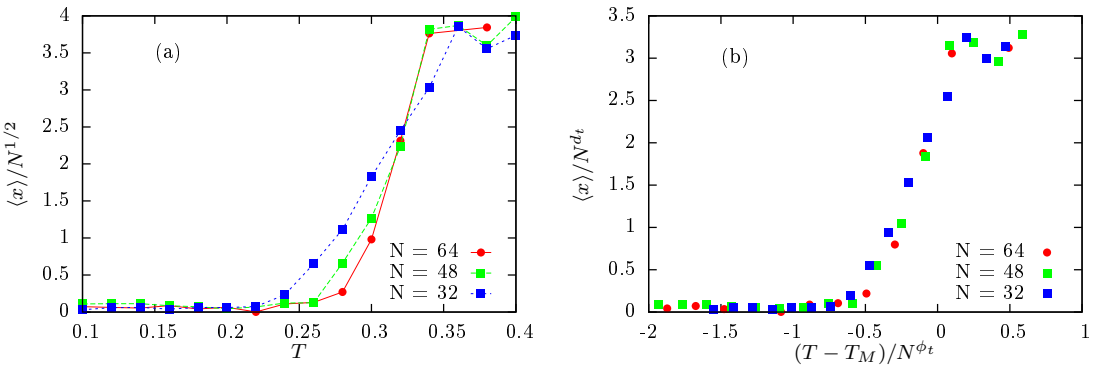

Figure 3.2: (a) represents the average end-to-end separation $\langle x\rangle/\sqrt{N}$ as a function of Temperature T and (b) is its data collapse.

sharper and sharper. In the thermodynamic limit, i.e., $N \to \infty$, this will become a step function. To find the melting temperature T_M we do a finite size scaling of the form

$$\langle x\rangle \sim N^{d_t}\mathcal{G}\left(\frac{|T-T_M|}{N^{\phi_t}}\right), \tag{3.3}$$

where d_t and ϕ_t are critical exponents, T_M is the melting temperature, and $\mathcal{G}$ is the scaling function. We found a good collapse for various chain lengths for the values $T_M = 0.33 \pm 0.02$, $\phi_t = 0.5 \pm 0.05$ and $d_t = 0.5 \pm 0.05$. The data collapse is shown in Fig. 3.2(b). From the above scaling (i.e., Eq. (3.3)), we get

$$\langle x\rangle \sim |T-T_M|^2, \tag{3.4}$$

showing that the order parameter $\langle x\rangle$ varies continuously as the melting temperature is approached. Therefore, the thermal denaturation of a polymer hairpin is a continuous transition.

3.2.2 Force-induced unzipping

Now we study the effect of pulling force g on the polymer hairpin. The first monomer is anchored at the origin and a fixed pulling force g is applied on the last monomer. The hairpin is allowed to equilibrate with a Langevin heat bath kept at a fixed temperature T. Once the system gets equilibrated we monitor the end-to-end separation of the hairpin which again serves as an order parameter. We increment the force in steps and repeat the above process.

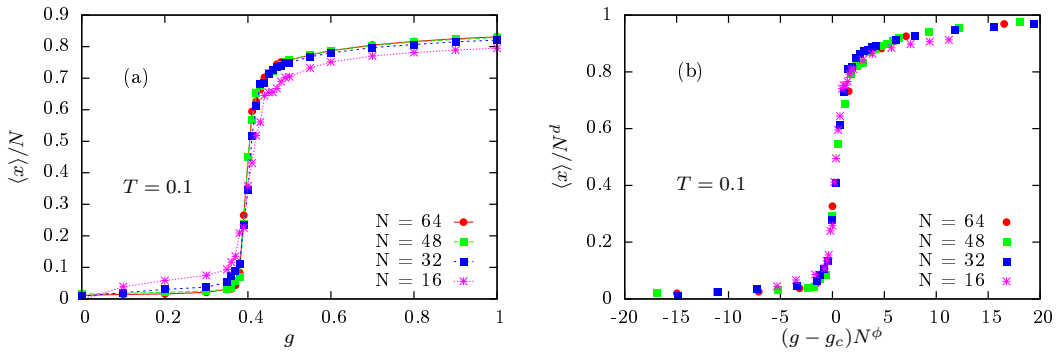

Figure 3.3: (a) Scaled average end-to-end separation $\langle x \rangle / N$ as a function of pulling force g for various chain lengths $N = 16, 32, 48$, and 64 at $T = 0.1$. (b) Collapse of data shown in (a).

In Fig 3.3(a) we have plotted the scaled average separation $\langle x \rangle / N$ as a function of pulling force g for various chain lengths $N = 16, 32, 48$, and 64 at $T = 0.1$. It clearly shows that for lower value of forces, the hairpin is in the zipped state where the complementary bases are in the bound state. On increasing the force g, the polymer remains in the bound state until the pulling force exceeds a critical value g_c, which depends on the temperature. For $g > g_c$, all the base pairs of the hairpin break as the polymer is in the unzipped state.

To obtain the critical force g_c, we use finite size scaling of the form

$$\langle x \rangle \sim N^d \mathcal{F}\left(|g - g_c| N^\phi\right), \tag{3.5}$$

where, d and ϕ are the critical exponents, g_c is the critical force, and $\mathcal{F}$ is the scaling function. In Fig. 3.3(b) we have shown the collapse of end-to-end separation data for various chain lengths. We obtain a nice collapse for $d = 0.95 \pm 0.05$, $\phi = 1.0 \pm 0.05$ and $g_c = 0.39 \pm 0.01$. From this scaling one can see that

$$\langle x \rangle \sim \frac{1}{\mid g - g_c \mid}, \tag{3.6}$$

which shows a jump in the order parameter as one approaches the critical force, g_c, implying a first order phase transition.

The isothermal extensibility, χ, which is the response of an external pulling force on the extension can be obtained by

$$\chi = \left.\frac{\partial \langle x \rangle}{\partial g}\right|_T \equiv \frac{1}{k_B T}\left(\langle x^2 \rangle - \langle x \rangle^2\right), \tag{3.7}$$

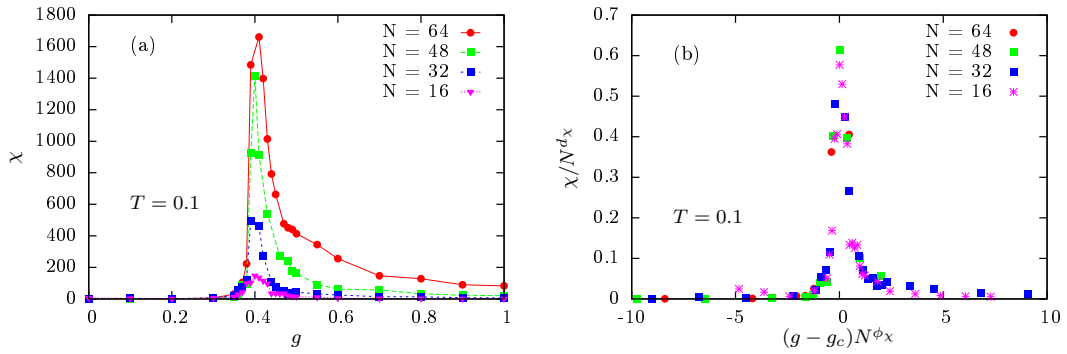

Figure 3.4: (a) Isothermal extensibility χ as a function of external pulling force g for chains of various lengths $N = 16, 32, 48$, and 64 at $T = 0.1$, and (b) the collapse of data shown in (a).

where $k_B = 1$ is the Boltzmann's constant.

In Fig. 3.4(a) we have shown the isothermal extensibility for chains of various lengths $N = 16, 32, 48$, and 64 at $T = 0.1$. As the pulling force increases, the value of χ also increases. It reaches a maximum at a force value $g = g_c$ and then starts decreasing again for $g > g_c$ values. The peak in χ, which shifts with chain length, is due to the finite length of the polymer. In the limit of infinite chain length χ shows a discontinuity at g_c. To obtain the value of critical force g_c we again resort to finite size scaling

$$\chi \sim N^{d_\chi} \mathcal{H}\left(|g - g_c| N^{\phi_\chi}\right), \tag{3.8}$$

where, d_χ and ϕ_χ are the critical exponents, g_c is the critical force, and $\mathcal{H}$ is the scaling function. We obtain a nice collapse for $g_c = 0.39 \pm 0.01$ and $\phi_\chi = 1.0 \pm 0.05$ and $d_\chi = 2.05 \pm 0.05$, which again shows

$$\chi \sim \frac{1}{|\ g - g_c|}, \tag{3.9}$$

a discontinuity in the response function, i.e., a first order phase transition.

We can repeat the above procedure at various temperatures and obtain the phase diagram of unzipping of a polymer hairpin for fixed pulling force. The phase diagram in force-temperature plane is shown in Fig. 3.5. The points are obtained from the simulations and the line is the best fit to the data. This is the phase boundary between the two phases. The polymer remains in the bound phase, i.e., having a hairpin configuration, below the phase boundary and remains in the open state above it. The phase diagram also shows that at the melting temperature $T = T_M$, the

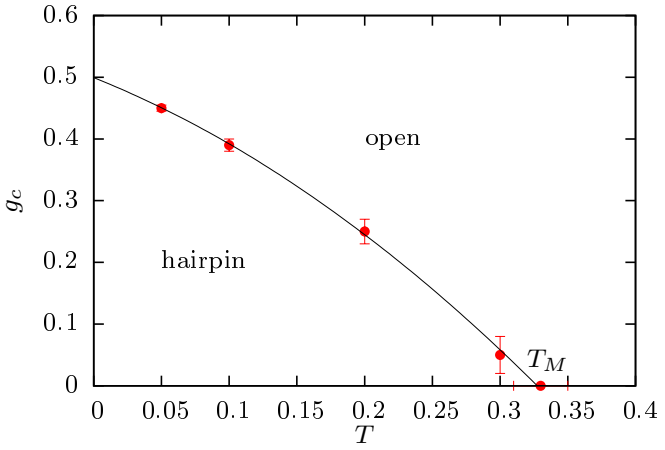

Figure 3.5: The force g_c vs T phase diagram of unzipping of a polymer hairpin for fixed pulling force

critical force needed to unzip the polymer hairpin is zero.

3.3 Conclusions

Using simple theoretical arguments and MD simulations we studied the thermal denaturation and force-induced unzipping of polymer haipins of lengths upto $N = 64$.

For the temperature, below the melting temperature, the polymer-hairpin remains in the bound state. But for $T > T_M$, the hairpin shows a transition from bound to unbound state. The order of this transition has been predicted from the scaling of the order parameter average end-to-end separation, $\langle x \rangle$ shown in Fig 3.2(b). This scales as $\langle x \rangle \sim N^{d_t} \mathcal{G}\left(\frac{T-T_M}{N^{\phi_t}}\right)$ with melting temperature $T_M = 0.33 \pm 0.02$ and exponents $\phi_t = 0.5 \pm 0.05$ and $d_t = 0.5 \pm 0.05$. From this scaling one can see that $\langle x \rangle \sim |T - T_M|^2$. The order parameter varies continuously as the melting temperature is approached. This result implies that thermal denaturation of a polymer hairpin is a continuous phase transition.

Polymer hairpin also shows transition from bound state to unbound state for the temperature values less than the melting temperature when we apply an external pulling force along the $x-$direction. The hairpin maitains its bound state for lower value of forces. And for a critical value of force the hairpin shows transition from bound to unbound state. The order of this transition has been supported by two evidences. First evidence is the scaling of average end-to-end separation shown in 3.3(b),

which scales as $\langle x \rangle \sim N^d \mathcal{F}(|g - g_c|N^\phi)$ for critical force $g_c = 0.39 \pm 0.01$ and exponents $\phi = 1.0 \pm 0.05$ and $d = 0.95 \pm 0.05$. From simple calculation it leads to, $\langle x \rangle \sim \frac{1}{|\ g - g_c|}$. So, there is discontinuity in the order parameter $\langle x \rangle$ as it approaches the critical value of force. And the second evidence is the thermal extensibilty χ which shows a peak at the critical value of force g_c. This scales as as $\chi \sim N^{d_\chi} \mathcal{F}(|g - g_c|N^\phi)$ with $g_c = 0.39 \pm 0.01$ and exponents $\phi_\chi = 1.0 \pm 0.05$ and $d_\chi = 2.05 \pm 0.05$ shown in Fig 3.4(b), which gives $\chi \sim \frac{1}{|\ g - g_c|}$. So, there is a discontinuity in response function also at $g = g_c$. Both the evidences have finally concluded that the force-induced unzipping of a polymer-hairpin is a first order phase transition.

Chapter 4

Summary

Molecular dynamics study of linear polymer and polymer hairpin has been carried out in this dissertation using large scale atomic/molecular massively parallel simulator. The structural properties of linear polymer are determined using coarse-grained bead spring model in the presence of a Langevin heat bath for various chain lengths upto $N = 256$. The longer chain lengths showed an ideal chain behaviour for the fundamental result (Eq 1.12). While the shorter chain, $N = 16$ have a deviation. This also suggets us that longer chainlengths give us more accurate results. The MSID for different chainlengths saturated at their correct end-to-end distance, which confirms the equilibration of the particular chain length. The size exonent obtained from the relation $\langle R_N^2 \rangle \sim N^{2\nu}$ is 0.63 ± 0.02. This value is slightly greater than $\nu \approx 0.59$, obtained in experiments. While the calculation of structure factor leads us to the the value $(= 0.59 \pm 0.01)$ which is in good agreement with the experimentally obtained value.

We also studied the unzipping of a polymer hairpin via two methods: thermal denaturation and force-induced unzipping. The simulations were carried out upto chainlength $N = 64$ with Langevin heat bath, keeping the one end anchored. The polymer hairpin maintains its zipped or bound state untill it reaches a temperature value, T_M. For $T > T_M$, it is always in open configuartion. While going from zipped to open configuration, it shows a transition. The end-to-end separation is the order parameter of the problem. From our estimation we found that, $\langle x \rangle \sim |T - T_M|^2$. So the order parameter is continuous as it approaches the melting temperature. This concluded that the behaviour of thermal denaturation of polymer hairpin is a continuous phase transition.

However in the case of force-induced unzipping, the polymer hairpin shows a first

order phase transtition on going from hairpin to open configuration at a critical value of force $g = g_c$. Below g_c, the hairpin is always zipped (bound state) and above g_c it is in open configuration. The order parameter, end-to-end separation and the response function, isothermal extensibilty both are found to be have a discontinuity at $g = g_c$, which makes sure that it is a first order phase phase transition. We only studied finite chain lengths, but if we go to higher chain lengths we can get the behaviour of average end-to-end separation with the force as a step function. In the limit of infinite chain length χ shows a discontinuity at g_c.

Appendix A

A.1 Equilibration of polymer chains

```
# Initialization
dimension 3
units lj
atom_style bond
special_bonds fene

# read data from initial configuration file
read_data chain16.data
mass * 1

neighbor 2 bin
neigh_modify every 1 delay 0 check yes

bond_style fene
bond_coeff * 30.0 1.5 1.0 1.0

# start with a pair-wise soft potential and allow the
# system to equilibrate. We use fix/limit so the system
# doesn't explode during the equilibration run

pair_style soft 1.0
pair_coeff * * 30.0

# create initial velocities at T=3.0
```

```
velocity all create 3.0 17786140 dist gaussian

fix 1 all nve/limit 0.01
run 1000
fix 1 all nve/limit 0.05
run 1000
fix 1 all nve/limit 0.1
run 500
unfix 1

# redefine the pair-wise potential to L-J repulsive potential
# with cut-off r_c = 2^{1/6}
pair_style lj/cut 1.12
pair_modify shift yes
pair_coeff  * * 1.0 1.0 1.12

# re-create initial velocities at T=3.0
velocity all create 3.0 1778663

fix 1 all nve
fix integrator all langevin 3.0 3.0 0.5 904297

thermo   100
timestep 0.008
run 10000000

# write equilibrate chain data
write_data equil.dat
```

A.2 Java code for hairpin-configuration

```
import java.lang.*;
import java.util.*;
import java.io.*;
public class config{
```

```
    public static void main(String[] args){
BufferedWriter output = null;
try{
PrintStream ps=new PrintStream(new FileOutputStream("data.txt"));
System.setOut(ps);
}
catch ( IOException e ) {
            e.printStackTrace();
        }
int N = 48;
System.out.println("" + "LAMMPS DESCRIPTION");
System.out.println("");
System.out.println("" + N + " " + "atoms" );
System.out.println("" + (N-1) + " " + "bonds");
System.out.println("");
System.out.println("" + N + " " + "atom types" );
System.out.println("" + 1 + " " + "bond types");
System.out.println("");
System.out.println("" + 0.0 + " "+ (N*2) + " "+ "xlo" + " "+ "xhi");
System.out.println("" + 0.0 + " "+ (N*2) + " "+ "ylo" + " "+ "yhi");
System.out.println("" + 0.0 + " "+ (N*2) + " "+ "zlo" + " "+ "zhi");
System.out.println("");
System.out.println("" + "Atoms # id mol type xu yu zu");
System.out.println("");
/*** Specify the y-coordinate for first and last atom ***/
double y = N/4 ;
for(int i=1; i<=N; i++){

if(i<=N/2){
System.out.println("" + i + " "+ 1 +  " " + i + " " + 4.0 + " " + y + " " + 0.0);
}
if(i>N/2){
System.out.println("" + i + " "+ 1 +  " " + i + " " + 5.12 + " " + y + " " + 0.0);
}
```

```
if(i<(N/2)){
y = y+1;
}
if(i>(N/2)){
y = y-1;
}
}
System.out.println("");
System.out.println("" + "Bonds");
System.out.println("");

for(int j=1; j<N; j++){
System.out.println( "" + j + " " + 1+ " "+ j+ " " + (j+1) );
}
    }
}
```

A.3 Polymer-hairpin under tension

```
# define temperature and force for the simulation
variable temp equal 0.1
variable maxforce equal 0.0

dimension 2
units lj
atom_style molecular
boundary p p p
neighbor 1 bin
neigh_modify every 1 delay 1

read_data data16.txt
# all monomers have mass 1
mass * 1

# Type 2 to 16 are mobile , 1 is fixed, and 16 is free
```

```
group mobile type 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
group fixed type 1
group end_monomer type 16

#pair interactions
pair_style lj/cut 1.12
# repulsive L-J potential
pair_modify shift yes
pair_coeff * * 1.0 1.0 1.12
# attractive L-J potential
pair_modify shift no
pair_coeff 1 16 1.0 1.0 2.5
pair_coeff 2 15 1.0 1.0 2.5
pair_coeff 3 14 1.0 1.0 2.5
pair_coeff 4 13 1.0 1.0 2.5
pair_coeff 5 12 1.0 1.0 2.5
pair_coeff 6 11 1.0 1.0 2.5
pair_coeff 7 10 1.0 1.0 2.5

#FENE type bond
bond_style fene
bond_coeff * 100.0 1.5 1.0 1.0
special_bonds fene

variable dt equal 0.01
timestep ${dt}
thermo 10000

fix integrator mobile nve
fix dynamics mobile langevin ${temp} ${temp} 0.4 252111
fix run2d all enforce2d

#make pictures
shell "mkdir img"
shell "rm img/*"
```

```
dump img all  image  100000 img/t*.jpg    type type adiam 1.0 bond type 0.4 zoom 1
dump_modify img backcolor white  boxcolor black
dump_modify img pad 6

#Add an external force to the end monomer
variable extforce equal ${maxforce}
fix externforce end_monomer  addforce v_extforce 0 0

variable eqrun equal 1000000
variable samplerun equal 100000000

#first equilibrate
run ${eqrun}
write_data equil16.dat
reset_timestep 0

#distance from one end to the other along the x axis
variable dist equal (x[16]-x[1])
variable t equal step*dt
fix avg all ave/time 1 1 100 v_t v_dist v_extforce file dist_vs_force ave one

thermo 10000
thermo_style custom step v_dist v_extforce
run ${samplerun}
```

Bibliography

[1] M. Rubinstein, R. H. Colby, Polymer Physics, Oxford(2003).

[2] JP Cotton, Journal de Physique Lettres **41**, 231–234(1980).

[3] M. P. Allen and D. J. Tildesley, Computer Simulations in Chemical Physics, Springer Science, 397–459(2012).

[4] C. Bustamante, S. B. Smith, J. Liphardt and D. Smith, Curr. Opin. Struct. Biol. **10**, 279(2000).

[5] G. S. Grest and K. Kremer, Phys. Rev. A **33**, 3628(1986).

[6] SJ Plimpton, LAMMPS Users Manual: Large-scale Atomic/Molecular Massively Parallel Simulator, Sandia Corporation: Sandia National Laboratories(2003).

[7] M. Perez, O. Lame, F. Leonforte and J-L Barrat, The Journal of chemical physics **128**, 234904(2008).

[8] L. H. Wong and A. L. Owczarek, J. Phys. A: Math. Gen. **36**, 9635–9646(2003).

[9] R. Kapri, S. M. Bhattacharjee and F. Seno, Phys. Rev. Lett. **93**, 248102(2004).

[10] R. Kapri, Phys. Rev. E **90**, 062719(2014).

[11] S. Kumar, R. Kumar and J. Wolfhard, Phys. Rev. E **93**, 010402(2016).

[12] S. M. Bhattacharjee, J. Phys. A **33**, L423(2000).

[13] D. Marenduzzo, A. Trovoto and A. Maritan, Phys. Rev. E **64**, 031901(2001).

[14] F. Ritort, J.Phys: Condens. Matter **18**, R531(2006).

[15] M. Doi, S. F. Edwards, The theory of polymer Dynamics, Oxford(1986).