Fermeture de bulles de dénaturation de l'ADN couplé à l'élasticité de l'ADN

Anil K. Dasanna

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par *l'Université Toulouse III – Paul Sabatier*
Discipline ou spécialité : *Physique de la matière*

Présentée et soutenue par Anil Kumar DASANNA
Le *30 Septembre 2013*

Titre : *Fermeture de bulles de dénaturation de l'ADN couplées à l'élasticité de l'ADN*

Title : *Closure of DNA denaturation bubbles coupled to DNA elasticity*

JURY

M^{me}	Maria BARBI	LPTMC, Paris	Rapporteur
M.	Nicolas DESTAINVILLE	LPT, Toulouse	Co-Directeur de thèse
M.	Richard FOURNIER	LPCE, Toulouse	Examinateur
M.	Marc JOYEUX	LSP, Grenoble	Rapporteur
M.	Manoel MANGHI	LPT, Toulouse	Directeur de thèse
M.	John PALMERI	LCC, Montpellier	Invité
M.	Andrea PARMEGGIANI	LCC, Montpellier	Examinateur

École doctorale : *Sciences de la matière*
Unité de recherche : *Laboratoire de Physique Théorique*
Directeurs de thèse : Manoel MANGHI **et** Nicolas DESTAINVILLE

Acknowledgements

This thesis is carried out at Laboratoire de Physique Théorique, Toulouse. First, I would like to thank to my supervisors, Prof. Manoel MANGHI and Prof. Nicolas DESTAINVILLE, from whom I have learnt the fundamentals in DNA biophysics. They had time to discuss whenever I knocked their doors. After the end of my PhD, I am now equipped with good skills in numerical simulations, for which all the credit goes to them. Their guidance through out my PhD is inevitable. Even though they had teaching duty, they were managed to give me enough time (more than enough). They made me to carry out the thesis topic step-by-step, to make it complete and concrete. In a simple terms, they were just amazing. Finally, I was very lucky to have them as my PhD advisors. Another person that I would like to thank is Prof. John PALMERI with whom all my PhD work is collaborated. Even though he was quite busy, I had very fruitful discussions with him.

I would like to thank all other members of LPT for making my stay very pleasant. Last but not least, I would like to thank my fellow PhD students and post-docs, Yasir, Phillip, Xavier, Vivek, Juan, David, Sumiran and Peter, with whom I had great company through out my PhD. I am also obliged to thank Sandrine for helping me all the time with computer related stuff.

Not to forget, I would like to thank all the Jury for accepting to read my thesis, especially the referees Prof. Maria Barbi and Prof. Marc Joyeux, and the director Prof. Andrea Parmeggiani. They carefully read my thesis by spending their valuable time and their comments improved my thesis considerably.

Finally, I would like to thank Université Paul Sabatier for providing the funding for my PhD.

Summary

Understanding the dynamics of biological processes such as transcription, translation or DNA duplication requires the knowledge of double stranded DNA (dsDNA) physics. The dynamics of dsDNA had mainly been tackled at two different scales: (1) at the macromolecule scale where the internal structure is ignored; (2) at the base-pair (bp) scale focusing on the dynamics of closing or opening bubbles of denaturation (segment of open base-pairs). Such studies do not consider the closure of thermalized or pre-equilibrated denaturation bubble which couples both internal degrees of freedom (base-pairing) and conformal degrees of freedom (chain's elastic degrees of freedom). Such a bubble also occurs during the final stage of transcription when RNA polymerase leaves the locally opened DNA. Altan-Bonnet *et al.* measured the closure times of denaturation bubbles of size 18 bps and found very long times in $20 - 100$ μs range at room temperature. Bubble lifetimes of about 1 μs have also been observed for DNA oligomers of size 14 bps in NMR measurements of imino protein exchange. In an attempt to explain the physical mechanisms behind these large timescales, we focused in this thesis on the closure of thermalized denaturation bubble using simple coarse-grained models.

In a first attempt, we considered the "ladder" coarse-grained model consisting of two semi-flexible interacting single strands (ssDNA) and a bending modulus which depends on the base-pair state (thus mimicking the effect of stacking interactions inside the double helix), with dsDNA segments being 50 times stiffer than ssDNA ones. We focused on closure dynamics of thermalized denaturation bubble using Brownian dynamics simulations and analytical arguments.

The closure of bubble occurs via two steps: (1) a fast zipping followed by (2) a metastable state. Fast zipping process is characterized by an anomalous exponent for the bubble size, $L(t)$, given by $L(t = 0) - L(t) \sim t^{1/1.4}$, and the zipping time. This has been already observed in polymer translocation and adsorption problems. When the size of the bubble

decreases, bending energy is stored inside the bubble, which eventually stops zipping, leading to the metastable state. The size of the metastable bubble is around 10 bps. Since the zipping is faster than diffusion of smaller arms (segments of dsDNA on each side of the bubble), the bubble closes by bringing the two arms closer enough until it falls into an "hairpin configuration". To relax the bending energy inside the bubble, arms need to be aligned. The final closure occurs (almost instantaneously) when both dsDNA arms are aligned via rotational diffusion, which sets the dominant timescale in the process of closure.

Total closure times of 0.1 to 4 μs are found for lengths of DNA $N = 20 - 100$ bps, that scale as $\tau_{\rm cl} \simeq N^{2.4}$. We thus have a good qualitative (because of the occurrence of a metastable state) agreement with experiments, but no quantitative matching.

To go further, we improved this model by allowing the two strands to inter-wind to form a double helix in the dsDNA molecules and thus adding an additional degree of freedom, the twist. We simulated DNA of lengths $N = 40 - 100$ bps with twist rigidity, $\kappa_\phi = 200 - 300\ k_{\rm B}T$ (corresponding to torsional rigidity on the order of 3×10^{-19} J.nm) using Brownian Dynamics. We showed that twist dynamics plays a key role in the closure of pre-equilibrated bubbles, which, here again, occurs in two steps. First, the large flexible bubble quickly winds from both ends (zipping regime), thus storing torsional energy in the bubble, which stops when it reaches a size ~ 10 bps. The closure of this metastable bubble depends both on the twist rigidity, κ_ϕ, and the length of DNA, N. For low κ_ϕ, an arms diffusion limited (ADL) regime is observed, as in the earlier ladder model, where the closure is limited by the diffusive alignment of the two dsDNA arms. For large κ_ϕ and not too large N, the bubble diffuses along the DNA and closes as soon as it reaches one DNA end (bubble diffusion limited (BDL) closure), with a closure time $\tau_{cl} \simeq N^{2.3}$ and for clamped ends (or long DNAs), the closure is temperature activated (TA), associated with crossing an energy barrier due to non-zero twist in the bubble. We constructed a 'phase' diagram based on the mean values of three closure time distributions in the plane of κ_ϕ and N. For clamped DNAs, the BDL region is replaced by a TA region. Total closure times of 0.1 to 100 μs are found with this model.

We thus found that chain bending and twisting properties of DNA plays an important role in the closure of denaturation bubbles. We recovered the experimental timescales, thus indirectly inferring the value $\kappa_\phi \simeq 290\ k_{\rm B}T$.

Résumé

Comprendre la dynamique des processus biologiques tels que la transcription, la translation ou la copie de l'ADN, nécessite d'approfondir la compréhension des phénomènes physiques qui régissent la dynamique de l'ADN double-brin (dsDNA). Celle-ci a jusqu'ici été étudiée à deux échelles bien distinctes: (1) à l'échelle de la macromolécule, qui ignore la structure moléculaire de l'ADN, et (2) à l'échelle de la paire de bases (bp), en se concentrant sur la dynamique d'ouverture et fermeture des bulles de dénaturation (segment de paires de bases ouvertes). En revanche ces études ne portent pas sur la fermeture de larges bulles de dénaturation pré-équilibrées, qui met en jeu un couplage entre les degrés de liberté internes d'appariement des paires de bases et ceux de configuration (ou d'élasticité) de la chaîne. Ces bulles apparaissent pourtant pendant la dernière étape de la transcription, lorsque l'ARN polymérase se dissocie de la région localement ouverte de l'ADN. Altan-Bonnet *et al.* ont mesuré, à température ambiante, des temps de fermeture de bulles de dénaturation de 18 bp de l'ordre de 20 à 100 μs. Des temps de vie de 1 μs ont également été mesurés pour des oligomères d'ADN de 14 bp dans des expériences de RMN. Dans ce travail de thèse, nous avons étudié, à l'aide de modèles numériques "gros grains", la fermeture de grandes bulles de dénaturation pré-équilibrées, afin de comprendre les mécanismes physiques cachés derrière ces longs temps de fermeture.

Dans un 1er temps, nous avons développé un modèle d'ADN en "échelle", composé de deux simples brins d'ADN (ssDNA) semi-flexibles en interaction avec un module de courbure qui dépend de l'état d'appariement des paires de bases (modélisant ainsi les interactions d'empilement dans la double hélice). Ce dernier est effectivement 50 fois plus grand pour l'ADN double-brin que simple-brin. A l'aide de simulations numériques de dynamique brownienne et d'arguments analytiques, nous avons étudié la dynamique de fermeture d'une grande bulle de dénaturation pré-équilibrée.

La fermeture se fait en deux étapes : (1) une fermeture rapide façon fermeture-éclair (ou "zipping"), suivie (2) d'un état métastable. Le zipping rapide est caractérisé par un

exposant anormal pour la taille de la bulle $L(t)$, selon $L(t=0) - L(t) \sim t^{1/1.4}$, et le temps de zipping. Cette loi a déjà été observée dans des problèmes de translocation ou d'adsorption de polymères. Lorsque la taille de la bulle diminue, de l'énergie de courbure est emmagasinée dans la bulle, ce qui, à un certain moment, arrête le zipping et conduit à une bulle métastable d'environ 10 bp. Comme le zipping est plus rapide que la diffusion des "bras" (segments de dsDNA flanquant la bulle des deux côtés), la fermeture de la bulle se fait en rapprochant les deux bras jusqu'à ce que l'ADN adopte une conformation en épingle à cheveux. La relaxation de l'énergie de courbure emmagsinée dans la bulle se fait alors en alignant les bras par diffusion rotationnelle. La fermeture définitive de la bulle se fait de façon quasi instantanée dès que ceux-ci sont alignés. Ce processus est le facteur limitant de la fermeture.

Pour des ADN de longueur $N = 20$ à 100 bp, on mesure ainsi des temps de fermeture complète de l'ordre de 0.1 à 4 μs, avec une loi d'échelle en $\tau_{\rm cl} \simeq N^{2.4}$. Grâce à l'apparition d'un état métastable, nous apportons donc une explication qualitative aux longs temps de fermeture mesurés expérimentalement mais pas un accord quantitatif.

Nous avons donc amélioré notre modèle en tenant compte de la géométrie hélicoîdale de l'ADN et en considérant donc un degré de liberté supplémentaire, la torsion. Nous avons simulé par dynamique brownienne des ADN de 40 à 100 bp avec un module de torsion $\kappa_\phi = 200 - 300\ k_{\rm B}T$ (correspondant à une rigidité de torsion de 3×10^{-19} J.nm). Nous avons montré que la dynamique de la torsion joue un rôle primordial dans la fermeture des bulles de dénaturation, qui là encore, se fait en deux étapes. D'abord, la grande bulle flexible se vrille par les deux bouts (régime zipping), emmagasinant de l'énergie de torsion, puis s'arrête lorsqu'elle atteint la taille d'environ 10 bp. La fermeture de cette bulle métastable se fait alors suivant trois scenarii distincts selon la valeur de κ_ϕ: Pour de faibles κ_ϕ, la fermeture est, comme dans le cas du modèle en échelle, limitée par la diffusion des bras jusqu'à ce qu'ils s'alignent (régime ADL). Pour de grands κ_ϕ et des N intermédiaires, la bulle diffuse le long de la chaîne et se ferme dès qu'elle atteint l'un des deux bouts (fermeture limitée par la diffusion de la bulle, régime BDL) conduisant à un temps de fermeture en $\tau_{\rm cl} \simeq N^{2.3}$. Pour des ADN "pinçés" aux deux bouts (ou de très longs ADN), la fermeture se fait selon un processus activé thermiquement (régime TA) et par le franchissement d'une barrière d'énergie de torsion. Nous avons construit un "diagramme des phases" dans le plan (N, κ_ϕ) à partir des moyennes des trois distributions de temps de fermeture associées aux trois scenarii. Pour des ADN pinçés aux deux bouts,

la région BDL pour les grands κ_ϕ est remplaçées par une région TA. Avec ce modèle, on mesure des temps de fermeture de 0.1 à 100 μs.

Nous montrons ainsi que la courbure et la torsion de l'ADN joue un rôle primordial dans la fermeture des bulles de dénaturation. Nous retombons sur des temps de fermeture en accord avec ceux mesurés expériementalement, ce qui permet une détermination de $\kappa_\phi \simeq 290\, k_\mathrm{B}T$.

Publications

[1] **Anil K.Dasanna**, Nicolas Destainville, John Palmeri and Manoel Manghi : *Strand diffusion-limited closure of denaturation bubbles in DNA* , Europhysics Letters, **98**, (2012) 38002.

[2] **Anil K.Dasanna**, Nicolas Destainville, John Palmeri and Manoel Manghi : *Slow closure of denaturation bubbles in DNA: twist matters*, Phys.Rev.E, **87**, (2013) 052703.

Contents

Chapter 1

Introduction

In this chapter, we present the DNA structure and its functional importance. Later, we describe the DNA denaturation and introduce the models that account for the denaturation dynamics. Then we present the recent experiments that studied the dynamics of denaturation bubbles.

1.1 Structure of DNA

Cells are building blocks of life. Cells contain a vast number of biomolecules that participate in different regulatory mechanisms. Two such important biomolecules are nucleic acids and proteins. Most of the functions inside the cell are done by proteins such as catalyzing chemical reactions, cell signaling, and many other processes. Deoxyribonucleic acid (DNA) is a such important and fascinating biomolecule. The main function of DNA or other nucleic acids, such as Ribonucleic acid (RNA), is to store the genetic information, which will be necessary in the formation of proteins. This is called the "Central Dogma of Molecular biology", as shown in Fig. 1.1.

A double stranded DNA (dsDNA) [1] is composed of two polymers, which are also called single stranded DNAs (ssDNA). Each ssDNA is a polymer of subunits called "nucleotides". A nucleotide is made up of three chemical units: the phosphate group (PO_4^-), the sugar group (deoxyribose) and the base group. A phosphate group in each strand is connected to two sugar groups and each sugar group is connected to one base and it repeats, as shown in Fig. 1.2a. There are four different kinds of bases: adenine (A), cytosine (C), guanine (G) and thymine (T). The bases attached to sugar groups of each strand are connected *via* hydrogen bonding interactions, thus forming a base-pair. The organization of DNA is as follows :

- The backbone structure of each ssDNA is made up of phosphate groups and sugars.
- The complementary bases of each ssDNA are connected *via* hydrogen bonding.
- Two ssDNAs have polarity called 3′-end and 5′-end, both run anti-parallel to each other.

The hydrogen bonding interactions take place only between A and T, G and C. Hydrogen bonding between G and C is much stronger than that of A and T as G–C consists of 3 bonds and A–T consists of 2 bonds only.

The secondary structure of DNA has been discovered in 1953 by Watson and Crick [2, 3] and by Franklin et *al.* [4]. The spatial structure of DNA is a double helix where two

FIGURE 1.1: "Central Dogma of Molecular Biology" showing transcription and translation processes, allowing the fabrication of proteins.

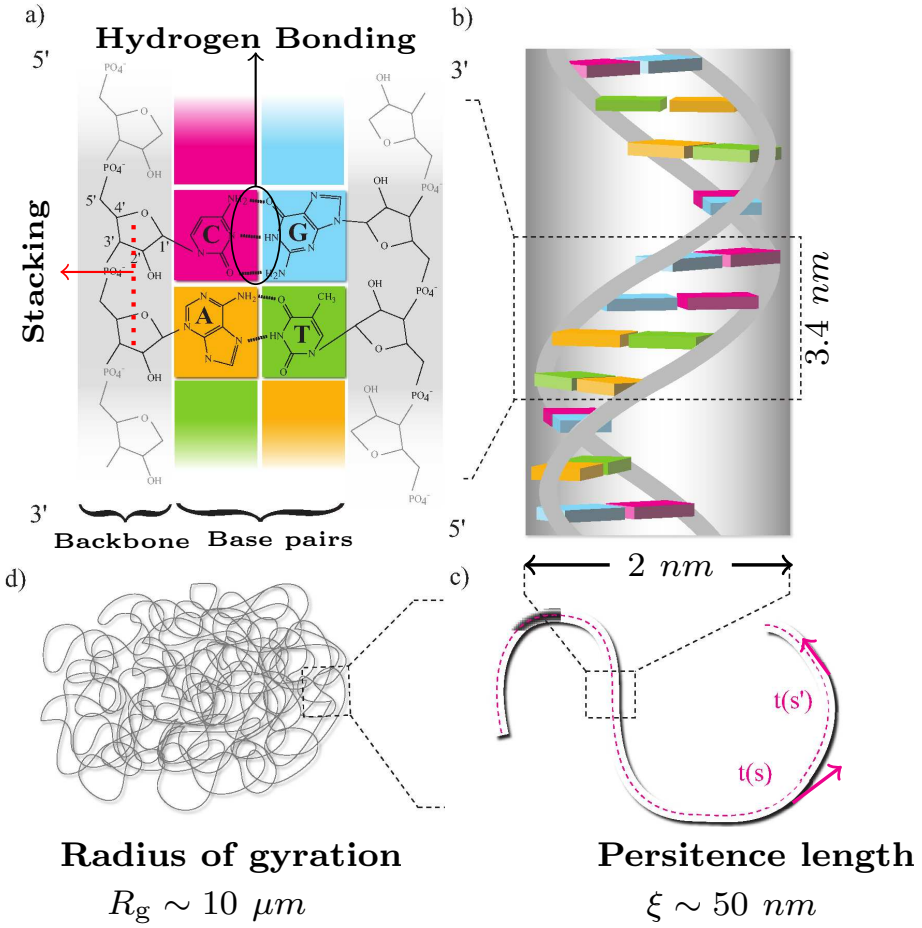

FIGURE 1.2: The structure of DNA is presented at different length scales. (a) The basic chemical composition of the primary structure of DNA is shown. The backbone of each strand is made up of phosphate and sugar groups. Four different bases, namely Adenine (A), Thymine (T), Guanine (G) and Cytosine (C) are attached to sugar groups. The bases of each ssDNA are connected *via* hydrogen bonding. (b) Secondary structure of DNA is shown. DNA is spatially organized as a double helix. Both strands are inter-wound around the same axis. The pitch of this double helix is 3.4 nm and the diameter is $\simeq$ 2 nm. (c) Tertiary structure of DNA. DNA at moderate length scales appears as a semi-flexible polymer with a persistence length, $\ell_{\rm ds} = 50$ nm. (d) DNA at longer length scales looks like a random coil of typical gyration radius of about 10 μs for a DNA of about 4 million base-pairs. Figure is taken from [1].

ssDNAs inter-wound around the helical axis as shown in Fig. 1.2b. The organization is such that the bases are inside and the backbones are outside the double helix. The geometrical parameters of this double helix are the pitch equal to 3.4 nm, and the diameter, equal to 2 nm. Along with these geometrical parameters, duplex DNA is not symmetrically wound. It has two grooves, a major groove and a minor groove which have unequal lengths. Even though the bases are projected inside the structure, they can be accessible from these grooves. Since the length of the major groove is around 2.2 nm and that of the minor one is 1.2 nm [5], this makes the bases more accessible to some proteins

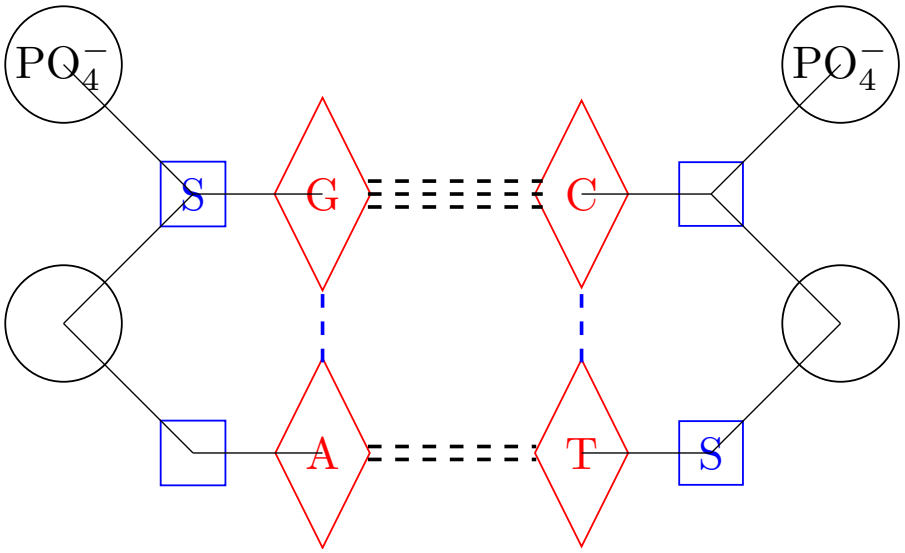

FIGURE 1.3: Hydrogen bonding interactions (black dotted lines) between Watson-Crick base-pairs and stacking interactions (blue dotted lines) between base-pair stacks are shown.

to bind the specific sites in DNA at major grooves during transcription [6]. The stability of the double helical DNA arises from two contributions :

- The hydrogen bonding interactions between Watson-Crick base-pairs.
- The stacking interactions between two consecutive base-pairs.

The first contribution is hydrogen bonding interactions between the complementary bases of DNA. Even though these hydrogen bonds are weak compared to other covalent bonds, they keep dsDNA stable at physiological temperatures. Quantum chemical calculations show that the hydrogen bonding energy of an A-T base-pair is 7 kcal mol^{-1} (11.76 $k_{\mathrm{B}}T_0$) and that of a G-C base-pair is 17 kcal mol^{-1} (28 $k_{\mathrm{B}}T_0$) [7], where T_0 is room temperature. The stacking interactions are mainly coming from the $\pi - \pi$ interactions between the aromatics rings of the base along each ssDNA as sketched in Fig. 1.3. As hydrogen bonding interactions, the stacking interactions also depend on the sequence. These stacking interactions have a minimum energy at a length almost equal to 0.34 nm, thus giving stabilized duplex DNA of pitch $\simeq$ 3.4 nm. The stacking energies between A-T and G-C base-pairs are $E^{\text{A-T}}_{\text{A-T}} = 14.0$ kcal mol^{-1} and $E^{\text{G-C}}_{\text{G-C}} = 17.00$ kcal mol^{-1} [8]. As the hydrogen bonding and stacking interactions depend on the sequence, the thermal stability of dsDNA depends on the sequence.

In eukaryotic cells, DNA is mainly localized inside the nucleus in the form of chromosomes with the help of proteins, the histones. DNA occupies nearly 10% of the total

volume of the cell. The typical diameter of nucleus is around 6 μm and the total length of DNA in human cell is around 2 m. It implies that DNA must be packed tightly inside the nucleus. The flexibility of DNA at moderate length scales can be described by its persistence length which enters in the Worm Like Chain (WLC) polymer model [9, 10] as shown in Fig. 1.2c. Persistence length is defined as the length over which the orientational vectors decorrelates, $\hat{t}(0) \cdot \hat{t}(s) = e^{-s/\ell_{\rm p}}$, where the orientational vectors, $\hat{t}(0)$ and $\hat{t}(s)$ are the tangent vectors of the mean dsDNA chain. If we take a polymer chain of length, $L >> \ell_{\rm p}$, the polymer chain looks like a coil as shown in Fig. 1.2d. The persistence length of DNA has been measured and is roughly 50 nm. It implies that DNAs of length much longer than the persistence length occupy a coil configuration whereas short DNAs of length comparable to its persistence length are more rigid.

1.2 DNA functions

The main functions of DNA include transcription & translation (Central Dogma of Molecular Biology, see below) and DNA replication during cell division. Along with these functions, DNA also interacts with proteins during its life cycle. DNA can be seen as the bank of genetic information. All the genetic information required in the process of protein synthesis is stored in DNA. The information is stored in the form of genes. Genes are the functional parts of the DNA, usually of length few hundreds to few thousand base-pairs. The functional genes are usually separated by repetitive parts of AT and GC.

Central Dogma of Molecular Biology [11]

Transcription and translation is a two step process as shown Fig. 1.1. The organization of a gene consists of many specific regions called promoter, coding region, regulatory regions and termination regions. All these regions have specific tasks during the stage of transcription. The genetic information of a gene is carried out to the functional protein by messenger RNA (mRNA), produced by the RNA polymerase. Transcription works mainly in three major steps :

- Initiation
- Elongation
- Termination

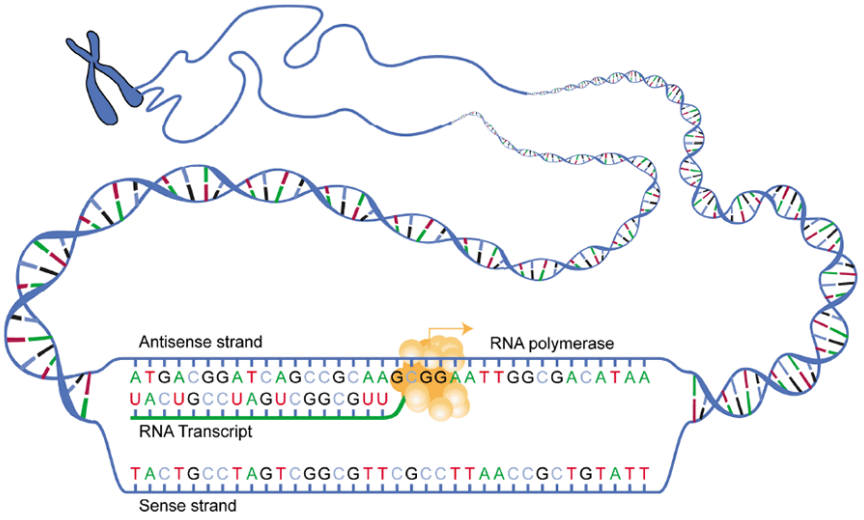

FIGURE 1.4: Sketch of RNA polymerase decoding the sequence and leaving behind the RNA transcript[1]. The different scales, from the chromosome to the base-pairs, are schematized.

The initiation step corresponds to the identification by the RNA polymerase of the promoter region of the gene and its binding to it. Once the initiation step is performed, subsequently elongation step starts. The RNA polymerase reads the sequence and moves forward. When it is moving forward after reading the sequence, it leaves behind the RNA transcript. When RNA polymerase reaches the termination region, it leaves the DNA. The translation step which is also a part of gene expression, includes decoding of mRNA which is produced during the stage of transcription. Ribosomes found inside the cytoplasm, decode the genetic information in mRNA into a specific chain of amino acids. This specific chain of amino acids folds, and thus becomes an active protein.

DNA replication [11]

Cell division is at the heart of inheritance. Before starting the cell division, DNA has to make its own copy. Cell division starts at the specific position at which the proteins break the hydrogen bonds between the bases of complementary strands, thereby creating the replication fork as shown in Fig. 1.5. Apart from helicase, DNA polymerase will assist in making the copies of the two strands, thus forming two copies of the parent DNA.

From a physical perspective, both the transcription and replication are very striking and fascinating. Studying DNA physics leads to better understanding of these biological processes.

[1]Source: http://en.wikibooks.org/wiki/Structural_Biochemistry/RNA_Polymerase_II
[2]Source: http://images.nigms.nih.gov/imageRepository/2543/DNA_Replication.jpg

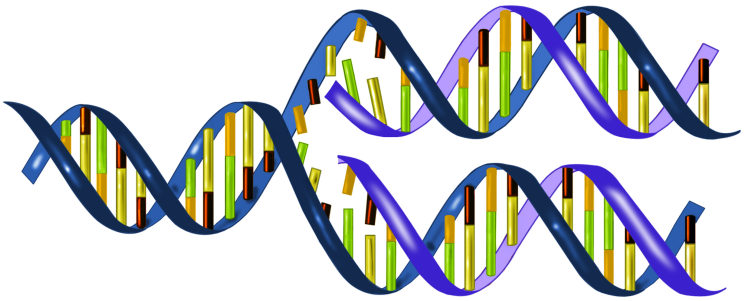

FIGURE 1.5: Sketch of DNA replication, showing the replication fork and the lagging two copies of the DNA[2].

1.3 DNA Denaturation

DNA in its double helix configuration is stable at physiological temperatures inside the nucleus of the cell. As already said, the stability of the double helix is provided by both hydrogen-bonding and stacking interactions of base-pairs. To study the DNA physics, experiments are done in solution. As one increases the temperature, these interactions/bonds start breaking. At a critical temperature, DNA completely opens up into two ssDNAs, which is also called "melting". DNA denaturation can be characterized by its "melting temperature" at which 50% of the DNA is completely broken and rest of the DNA is closed. This melting temperature is denoted by T_{m}. As the strength of hydrogen-bonding and stacking depends on the sequence, the melting temperature also depends on the sequence of DNA.

Experimentally, the thermal denaturation of DNA is studied using UV absorption technique. The fraction of open base-pairs results in an increase of UV absorption at around 260 nm. Using this technique, it was found that homopolymer DNAs denature in a single interval of temperature, whereas heteropolymer DNAs denature in multiple steps [12]. As an example for heteropolymer DNA, the typical melting profile of pNT1 DNA plasmid (which is cut to make it linear) is shown in Fig. 1.6. The melting temperature is $T_{\mathrm{m}} \simeq 75°C$. The parameter, $1-\theta(T)$ denotes the fraction of open base-pairs of DNA in aqueous solution, whereas $-\frac{\mathrm{d}\theta}{\mathrm{d}T}$ denotes the differential melting profile. The fraction of open base-pairs, $1-\theta(T)$ is computed as

$$1-\theta(t) = \frac{A(T)-A_{\mathrm{n}}}{A_{\mathrm{d}}-A_{\mathrm{n}}} \tag{1.1}$$

where $A(T)$, A_{n} and A_{d} are UV absorbance of DNA solution, UV absorbance of DNA in double stranded state and UV absorbance in single stranded state respectively. The

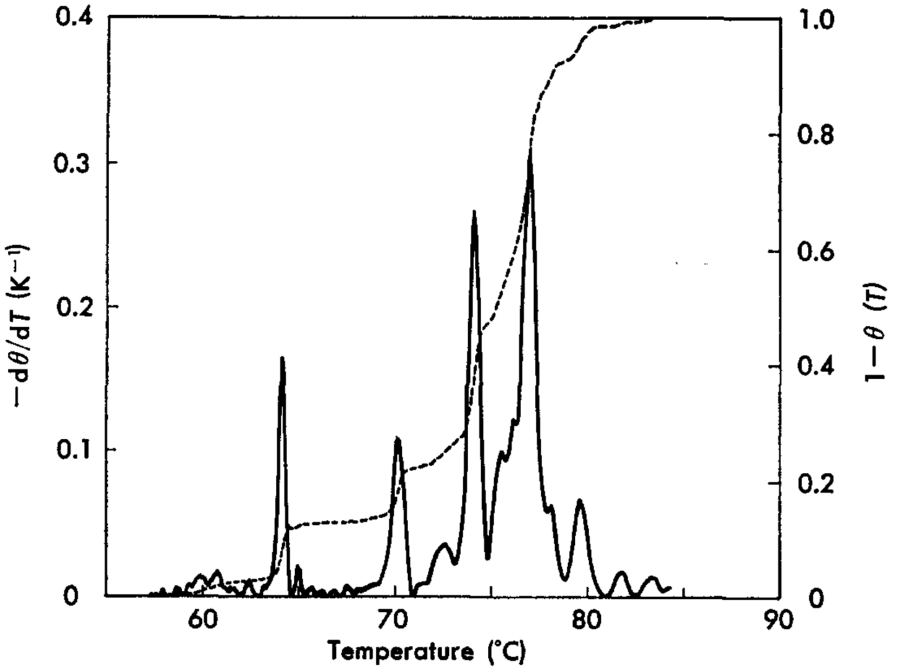

FIGURE 1.6: Melting profile of pNT1 DNA, showing the cumulative profile (dotted line) and differential melting profile (solid line). Taken from [15].

melting profile for homopolymer DNA is usually very sharp. The melting profiles for two homopolymer DNAs are shown in Fig. 1.7, which seem to be relatively sharper than that of heteropolymer DNA shown in Fig. 1.6. Apart from UV absorption technique, DNA denaturation can be studied using Fluorescence Correlation Spectroscopy (FCS) [13]. In FCS, a base-pair is tagged with a fluorophore and a quencher as shown in Fig. 1.14 below. When the base-pair is closed, the fluorophore is in proximity of the quencher, suppressing the fluorescence. When the base-pair is open, the fluorescence is restored, thus leaving the opening and closing of a base-pair into fluctuations in fluorescence amplitude [14]. Melting map of DNA gives the probability of the fraction of open base-pairs for a given temperature. The complete denaturation of DNA takes place mainly in two pathways:

- The first pathway is the formation of local openings which are also called "bubbles".
- The other pathway is the opening at one end of the DNA.

Note that at room temperature, the probability for the formation of the bubble is almost zero.

Understanding DNA denaturation helps in understanding DNA transcription and replication as the formation of a bubble plays a crucial role. In the below sections, we study DNA denaturation from a physical point of view by reviewing models existing in the literature.

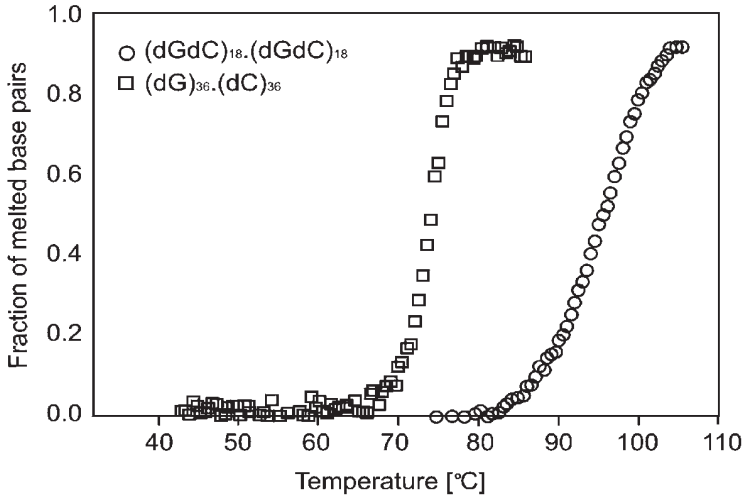

FIGURE 1.7: Normalized UV absorption melting curves for $(dG)_{36}{\cdot}(dC)_{36}$ with melting temperature of $74°C$ and $(dGdC)_{18}{\cdot}(dGdC)_{18}$ with melting temperature of $96°C$. Taken from [16].

1.4 DNA Denaturation: Equilibrium

1.4.1 Poland-Scheraga model

A simple model for DNA denaturation was first introduced by Poland and Scheraga in 1966 [17, 18]. The model is inspired by the Ising model for magnetic systems. It considers DNA as a succession of bound and unbound (bubbles) segments as shown in Fig. 1.8. It is a discrete model at the base-pair level. The model considers two states for a base-pair, either 1 or 0, which account for open and closed states. In this model, the bound or ds segments are approximated to be infinitely rigid, whereas unbound or ss segments are infinitely flexible. However, the dsDNA is only around 50 times stiffer than ssDNA which already points out one limitation of the PS model. As in the Ising model, the equilibrium weight for the closed base-pair is $w = e^{-\beta\epsilon}$, so the statistical weight for a bound segment of length ℓ_1 becomes, $w = e^{-\beta\epsilon\ell_1}$. The statistical weight for unbound segment, or bubble, of length ℓ_2 is given by

$$w = \frac{s^{\ell_2}}{\ell_2^c} \tag{1.2}$$

where s is a non-universal constant and c is the loop exponent which depends on the loop configuration. This statistical weight is an entropic contribution of the bubble. The nature of the single strands in the bubble are assumed to be polymeric. The origin of the statistical weight of the bubble shown in Eq. (1.2) is the entropic cost in closing a flexible polymer loop of length 2ℓ, where ℓ is the bubble size. Note that this weight is

true only for large bubbles and also the bubble size should be larger than the persistence length of single strand DNA. The type of phase transition depends on the value of the loop exponent, c

$$c = \begin{cases} 1 < c < 2 & \text{Continuous} \\ c > 2 & \text{Discontinuous} \end{cases} \tag{1.3}$$

For $c < 1$, there will be no crossover which means DNA is always closed (only if ϵ does not depend on temperature, T). The value of the loop exponent can be estimated using the random walk configurations and it turned out that for a phantom flexible loop, $c = \frac{d}{2}$, where d is the spatial dimension. So, for $d = 2$ and $d = 3$, the transition is always continuous. The model was further refined by Fisher by including excluded volume interactions in the loop [19, 20]. The value of c in this case becomes, $d\nu$, where ν is the Flory's exponent, which still makes the transition continuous in 2 and 3 dimensions. The excluded volume interactions between denatured loops and bound segments makes the transition first order [21, 22]. Note however that in practice, DNA molecules are always of finite length and studied in solution.

Poland-Scheraga model is also applied in studying bubble dynamics using the Fokker-Plank equation approach [23, 24], stochastic dynamics [25] or numerical simulations [26]. This model does not consider the origin of loop entropy cost explicitly, which actually has its origin in conformational degrees of freedom of the molecule.

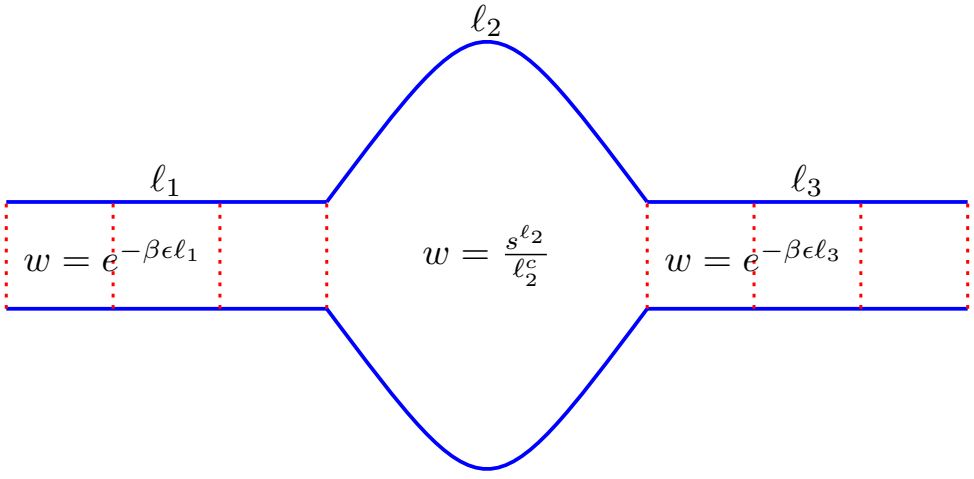

FIGURE 1.8: A sketch of a DNA with bound segments of length, ℓ_1, ℓ_3 and unbound segment of length, ℓ_2. Corresponding Poland-Scheraga statistical weights for bound and unbound segments are also shown.

1.4.2 Palmeri-Manghi-Destainville (PMD) model

Recently, Palmeri et *al.* [27, 28] came over this, by adding conformational degrees of freedom of the chain to base-pairing degrees of freedom. The effect of bending fluctuations of the chain is explicitly taken into account. The DNA of length N is completely described by a set of $2N$ variables, $\{\sigma_i, \mathbf{t}_i\}$, where σ_i is the Ising variable of i^{th} base-pair and $\mathbf{t}_i$ is the unit tangent vector of i^{th} base-pair which is given by spherical polar coordinates. The Hamiltonian becomes

$$\mathcal{H}[\sigma_i, \mathbf{t}_i] = \sum_{i=1}^{N-1} \kappa_{i+1,i}(1 - \mathbf{t}_{i+1} \cdot \mathbf{t}_i) - \sum_{i=1}^{N-1} \left[J\sigma_{i+1}\sigma_i + \frac{K}{2}(\sigma_{i+1} + \sigma_i) \right] - \mu \sum_{i=1}^{N} \sigma_i \qquad (1.4)$$

The first term in Eq. (1.4) is the bending energy of a discrete worm-like chain with local bending rigidity, $\kappa_{i,i+1}$, which depends on the base-pair state. The second and third terms in Eq. (1.4) belong to the Ising part which models base-pairing. The bending constant, $\kappa_{i+1,i}$ depends on the Ising variables, which accounts for the effect of stacking together with J. The model is solved using a transfer matrix approach. The model has also been extended by considering torsional degrees of freedom and loop entropy [29]. The model by PMD can be applied in studying bubble closure dynamics using kinetic Monte Carlo simulations which are consistent with the results of a simple coarse-grained model [30]. In the next section, we discuss DNA models that are more efficient in understanding the dynamics of DNA denaturation.

1.5 DNA Denaturation: Dynamics

1.5.1 Peyrard-Bishop model

This model was developed by Peyrard and Bishop in 1989 [31]. The model has also one degree of freedom per base-pair as in the PS model [17]. But unlike the PS model where the base-pair state, closed or open, is given by a discrete number, either 0 or 1, in this model, the base-pair state is given by a continuous variable, y. The Hamiltonian of this model is

$$\mathcal{H} = \sum_n \left(\frac{1}{2} m \dot{y}_n^2 + V(y_n) \right) + \sum_n W(y_n, y_{n+1}) \qquad (1.5)$$

where m is the reduced mass of the each base-pair. There are two potentials contributing to the Hamiltonian, $\mathcal{H}$. The first contribution, $V(y_n)$, which is the second term in

Eq. (1.5), is the hydrogen interaction between the bases of the two strands. The second contribution, $W(y_n, y_{n+1})$ to the Hamiltonian is the third term in Eq. (1.5) which mimics the stacking interaction between two consecutive base-pairs, as shown in Fig. 1.9. The hydrogen bonding interaction potential for the n^{th} base-pair is modeled by the Morse potential

$$V(y_n) = D\left(e^{-\alpha y_n} - 1\right)^2 \tag{1.6}$$

where α is the parameter which sets the inverse of the width of the potential and D is the depth of the potential. The Morse potential has two stable positions, one at $y = 0$, which corresponds to the closed state of a base-pair, and one at $y = \infty$, which corresponds to the open state. For short displacements in y from its equilibrium position, $y = 0$, the potential is similar to an harmonic one, whereas for larger deviations from $y = 0$, the potential saturates to zero, thus leading to non-linear bond breaking effects. The sharpness of the denaturation profile (refer to Fig. 1.7) is due to the entropic effects of the denatured part of DNA. To explain this sharpness of the denaturation profile, a non-linear stacking potential is considered [32]. The stacking interaction is

$$W(y_n, y_{n+1}) = \frac{K}{2}\left(1 + \rho e^{-\delta(y_{n+1}+y_n)}\right)(y_{n+1} - y_n)^2 \tag{1.7}$$

where the parameters, ρ and δ refer to anharmonic stacking interaction. When $\rho = 0$, the stacking interaction becomes harmonic. As the base-pair distance increases, the stacking interaction constant decreases from $\approx K(1+\rho)$ to K. whereas the harmonic approximation of stacking interaction is

$$W(y_n, y_{n+1}) = \frac{K}{2}(y_{n+1} - y_n)^2 \tag{1.8}$$

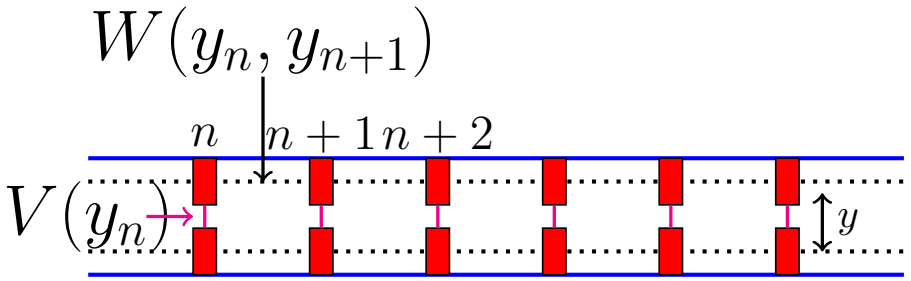

FIGURE 1.9: Sketch of model DNA showing the hydrogen interaction term and stacking term between the bases.

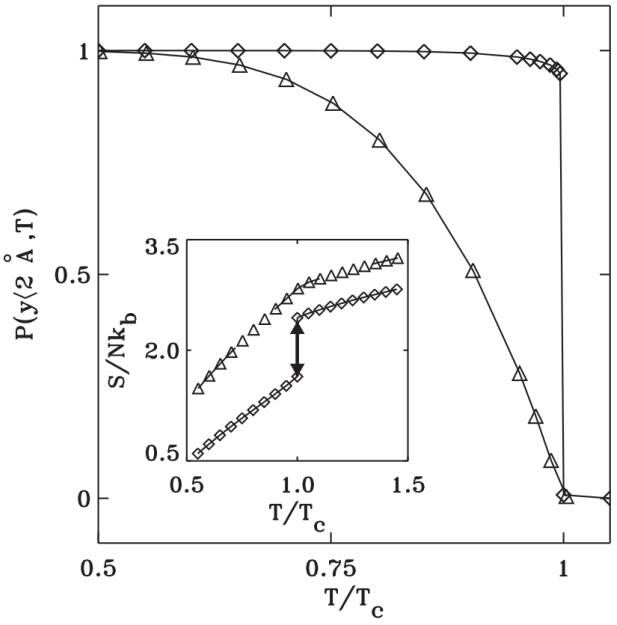

FIGURE 1.10: Theoretical prediction of fraction of closed base-pairs for both harmonic (△) and non-harmonic (⋄) stacking potentials against temperature, T. Inset: entropy plotted against temperature. Taken from [34].

The partition function for this model is written as

$$\mathcal{Z} = \int \prod_n \mathrm{d}p_n \mathrm{d}y_n e^{-\beta H(p_n, y_n)} \tag{1.9}$$

where $H(p_n, y_n)$ is given in Eq. (1.5) and $p_\mathrm{n} = m\dot{y}_\mathrm{n}$. This equation can be solved using transfer integral method and one arrives at [33]

$$\frac{-1}{2K}\frac{\mathrm{d}^2\phi(y)}{\mathrm{d}y^2} + \beta^2 V(y)\phi(y) = \beta^2\tilde{\epsilon}\phi(y) \tag{1.10}$$

with an average base-pair stretching

$$\langle y \rangle = \int_{-\infty}^{+\infty} |\phi_0(y)|^2 \mathrm{d}y \tag{1.11}$$

which is analogous to the Schrödinger equation for a particle in Morse potential. Note that Eq. (1.10) is a result of many approximations and has been derived for harmonic stacking potential. In Eq. (1.10), $\phi(y)$ is the eigenfunction of the transfer integral. To compute the fraction of closed or open base-pairs, a cut-off value for y is needed to define the open or close states. By choosing $y_0 = 2\mathring{A}$, the fraction of closed base-pairs are derived either analytically or by molecular dynamics simulations for both harmonic and non-harmonic stacking interactions, which is shown in Fig. 1.10. From Fig. 1.10, it

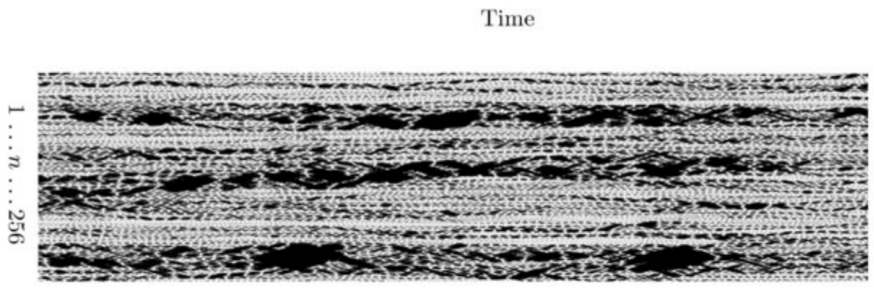

FIGURE 1.11: Molecular Dynamics simulation of model DNA of length 256 bps at temperature, $T = 340\ K$. The base-pair distance is represented in gray scale. The white color means that the base-pair is closed, whereas black color means that the base-pair is open. Taken from [34], which does not provide the simulation time in real units.

is clear that nonlinear stacking interaction results in sharpness in denaturation profile unlike in harmonic stacking case. So, the harmonic approximation to stacking does not explain the sharpness of the denaturation profiles as shown in Fig. 1.10.

The dynamics of this model can be studied either using molecular dynamics simulations [33, 35] or nonlinear analysis [36, 37]. The numerical simulations of the dynamics of a model DNA of length 256 bps is studied by keeping in contact with a thermal bath simulated according to Martyna et *al.* [38]. In Fig. 1.11, a typical realization of molecular dynamics simulation at temperature, $T = 340\ K$ is shown. Mainly, there are two kinds of patterns in the realization. The first apparent pattern is the opening of few tens of base-pairs which lasts over a finite time. This pattern corresponds to "denaturation bubbles" in DNA. The other pattern is fast opening and closing of a few base-pairs which corresponds to "breathing" of DNA observed in experiments.

The non-linear dynamics of denaturation for this model DNA is studied using the harmonic stacking potential given in Eq. (1.8). The equation of motion corresponding to the Hamiltonian given in Eq. (1.5) is

$$\frac{\mathrm{d}^2 Y_n}{\mathrm{d}\tau^2} = S\left(Y_{n+1} + Y_{n-1} - 2Y_n\right) - \frac{\mathrm{d}V(Y_n)}{\mathrm{d}Y_n} \tag{1.12}$$

where $S = K/(D\alpha^2)$, $Y_n = \alpha y_n$ are dimensionless quantities. Eq. (1.12) forms coupled non-linear differential equations. A localized breather-like solution is obtained [31, 37] using a small-amplitude expansion, $Y_n = \epsilon\phi_n$ which leads to non-linear Schrödinger (NLS) type equation. The breather solutions to this equation are thought to be representing the "breathing of DNA", a fast opening and closing of bubbles in DNA which is also observed in molecular dynamics simulations as shown in Fig. 1.11.

Even though this model captures the experimental denaturation profiles [39], it lacks few

important properties of DNA. As the helical axis of the PB model DNA is always straight, the bending fluctuations of the DNA are not captured in this model. The DNA breathing appears as the localized breather solutions of NLS equation, but the real dynamics of DNA in solution is over-damped. In this model, DNA breathing comes from the non-linearity in the Morse potential. The torsional degrees of freedom of the base-pairs are also relaxed in this model.

1.5.2 Barbi-Cocco-Peyrard model

The torsional degrees of freedom are neglected in the simple Peyrard-Bishop model. However, they should be considered for the simple fact that dsDNA has a double helical structure. The dynamics of bubble, opening or closing of base-pairs, are coupled with twisting or untwisting of the two single strands. The present model was developed by Barbi et *al.* [40, 41] in 1999. The model is an extension of the Peyrard-Bishop model by allowing the base-pair to rotate around the helical axis (twisting) defined by the center of mass of the two single strands. Each base-pair has two degrees of freedom, stretching y and twisting φ, unlike in the PB model where only one degree of freedom was considered per base-pair. The twisting angle, φ_n, is defined as the angle by which the n^{th} base-pair is rotated with respect to $(n-1)^{\text{th}}$ base-pair. Each single strand is modeled by a flexible rod of bases of mass m. As in PB model, the base of each strand interacts and forms hydrogen bonding interactions which is also modeled by the Morse potential. The total Hamiltonian of the model is

$$\mathcal{H} = \sum_n \left(\frac{1}{2} m(\dot{y}_n^2 + y^2 \dot{\phi}_n^2) + V_h(y_n) + K(h_n - H)^2 + V_s(y_n, y_{n-1}) \right) \tag{1.13}$$

Where $V_h(y_n)$ is the hydrogen bonding potential given by Eq. (1.6) and $V_s(y_n, y_{n-1})$ is the stacking potential given by Eq. (1.5). In Eq. (1.13), the potential $K(h_n - H)^2$ keeps the distance in between two consecutive base planes close to H. Thus the distance, h_n, is related to y_n, y_{n-1} and $\theta_n = \phi_n - \phi_{n-1}$ as follows:

$$h_n = \sqrt{L^2 - y_n^2 - y_{n-1}^2 + 2 y_n y_{n-1} \cos(\theta_n)} \tag{1.14}$$

The helicoidal structure arises when $H < L$ for which the potential $K(h_n - H)^2$ reaches its minimum for non-zero twist angle. Note that there is a slight difference between model described in [40, 42], where the backbone is considered to be a flexible rod and the distance between base-pair planes is kept fixed, and the one described in [41], where the

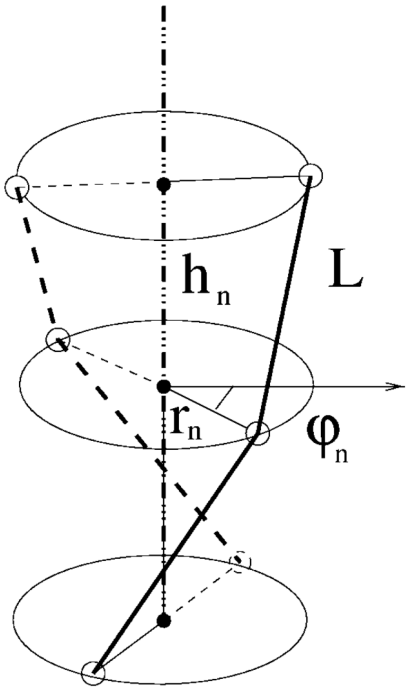

FIGURE 1.12: Sketch of a model DNA. Each base-pair is modeled by the base-pair distance r_n, twisting angle φ_n, and axial distance h_n between the base-pair planes. L is the backbone length along each strand. Taken from [41].

backbone is considered to be a rigid rod and the distance between the base-pair planes is allowed to fluctuate. But both models give rise to the same results.
Even though this helicoidal model of DNA is slightly different in geometrical considerations from the original PB model, the model gives rise to almost the same features. The bubbles or breathers also appear as localized excitations as the solution of NLS equation [43, 44]. However, the bubbles in this model are associated with untwisting as well. This model also lacks the conformational fluctuations of the whole DNA chain.

1.5.3 Kim-Jeon-Sung model

Here DNA is modeled by two interacting semiflexible chains, each made of N beads [45]. A sketch of model DNA is shown in Fig. 1.13. The Hamiltonian, $\mathcal{H} = \mathcal{H}_{\rm el}^{(1)} + \mathcal{H}_{\rm el}^{(2)} + \mathcal{H}_{\rm int}$, has three terms. The elastic energy of strands $i = 1, 2$ is

$$\mathcal{H}_{\rm el}^{(i)} = \sum_{n=2}^{N-1} \left[\frac{k_{\rm s}}{2} \left(\mathbf{r}_{n+1}^{(i)} - \mathbf{r}_{n}^{(i)} \right)^2 + \frac{k_{\rm b}}{2} \left(\mathbf{r}_{n-1}^{(i)} - 2\mathbf{r}_{n}^{(i)} + \mathbf{r}_{n+1}^{(i)} \right)^2 \right] \tag{1.15}$$

The first term in Eq. (1.15) represents the stretching energy between the two consecutive beads along each chain with a stretching constant, $k_{\rm s}$. The second term is the bending

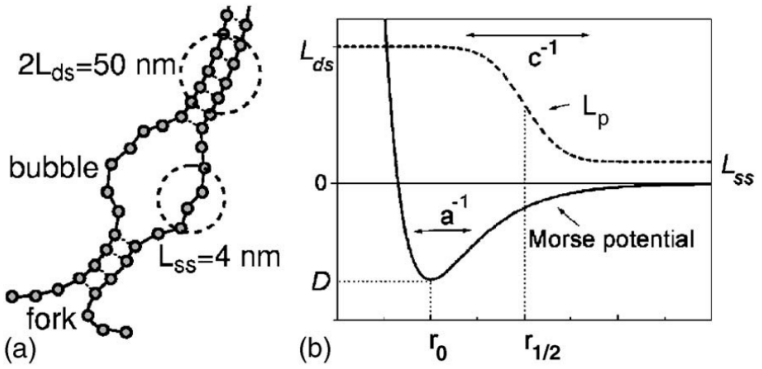

FIGURE 1.13: (a) Sketch of model DNA with a denatured bubble and a fork on one end of the DNA. (b) Variability of persistence length of dsDNA, L_{p} with base-pair distance r. Taken from [46].

energy with bending constant, k_{b} which is related to the chain persistence length. The interacting potential, $\mathcal{H}_{\mathrm{int}}$ is modeled by the Morse potential which is given in Eq. (1.6). The stacking interactions are introduced by changing the bending constant, k_{b} with intra base-pair distance, r which is shown in Fig. 1.13(b). This stacking interaction ensures that the DNA is stiffer in double stranded state than in single stranded state (bubble or fork), as shown in Fig. 1.13(a).

The stacking interactions are introduced similar to PBD model [32]. The effect of stacking interactions [46] and the effect of DNA breathing on the flexibility of DNA [47] have been investigated. The model predicted that the lifetime of bubbles increases non-linearly with the bubble size, giving several nano seconds. The helicity of the duplex and the effect of sequence are ignored in this model.

1.5.4 Metzler et *al.* model

Metzler and co-workers studied DNA bubble dynamics or DNA breathing based on the Poland-Scheraga (PS) model [48]. The dynamics of a single bubble of size n, sandwiched between the two stiffer dsDNA arms was studied. The free energy of PS model for such configuration of DNA is written as [48]

$$\beta\mathcal{F} = n\gamma(T) + \gamma_1 + c\ln(n+1) \tag{1.16}$$

where c is the loop entropy exponent, γ_1 is the initial energy barrier for the creation of the bubble and γ is the Gibbs free energy per base-pair (in units of $k_{\mathrm{B}}T$). Using a Fokker-Plank equation for the probability density function, characteristic times for bubble closure and bubble opening are found. The characteristic time for the closure

of the bubble of initial size n_0 depends on the temperature. At temperatures below the melting temperature, $T < T_m$, the characteristic time τ is

$$\tau = \frac{n_0}{D\gamma} \tag{1.17}$$

where D is the diffusion constant associated with the closing and opening of a base-pair, which is taken as a free parameter. Thus, at temperatures $T < T_m$, the bubble closure time, τ, scales linearly with the initial bubble size, n_0. At temperature $T = T_m$, the characteristic closure time for the bubble with initial size n_0 is $\tau = n_0^2/(2D)$ which is a quadratic dependence. Later on, a stochastic approach using Gillespie scheme [49, 50] and the Master equation approach was proposed to study the bubble dynamics from which the relevant quantities like the probability density function for the bubble size, characteristic timescales for opening or closure of bubble were measured [51, 52, 53]. One should note that all these approaches are based on PS model. In all these approaches, the rates associated with closing or opening of a base-pair are related to the statistical weights of the bubble from PS model. There are two major approximations employed here:

1. The statistical weights of PS model are equilibrium quantities ;
2. The statistical weight of the unbound segment (bubble) is correct for large bubbles only as for short bubble persistence length of ssDNA plays a role;

1.6 Bubble Dynamics: Experiments

Altan-Bonnet et *al.* [13] performed fluorescence correlation spectroscopy experiments on synthetic DNA constructs to study bubble dynamics. Heteropolymer DNA constructs of length $N = 30$ bps are used in experiments. Three DNAs of different sequence are used in experiments, namely M_{18}, A_{18} and $(AT)_9$. All the three sequences have the same end constructs, but they differ in middle part of the sequence which is either A or T. The typical DNA construct, M_{18}, is shown in Fig. 1.14.

Bubble dynamics is studied by measuring the correlation function of the fluorescence amplitude coming from a solution of internally tagged DNA constructs. From the correlation function, the breathing timescale is measured. In Fig. 1.15, an Arrhenius plot of breathing timescales is plotted against temperature for the three DNA constructs mentioned earlier with a slight dependence of the intercept on the sequence/DNA construct. It is clear that

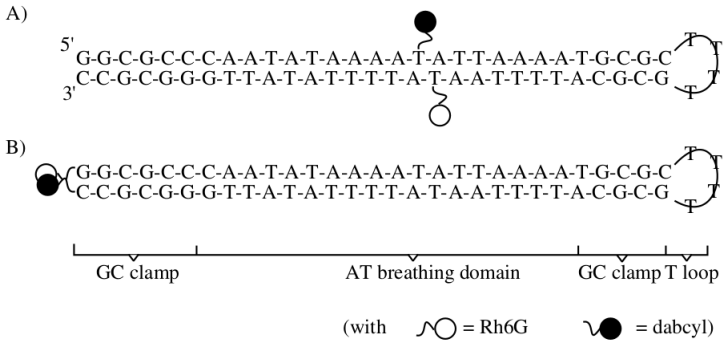

FIGURE 1.14: Sketch of M_{18} DNA constructs: (A) with internal fluorescent tagging; (B) with end fluorescent tagging. Taken from [13].

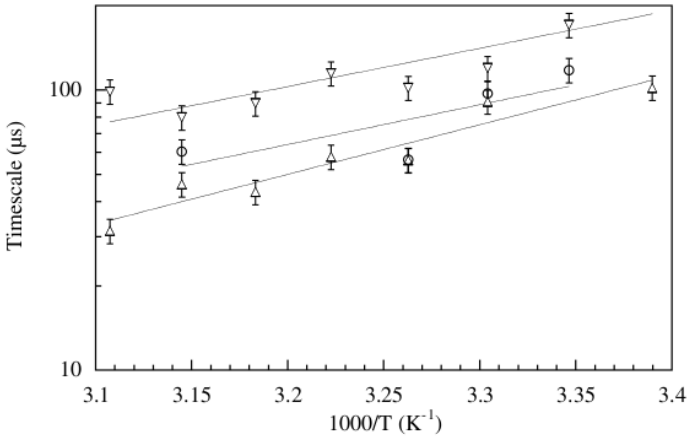

FIGURE 1.15: Arrhenius plot of breathing timescale vs temperature for three DNA constructs, M_{18} (○), A_{18} (△) and $(AT)_9$ (▽). Taken from [13].

the breathing timescales measured are on the order of tens of μs. The Arrhenius plot shown in Fig. 1.15 has an activation energy of around 7 kcal mol^{-1} $\simeq 12\ k_B T_0$. The total breathing timescales for the bubble of length 18 bps are found to be $20 - 100\ \mu s$ at room temperature.

Apart from FCS measurements, NMR measurements of the imino proton exchange [54] on DNA oligomers of 14 bps also observed bubble lifetimes of about 1 μs. *In vivo*, this situation also arises in the final stage of transcription when RNA polymerase leaves the locally open DNA (see Fig. 1.4).

Chapter 2

Coarse-Grained numerical modeling of dsDNA

In this chapter, we review the coarse-grained models of DNA available in the literature, the two coarse-grained models developed by us and our numerical algorithm. Coarse-graining is a mesoscopic representation of any biophysical system replacing all degrees of freedom by a smaller amount of effective degrees of freedom, i.e smoothing away the unessential microscopic details at the spatial and time scales of interest. Coarse-grained models are useful, as all-atom simulations are almost forbidden for large systems in the range of 50-100 bps or timescales in the μs range.

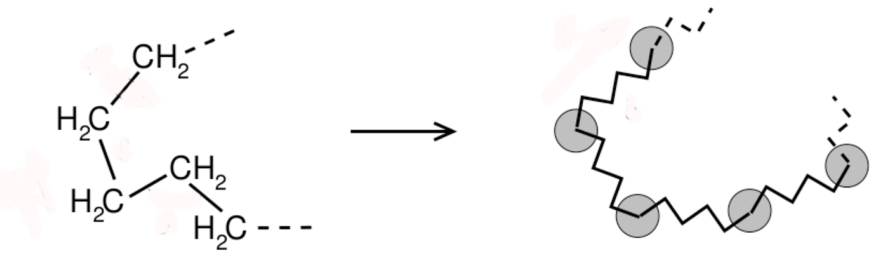

FIGURE 2.1: Illustration of a polymer coarse-grained from its chemical structure. Each CH_2 group is replaced by a bead and each C–C bond is replaced by a spring. Taken from J. Baschnagel et *al.* [56]

2.1 Basics of Coarse-Graining

In this section, we describe the basic ingredients of coarse-graining of bio-polymers. Coarse-graining of bio-polymers dates back to the 1970s. Bishop et *al.* [55] first simulated a single polymer chain in a solvent in 1979 using molecular dynamics simulations. Coarse-graining of bio-polymers is done by replacing each atom or group of atoms by an effective bead. The shape of the bead depends on the resolution, which in turn modifies the friction coefficient. In most of the cases, it is chosen to be spherical. The consecutive beads are connected usually *via* springs whose equilibrium distance is the bond distance. These models are also called bead-spring models. All the equilibrium parameters like bond lengths, bending angles and dihedral angles are maintained by the use of local potentials. The "covalent bonds" between the beads are modeled using harmonic potentials

$$V_{\text{har}}\left(r_{ij}\right) = \frac{\kappa_s}{2}\left(r_{ij} - r_0\right)^2 \tag{2.1}$$

where κ_s is the strength of the stretching potential. The distance between the beads i and j is $r_{ij} = |\mathbf{r}_{ij}|$, where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ and r_0 is the equilibrium bond length.
Flexibility or stiffness of the polymer chains is controlled by the bending angle potentials. The angle constraint between two consecutive bond vectors is a result of geometric constraints along the polymer chains. Available potentials used in the literature are:

$$V_{\text{bend}}\left(\theta_{ijk}\right) = \frac{\kappa_\theta}{2}\left(\theta_{ijk} - \theta_0\right)^2 \tag{2.2}$$

$$V_{\text{bend}}\left(\theta_{ijk}\right) = \frac{\kappa_\theta}{2}\left(\cos(\theta_{ijk}) - \cos(\theta_0)\right)^2 \tag{2.3}$$

The angle θ_{ijk} is between the vectors $\hat{\mathbf{r}}_{ij} = \frac{\mathbf{r}_{ij}}{r_{ij}}$ and $\hat{\mathbf{r}}_{jk} = \frac{\mathbf{r}_{jk}}{r_{jk}}$ defined as shown in Fig. 2.2, $\cos\left(\theta_{ijk}\right) = \hat{\mathbf{r}}_{ij} \cdot \hat{\mathbf{r}}_{jk}$ and θ_0 is the equilibrium angle. The strength of the potential κ_θ is related to the persistence length ℓ_p of the polymer as described in Chapter 1.

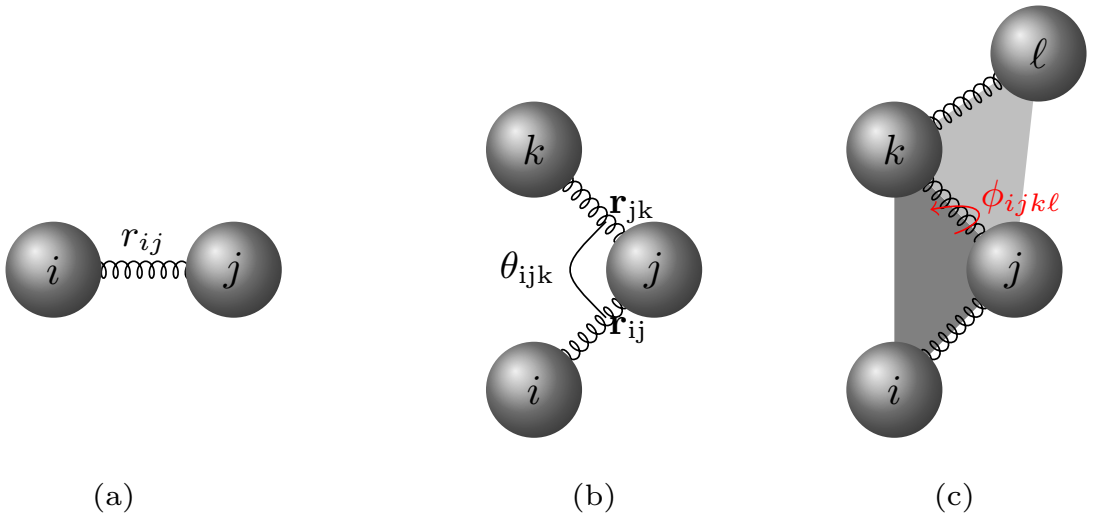

FIGURE 2.2: (a) Sketch of a harmonic potential between the beads i and j with an equilibrium distance r_0. (b) Sketch of a bending angle between the beads i, j and k and θ_{ijk} is the bending angle. (c) Sketch of a dihedral angle, $\phi_{\mathrm{ijk}\ell}$ between the two planes defined by ijk and $jk\ell$ is shown.

Torsion or dihedral angle potentials are used to constrain the rotation around the bond. The dihedral angle is defined for four consecutive beads, it is the angle between the two consecutive planes formed by the four atoms. Available potentials used in the literature are:

$$V_{\mathrm{dihedral}}\left(\phi_{ijk\ell}\right) = \frac{\kappa_\phi}{2}\left(\phi_{ijk\ell} - \phi_0\right)^2 \tag{2.4}$$

$$V_{\mathrm{dihedral}}\left(\phi_{ijk\ell}\right) = \kappa_\phi\left[1 - \cos(\phi_{ijk\ell} - \phi_0)\right] \tag{2.5}$$

$\phi_{ijk\ell}$ is the angle between the planes formed by $i\,j\,k$ and $j\,k\,\ell$, $\cos\left(\phi_{ijk\ell}\right) = \hat{\mathbf{n}}_{ijk} \cdot \hat{\mathbf{n}}_{jk\ell}$ as shown in Fig. 2.2, where $\hat{\mathbf{n}}_{ijk}$ and $\hat{\mathbf{n}}_{jk\ell}$ are unit vectors normal to the planes formed by the beads ijk and $jk\ell$ respectively. The equilibrium dihedral angle is denoted by ϕ_0. The range of dihedral angle, ϕ, is from 0 to 2π.

Apart from the bonded interactions, bending angle and dihedral angles and hydrogen interactions, there are long range attractive forces and short range repulsive ones. Short range repulsive forces are due to the overlap of the electron clouds. Attractive forces are due to the long range van der Waals interactions. Most commonly van der Waals interactions are modeled by a Lennard-Jones potential

$$V_{\mathrm{LJ}}\left(r_{ij}\right) = 4\,\epsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \tag{2.6}$$

where ϵ is the depth of potential well and σ is the inter-particle distance at which the

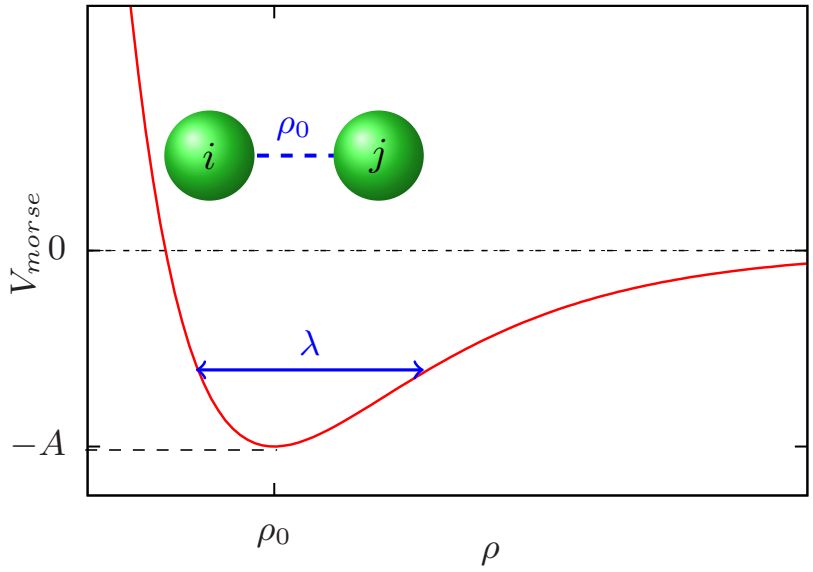

FIGURE 2.3: Sketch of a Morse potential of depth A, width λ and equilibrium distance ρ_0.

potential is zero. Excluded volume interactions are commonly modeled by a truncated repulsive Lennard-Jones potential

$$V_{\mathrm{LJ}}^{\mathrm{repulsive}}(r_{ij}) = \begin{cases} V_{\mathrm{LJ}}(r_{ij}), & \text{for } r \leq r_{\mathrm{m}} = 2^{\frac{1}{6}}\,\sigma \\ 0, & \text{for } r > r_{\mathrm{m}} = 2^{\frac{1}{6}}\,\sigma \end{cases} \tag{2.7}$$

From the above definition, one can check that the force is zero at r_{m}. This potential includes only the repulsive part of Lennard-Jones potential.
Hydrogen bonds (in dsDNA) are usually modeled by short range potentials like Morse or Lennard-Jones ones. The Morse potential [57] is widely used in the literature as it is analytically tractable. We used a Morse potential to model hydrogen bonds in DNA. The Morse potential explicitly includes bond breaking effects as it has two extrema, one at equilibrium distance and other at infinity

$$V_{\mathrm{Morse}}(\rho_{ij}) = A\left(e^{-\frac{2\left(\rho_{ij}-\rho_0\right)}{\lambda}} - 2\,e^{-\frac{\left(\rho_{ij}-\rho_0\right)}{\lambda}}\right) \tag{2.8}$$

where A is the Morse potential depth, λ is the width of the potential and ρ_0 is the equilibrium distance, as shown in Fig. 2.3.

As the DNA is charged, one has to consider the electrostatic forces in modeling DNA. The electrostatic forces or Coulombic interactions can be modeled using Debye-Hückel

approximation [58]. The electrostatic potential between charges q_i and q_j is given by:

$$V_{\text{Coulombic}} = \frac{q_i q_j}{4\,\pi\,\epsilon_0 \epsilon_k r_{ij}} e^{-r_{ij}/\kappa_D} \tag{2.9}$$

where κ_{D} is Debye length, ϵ_k is the dielectric constant of the medium, water, is 78 and ϵ_0 is the permittivity of the vacuum.

2.2 Overview of Coarse-Grained models of DNA available in the literature

Coarse-grained models of DNA are constructed with the same inspiration as for polymer models. As the double-stranded DNA consists of two complementary single-stranded DNAs, the interaction potentials mainly have two contributions, one resulting from the single strands elasticity and the other coming from the interaction between two complementary strands. The level of coarse-graining depends on the level of description of the nucleotide. All the coarse grained models can be classified based upon the number of beads per nucleotide or even nucleotide per bead [59]. Here, before introducing ours, we review a few coarse-grained models of DNA taken from the literature.

2.2.1 Two beads per nucleotide

Model by Zhang and Collins (1995) [60] :
In this model, the backbone and the bases are modeled by a single bead, but base beads have two or three hydrogen bonding interaction sites depending upon the sequence as shown in Fig. 2.4. The bases are attached to the backbone beads with a flexible rod. The consecutive backbone beads are connected *via* elastic rods of equilibrium length, $\boldsymbol{r}_{\text{b}}$. For simplicity, the motion of beads within each base is restricted to a two-dimensional plane as shown in Fig. 2.4. The model potential energy is:

$$V = \sum_{i=1,2} \sum_{n} \left[V^{(i)}_{\text{har}}(r_n) + V^{(i)}_{\text{bend}}(\theta_n) + V^{(i)}_{\text{dihedral}}(\tau_n) \right] + \sum_{n} (V_{\text{LJ}} + V_H) \tag{2.10}$$

The first three potentials are stretching [Eq. (2.1)], bending [Eq. (2.3)] and torsion [Eq. (2.5)] potentials applied to each strand separately. The stretching parameter r_n, bending angle θ_n and torsion angle τ_n run over all base-pairs. The index i runs over the two strands.

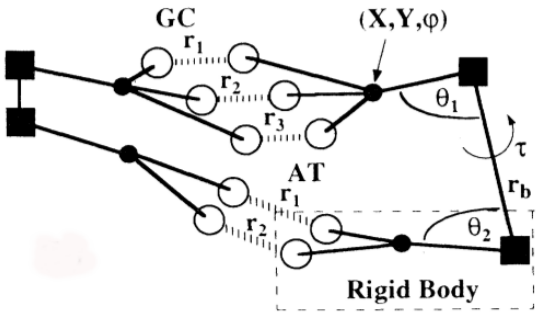

FIGURE 2.4: Schematic representation of the model DNA by Zhang and Collins [60]. Backbone sites (black squares), hydrogen bonding sites (white circles) and centers of mass (black circles) are indicated. Bond lengths, angles and torsion angles are defined. Taken from [60]

The stretching constant κ_s, bending constant κ_θ and torsion constant κ_τ are chosen equal to 1 kcal mol^{-1} Å^{-2}, 10 kcal mol^{-1}rad^{-2} and 40 kcal mol^{-1} respectively. The equilibrium parameters r, θ and τ depend on the sequence. The last two potentials in Eq. (2.10) are the Lennard-Jones potential Eq. (2.6) and the hydrogen bonding potential. The potential depth ϵ and distance σ in the Lennard-Jones potential are chosen equal to 1 kcal mol^{-1} and 4.0 Å respectively. The hydrogen-bonding potential is implemented by

$$V_H = V_0 \{\exp[-\alpha(r - r_{\text{eq}})] - 1\}^2 - \frac{V_0}{4}\{1 + \tanh[\beta(r - r^*)]\} \tag{2.11}$$

The second term of Eq. (2.11) mimics the formation of hydrogen bonds with the solvent molecules as soon as the distance between the complementary base-pairs is larger than the cut-off value. The depth of the Morse potential V_0 is chosen around 3.5 kcal mol^{-1}. The width of the Morse potential α is chosen to be 1.96 Å^{-1}. The solvent interaction factor β and solvent cut-off distance r^* are chosen to be 2.0 Å^{-1} and 4.0 Å respectively. Note that the above parameters depend on the sequence of DNA. The authors used molecular dynamics at constant temperature to simulate the above DNA model using the Nosé-Hoover thermostat [61, 62, 63]. The time step is chosen to be 1 ps. The model predicts sharp melting profiles but a slow melting kinetics (complete melting transition for 100 bps of DNA occurs only after 10 ns).

Model by G.C.Schatz et *al* (2000) [64] :

In this model, each nucleotide is also modeled by a backbone (the phosphate group) and a base (sugar and base). The model is a simplification of the previous one [60]. The model is sketched in Fig. 2.5. The total energy of the system is similar to Eq. (2.10),

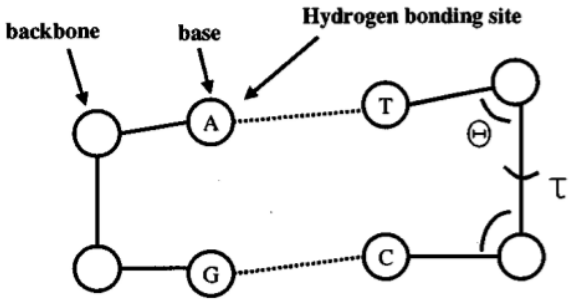

FIGURE 2.5: Representation of the model DNA of Schatz et *al.*, showing the backbones and the bases. Four possible bases are shown A,T,G and C. Taken from [64].

where all the potentials are the same except for the hydrogen bonding one:

$$V_H = (V_{H_1} - V_{H_2})\, f(\phi) \tag{2.12}$$

where, $V_{H_1} - V_{H_2} = V_H - V_0$, where V_H is given in Eq. (2.11), and

$$f = \begin{cases} \frac{1}{2}\left[\cos(\gamma\phi) + 1\right] & \text{if } \phi_{\min} \leq \phi \leq \phi_{\max} \\ 0 & \text{elsewhere} \end{cases} \tag{2.13}$$

The hydrogen bonding potential is thus modified in order to include an angle dependent term, $f(\phi)$. The hydrogen bonding potential depends not only on the acceptor-receptor distance, but also on the backbone-donor-acceptor angle ϕ, restricted to a finite interval $[\phi_{\min}\, \phi_{\max}]$. The factor γ makes sure that the hydrogen bonding vanishes at the boundaries ($\phi_{\min}$ and $\phi_{\max}$). The values of $\phi_{\min}$ and $\phi_{\max}$ are chosen to be -0.25 rad and 0.25 rad respectively. The model parameters for hydrogen bonding and excluded volume interactions also depend on the sequence. The authors have used Langevin dynamics, integrated using the Verlet algorithm [65]. The time step used in the simulations is $\Delta t = 10$ fs. The model is tested for many duplex decamers by comparing denaturation profiles to experimental ones.

Model by M.Sayar et *al.* (2010) [66] :

The model is composed of two beads per nucleotide, P for phosphate and sugar; B for base, as shown in Fig. 2.6. The total potential of the system is now

$$V = V_{\text{har}} + V_{\text{dihedral}} + V_{\text{LJ}} \tag{2.14}$$

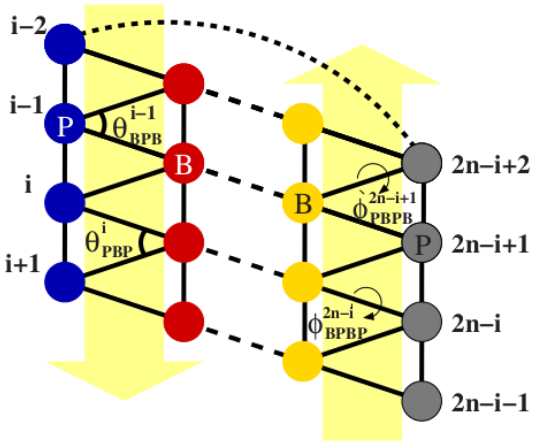

FIGURE 2.6: Schematic representation of a model DNA. Intra-strand bonds are shown by solid lines. Inter-strand bonds are shown by dashed lines. ϕ is the dihedral angle formed by four consecutive beads. Taken from [66].

The stretching potential (thick black lines in Fig. 2.6) is applied between the intra-chain beads and ensures the connectivity. Unlike in the other models, there is no explicit bending angle potential in the model, but angles are maintained *via* a stretching potential. So, stretching potentials in this model not only maintain the equilibrium distances, but also the equilibrium angles in the system. The stretching potential, V_{har}, is given in Eq. (2.1). The constants k_{s} and r_0 depend on the bond type (PP or PB or BB). The dihedral potential is applied to each strand to maintain the helicity. The dihedral potential, V_{dihedral} is given in Eq. (2.5). Excluded volume interactions V_{LJ} are implemented using the Lennard Jones potential, Eq. (2.7). ESPResSO package [67] is used for coarse-grained Molecular Dynamics simulations using a Langevin thermostat. All beads (including backbone and base) have the same mass, 170 atomic mass units. The equations of motion are integrated using the velocity-Verlet algorithm [65]. The time step used in simulations is $\Delta t = 0.1$ fs. The model captures the DNA pitch, the major and minor grooves of DNA as well. However, the persistence length is under-estimated, equal to 20 nm (60 bps). The model is used to study the formation of supercoils in DNA minicircles.

2.2.2 Three beads per nucleotide

Model by Knotts et *al* (2007) :

The model [58, 68] uses three beads per nucleotide, as shown in Fig. 2.7. The three beads correspond to the phosphate group, the sugar group and the base respectively. The total

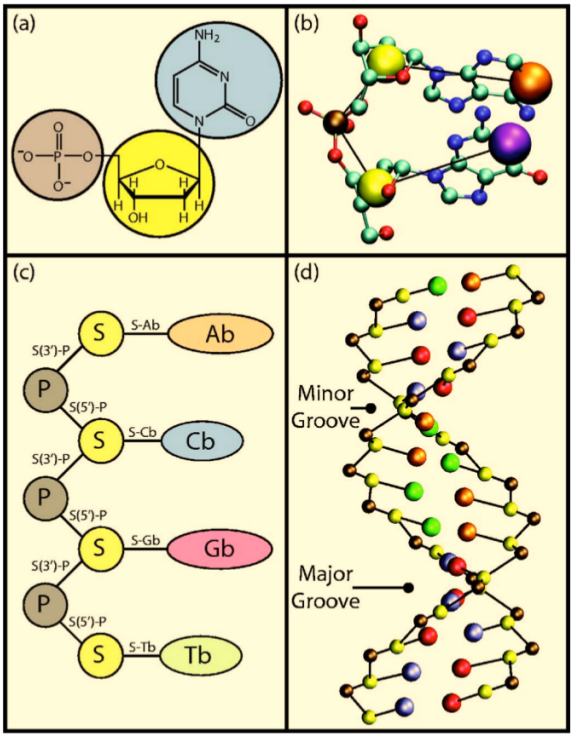

FIGURE 2.7: Schematic coarse-graining procedure for dsDNA. (a) Grouping of atoms. (b) Mapping of atomistic details to coarse-graining. (c) Topology of a single strand. (d) Model DNA of 13 bps with its minor and major grooves. Taken from [58].

potential of the system is

$$V = V_{\text{bond}} + V_{\text{bend}} + V_{\text{dihedral}} + V_{\text{stack}} + V_{\text{H}} + V_{\text{LJ}} + V_{\text{Coulombic}}$$

where the various potentials are

$$V_{\text{bond}} = \sum_{i}^{N_{\text{bond}}} \left[k_1 \left(d_i - d_{0_i} \right)^2 + k_2 \left(d_i - d_{0_i} \right)^4 \right] \tag{2.15}$$

$$V_{\text{H}} = \sum_{bps}^{N_{bp}} 4\,\epsilon_{bp_i} \left[5 \left(\frac{\sigma_{bp_i}}{r_{ij}} \right)^{12} - 6 \left(\frac{\sigma_{bp_i}}{r_{ij}} \right)^{10} \right] \tag{2.16}$$

The first three terms of Eq. (2.15) are intra-chain sequence specific potentials. The angle bending potential V_{bend} is given in Eq. (2.2). The dihedral potential V_{dihedral} is given in Eq. (2.5). Stacking (V_{stack}) and excluded volume interactions (V_{ex}) are modeled using Eq. (2.6). Electrostatic potential ($V_{\text{Coulombic}}$) is given in Eq. (2.9). The model parameter values are parameterized by the experimental melting curves. The molecular dynamics simulations are done in the NVT ensemble. Temperature is maintained using the Noosé-Hoover thermostat [61, 62, 63]. The time step used in simulations is $\Delta t = 1$ fs.

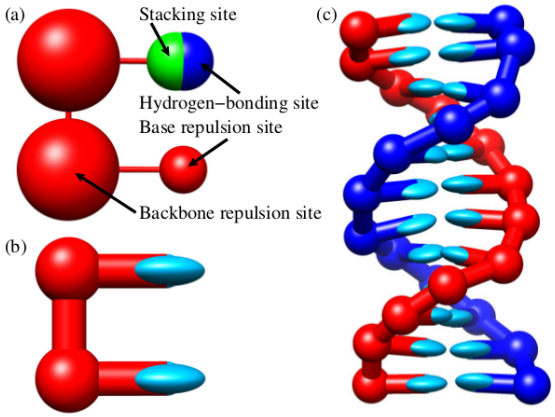

FIGURE 2.8: (a),(b)Model interaction sites. The size of the beads represents the range of excluded volume interactions. (c)12 bp of DNA as represented by the model. Taken from [70].

The model gives a value for the ssDNA persistence length close to the experimental value but $\ell_{\text{ds}} \approx 20$ nm for dsDNA, much smaller than the actual value of 50 nm. The sequence specificity is implemented by changing base-pair interactions and stacking interactions. The model correctly reproduces the experimental melting curves and the salt dependence of thermal melting. This model is slightly adjusted to give rise to correct thermal and mechanical denaturation properties for long sequences by Florescu et *al.* [69].

Model by Ouldridge et *al* (2011) : The model developed by Ouldridge et *al* [70, 71] consists in rigid nucleotides as shown in Fig. 2.8, with three colinear interaction sites. The interaction sites model backbone, base stacking and hydrogen bonding interactions. The total potential is

$$V = V_{\text{FENE}} + V_{\text{H}} + V_{\text{stack}} + V_{\text{c\_stack}} + V_{\text{excluded}} \tag{2.17}$$

Where, V_{FENE} is the finitely extensible non-linear elastic spring (FENE) [72, 73] between the nucleotides along each strands.

$$V_{\text{FENE}}\left(r_{ij}\right) = -0.5\,\kappa_s\,R_0^2\,\ln\left[1-\left(\frac{r_{ij}}{R_0}\right)^2\right] \tag{2.18}$$

In the FENE potential, R_0 is the maximum extent of the bond, where the potential diverges. The FENE potential should be used along with the Lennard-Jones Potential Eq. (2.6) which makes the potential smoother. In Eq. (2.17) $V_{\text{c\_stack}}$ is a cross stacking potential between two complementary strands. Even though there is no explicit torsion,

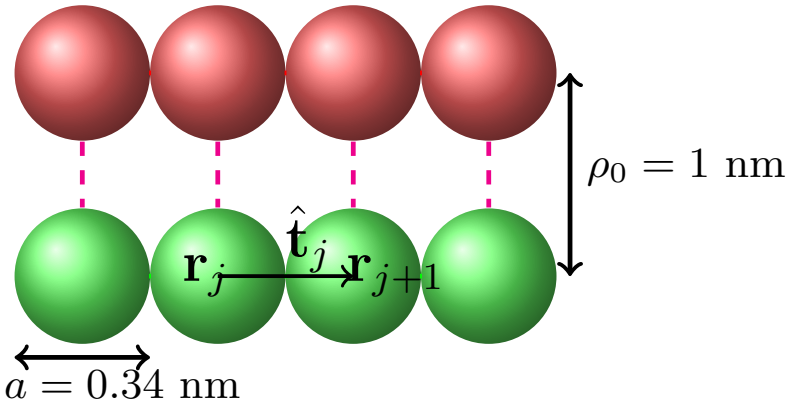

FIGURE 2.9: "Ladder" DNA consisting of two semi-flexible strands, where a is the bead diameter and ρ_0 is the equilibrium base-pair separation.

stacking and cross-stacking interactions give effective torsion to the model. The simulations are done using the Virtual Move Monte Carlo [74] method, used for bound systems. The model yields a ssDNA persistence length of 2 to 5 base-pairs and that of a dsDNA one $\ell_{\rm ds} \simeq 125$ bps. The model captures most of the thermodynamic and mechanical DNA properties.

There are many other coarse-grained models defined in the literature [41, 75, 76, 77, 78, 79, 80]. Most of the models that are defined above are computationally expensive if one is interested in reaching timescales on the order of μs or long size DNA molecules. In the next section, we describe two simple coarse-grained models of dsDNA developed by us [30, 81], allowing us to simulate times up to 100 μs : the ladder and the helical model.

2.3 Ladder DNA Model

The model is developed by us with the inspiration of PMD model 1.4.2, defined in Chapter 1. This model will be used in Chapter 3. An homopolymer DNA is modeled by two interacting bead-spring chains each made of N beads located at $\mathbf{r}_i$ as shown in Fig. 2.9, thus forming a "ladder", essentially planar up to thermal fluctuations, at short length scales. At large scales, it naturally becomes semiflexible.

The Hamiltonian, $\mathcal{H} = \mathcal{H}_{\rm el}^{(1)} + \mathcal{H}_{\rm el}^{(2)} + \mathcal{H}_{\rm int}$, has three terms. The elastic energy of strands $i = 1, 2$ is

$$\mathcal{H}_{\rm el}^{(i)} = \sum_{j=1}^{N-1} \left[\frac{\kappa_s}{2} \left(|\mathbf{t}_j| - a_0 \right)^2 + \kappa_{b,j} \left(1 - \hat{\mathbf{t}}_j \cdot \hat{\mathbf{t}}_{j+1} \right) \right] \tag{2.19}$$

where $\mathbf{t}_j = \mathbf{r}_{j+1} - \mathbf{r}_j$ and $\hat{\mathbf{t}}_j = \mathbf{t}_j/|\mathbf{t}_j|$. The first term of the rhs. of Eq. (2.19) is the stretching energy with stretching modulus $\beta\kappa_s = 100$ ($\beta^{-1} = k_{\rm B}T_0$ where T_0 is the room temperature) and $a_0 = 0.34$ nm is both the bead diameter and the equilibrium distance between two beads in each strand. The second term is the usual bending energy with a bending modulus $\kappa_{b,j}$ that depends on the local chain configuration as in Kim-Jeon-Sung model, described in Chapter 1. The interaction energy between the two strands (the hydrogen bonding between two complementary bases) is modeled *via* a Morse potential [31] of width λ and depth A which is given in Eq. 2.8. where $\rho_j = |\mathbf{r}_j^{(1)} - \mathbf{r}_j^{(2)}|$ is the distance between complementary bases at position j along the chain and $\rho_0 = 1$ nm is the equilibrium distance. The stacking interaction comes from the second term of rhs of Eq. (2.19) with a bending modulus κ_b which depends on ρ, interpolating from $\kappa_{\rm ds}/2 = 75\,k_{\rm B}T$ [1] for the dsDNA state to $\kappa_{\rm ss} = 3\,k_{\rm B}T$ for the single stranded one, according to [45]:

$$\kappa_{b,j} = \frac{\kappa_{ds}}{2} - \left(\frac{\kappa_{ds}}{2} - \kappa_{ss}\right) f(\rho_{j-1})f(\rho_j)f(\rho_{j+1}) \tag{2.20}$$

$$f(\rho_j) = [1 + \mathrm{erf}(\frac{\rho_j - \rho_b}{\lambda'})]/2 \tag{2.21}$$

where λ' is the width of the transition and $\rho_b = 1.5\rho_0$. The variable bending modulus depends on three consecutive base-pair distances, which provides cooperativity. We chose $\lambda = 0.2$ nm, $\lambda' = 0.15$ nm and $\rho_0 = 1$ nm. We have chosen $\beta A = 8$ such that an initial dsDNA remains always closed in the longest simulation run. The threshold value for ρ, discriminating between open and closed states, is fixed at 1.13 nm. The results are insensitive to slight change in this parameter. We used a Brownian Dynamics scheme for simulating the system, which will be discussed in section 2.5. We checked the equilibrium properties like, persistence lengths of ssDNA and dsDNA, which are close to 3 bps and 150 bps respectively.

2.4 Helical Model

We developed a new model which considered torsional degrees of freedom and it will be used in Chapter 4. The homopolymer DNA is now represented by two inter-wound polymer chains. Each bead-spring chain is made of N beads of radius 0.17 nm located at $\mathbf{r}_i$. Two consecutive beads along the same strand $\mathbf{r}_i$ and $\mathbf{r}_{i+1}$ are connected by strong springs. The distance between these two consecutive beads is maintained to 0.357 nm,

[1] The factor 1/2 comes from the fact that the dsDNA is made of two such strands.

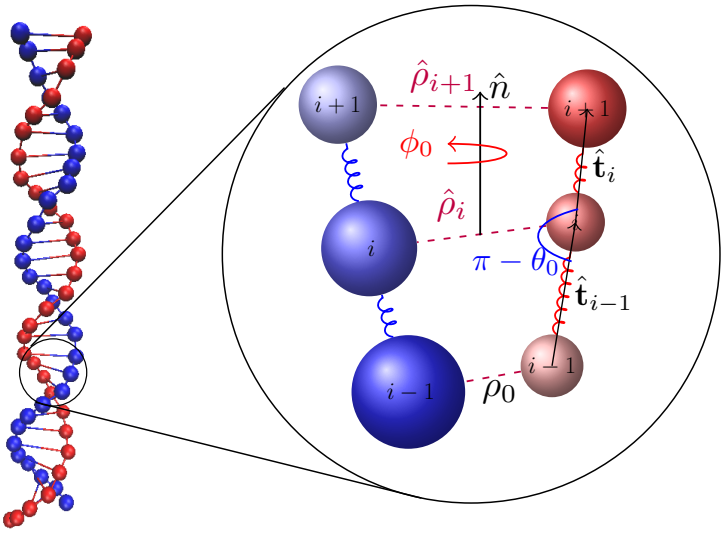

FIGURE 2.10: Geometry of the helical Model. P_i, P_{i-1}, P_{i-2} and Q_i, Q_{i-1}, Q_{i-2} represents two complemetary strands. θ_0 is equilibrium bending angle along each strand. ρ_0 is the equilibrium base-pair distance, $\hat{n}$ is the helical axis around which twist is defined and ϕ_0 is the equilibrium twisting angle.

a value larger than 0.34 nm as in Barbi-Cocco-Peyrard model, described in Chapter 1. The model geometry is shown in Fig. 2.10.

The total Hamiltonian is now

$$\mathcal{H} = \mathcal{H}_{\rm el}^{(1)} + \mathcal{H}_{\rm el}^{(2)} + \mathcal{H}_{\rm twist} + \mathcal{H}_{\rm int} \tag{2.22}$$

The first two contributions are elastic energies of the two strands which include stretching energy and angle bending energy. The third and fourth terms are torsion and hydrogen-bonding interactions respectively. These two potentials are intra-strand potentials.
The elastic energy of each strand goes as follows

$$\mathcal{H}_{\rm el}^{(i)} = \sum_{j=0}^{N-1} \frac{\kappa_{\rm s}}{2}(|\mathbf{t}_j| - a_0)^2 + \sum_{j=0}^{N-1} \frac{\kappa_\theta}{2}(\theta_j - \theta_0)^2 \tag{2.23}$$

The first term in Eq. (2.23) is the stretching energy with stretching constant $\kappa_{\rm s} = 100\ k_{\rm B}T_0$ and equilibrium distance $a_0 = 0.357$ nm. We kept large stretching constant to make the harmonic bonds stiffer. The second term is the angle bending energy. The angle bending potential keeps the angle between two consecutive tangent vectors along each strand (as shown in Fig. 2.10) constant. We thus now model the single strands as Freely Rotating Chains (FRC) [82].
The torsion energy is modeled by a simple harmonic potential around the mean twist

angle

$$\mathcal{H}_{\text{twist}} = \sum_{i=1}^{N-1} \frac{\kappa_\phi}{2}(\phi_i - \phi_0)^2 \tag{2.24}$$

The angle ϕ_i is measured between the two consecutive base-pair vectors $\boldsymbol{\rho}_i = \mathbf{r}_i^{(1)} - \mathbf{r}_i^{(2)}$ and $\boldsymbol{\rho}_{i+1} = \mathbf{r}_{i+1}^{(1)} - \mathbf{r}_{i+1}^{(2)}$. κ_ϕ is equivalent to the twist rigidity C in the Rod Like Chain (RLC) model [83]. The values of κ_θ and κ_ϕ are chosen to be 600 $k_{\text{B}}T_0/\text{rad}^2$ and 300 $k_{\text{B}}T_0/\text{rad}^2$ respectively to get reasonable dsDNA persistence length (ℓ_{ds}). The equilibrium value of ϕ is $\phi_0 = 0.62$ rad which is meant to keep the pitch of dsDNA close to 10 bps and θ_0 is chosen to be 0.41 rad in accordance with ϕ_0 [2]. The hydrogen bonding interaction is implemented *via* a Morse potential which is given in Eq. 2.8. The distance between the complementary bases of base-pair j is $\rho_j = |\boldsymbol{\rho}_j|$ and ρ_0 is the equilibrium distance which is set to be 1 nm. The depth and width of the Morse potential are 8 $k_{\text{B}}T_0$ and 0.2 nm respectively [30].

We have also implemented a stacking interaction by changing κ_ϕ with ρ, the distance between complementary bases:

$$\kappa_{\phi,j} = \kappa_\phi - (\kappa_\phi - \kappa_{\phi,ss})\ f(\rho_j) f(\rho_{j+1}) \tag{2.25}$$

where

$$f(\rho_j) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{\rho_j - \rho_b}{\lambda'}\right)\right]$$

$\kappa_{\phi,j}$ depends on the base-pair states ρ_j and ρ_{j+1}. We have chosen $\kappa_{\phi,ss}$ to be 0 $k_{\text{B}}T_0/\text{rad}^2$ since when the base-pair is denatured, the two ssDNA's are independent. We have chosen $\lambda' = 0.15$ nm which controls the width of the transition and $\rho_b = 1.5\,\rho_0$. A slight change in ρ_b and λ' does not change the results significantly. In reality, stacking interactions also exist in single stranded DNA and affect the flexibility of ssDNA [84, 85], which is lacking in our present model. The present model can be extended to include stacking interactions in single stranded DNA as a future direction. In both Ladder and Helical models, we did not implement the excluded volume interactions as we are mainly interested in short size DNAs. In any case, we checked the role of excluded volume interactions for a few realizations for both the Ladder model and the Helical model, the inclusion of excluded volume interactions did not change the mechanisms of closure.

[2] Because θ_0 and ϕ_0 are geometrically related.

2.4.1 Drawbacks of previous models

We anticipate that the dynamics of denaturation depends on the values of persistence lengths of dsDNA ℓ_{ds} and ssDNA ℓ_{ss}. To study the closure dynamics of denaturation bubbles, the model DNA needs to have values of persistence lengths experimentally relevant. Most of the coarse-grained models of DNA lack acceptable values for, either ℓ_{ds}, or ℓ_{ss}. The coarse-grained model developed by Knotts et *al* [58] results in $\ell_{ds} \simeq 60$ bps which is smaller than the actual value 150 bps. An extension of the same model by Sambriski et *al* [68] results in a good persistence length of dsDNA, $\ell_{ds} \simeq 150$ bps, but it lacks the good value for $\ell_{ss} \simeq 36$ bps. Similarly, the model developed by M. Sayar et *al* [66] results in ℓ_{ds} value around 96 bps. Even though other coarse-grained models [70, 71] results in good values of ℓ_{ds} and ℓ_{ss}, they are computationally expensive to reach the timescales (microseconds) that we are interested in. To avoid this, we developed a simple yet efficient model which results in good equilibrium values, like persistence lengths and pitch, which are discussed in the following section.

2.4.2 Equilibrium properties

A typical equilibrium configuration of our model DNA for $N = 30$ bps is shown in Fig 2.11. This simple model yet captures most of the essential features of the DNA: i) The *Right*

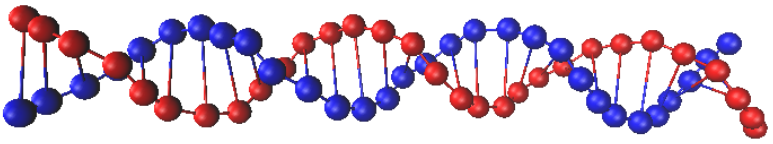

FIGURE 2.11: Equilibrium configuration of helical DNA of length 30 bps at room temperature.

Handed helicity is maintained by choosing the positive sign of the determinant of two base-pair vectors and normal vector. ii) We estimated the persistence length of dsDNA, $\ell_{ds} \simeq 160$ bps. The persistence length can be modified by tuning both angle bending and twisting potentials as they modify the local stiffness. iii) The pitch is found to be 12 bps. iv) The persistence of ssDNA, ℓ_{ss} is found to be 11 bps.

The method presented in [66] is used to compute ℓ_{ds} for $N = 150$ bps. We computed the correlation function between the beads of a single strand, $C(s) = \langle \hat{\mathbf{r}}(0) \cdot \hat{\mathbf{r}}(s) \rangle$, where $\hat{\mathbf{r}}$

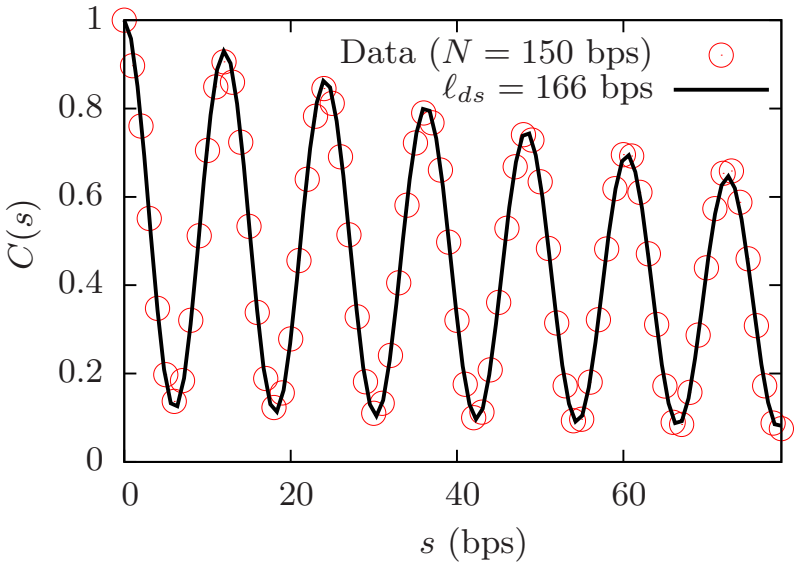

FIGURE 2.12: The correlation function $C(s)$ obtained for $N = 150$ bps dsDNA (The solid line is a fit from Eq. (2.26)).

is the unit vector connecting two consecutive beads along the single strand. We fit the correlation function with

$$C_{\rm th}(s) = e^{-s/\ell_{\rm p}} \left[a + (1-a)\cos(2\pi s/\lambda)\right] \tag{2.26}$$

where $\ell_{\rm p}$ is the persistence length and λ is the helical pitch. In Fig. 2.12, $C_{\rm th}(s)$ (solid line) is fitted to the data obtained for $N = 150$ bps of DNA. The persistence length of dsDNA is found to be $\ell_{\rm ds} \approx 160$ bps. The pitch is found to be 12 bps. Even though ϕ_0 was chosen to be 0.62 rad, other microscopic potentials such as the bending potential, the intra-strand one and thermal fluctuations modify the equilibrium value at room temperature. We also estimated the persistence length of ssDNA. We used 80 bases of ssDNA to compute the persistence length, $\ell_{\rm ss}$. In this case, the persistence length purely results from angle bending potential controlled by both κ_θ and θ_0 as there is no torsional potential in ssDNA. Theoretically using the Freely Rotating Chain model (FRC), the persistence length, $\ell_{\rm ss}$ is purely controlled by the equilibrium bending angle θ_0, and does not depend on the strength of the potential, κ_θ:

$$\ell_{\rm ss} = -\frac{a}{\ln(\cos(\theta_0))} \tag{2.27}$$

In Fig. 2.13, the persistence length of ssDNA is estimated by fitting the tangent-tangent correlation function at short distances with $\exp(-s/\ell_{\rm ss})$. For three values of θ_0, $\ell_{\rm ss}$ is found and plotted along with Eq. (2.27) in the inset of Fig. 2.13. As shown in Fig. 2.13, the persistence length is found to be 11 bps ($\simeq 4$ nm). Even though the experimental

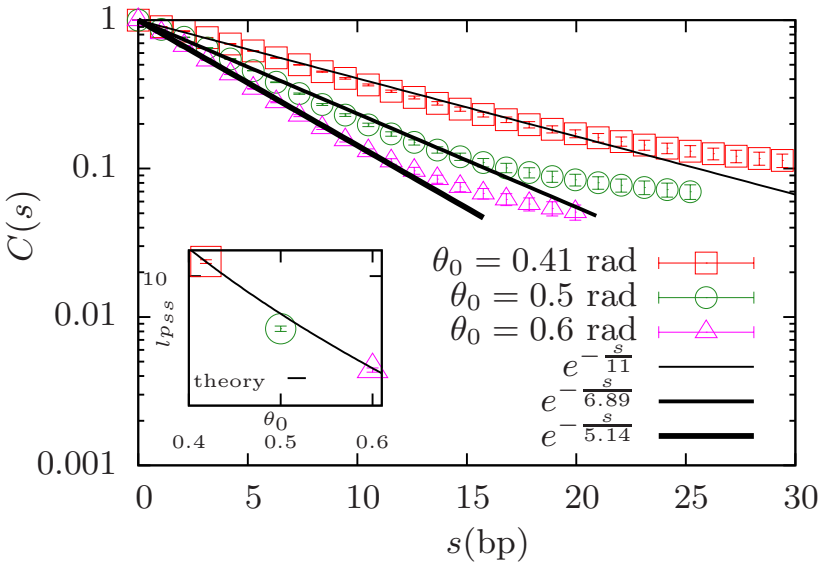

FIGURE 2.13: Linear-log plot of Correlation function $C(s)$ for three different θ_0 values, fitted with tangent-tangent correlation function $\exp(-s/\ell_{\rm ss})$. Inset: $\ell_{\rm ss}$ values corresponding to θ_0 are plotted along with theoretical prediction given by Eq. (2.27).

persistence length of ssDNA is known to be on the order of 1 nm [86], $\ell_{\rm ss} \simeq 4$ nm is also found by gel electrophoresis [87]. Moreover, $\ell_{\rm ss}$ depends on the salt concentration [88]. It is possible to change the persistence length of ssDNA in our model, but then one has to relax the dsDNA pitch.

Thus with a limited number of parameters in the model, we selected a parameter set which results in values of equilibrium parameters comparable to experiments, contrary to the other models presented in this chapter.

2.5 Numerical scheme

We use Brownian dynamics algorithm to simulate the coarse-grained models. Brownian motion is an erratic motion of colloidal particles caused by the random collision of solvent particles. This was first observed by Robert Brown [89]. The motion of these colloidal particles in the bath is given by the Langevin equation [90]

$$m\,\ddot{\mathbf{r}}_i(t) = -\nabla_{\mathbf{r}_i}\mathcal{H}(\{\mathbf{r}\}) - \zeta_i\dot{\mathbf{r}}_i(t) + \boldsymbol{\xi}_i(t) \tag{2.28}$$

where $\mathbf{r}_i$ is the position of the bead i. The first term on the right hand side of Eq. (2.28) is the deterministic force coming from the interaction potentials. The second term is the damping term, where ζ_i is the friction coefficient of beads. The last term is an effective

random force due to solvent particles. The basic assumption is that the timescale of solvent particles is much faster than that of colloidal particles, so that one can average out the fast modes of relaxation of solvent particles. The average force is zero :

$$\langle \boldsymbol{\xi}_i(t) \rangle = 0 \tag{2.29}$$

and $\boldsymbol{\xi}_i(t)$ follows the fluctuation-dissipation theorem [91],

$$\langle \boldsymbol{\xi}_i(t) \cdot \boldsymbol{\xi}_j(t') \rangle = 6\zeta_i k_{\mathrm{B}} T \delta_{ij} \delta(t - t') \tag{2.30}$$

Since we are dealing with the particles or beads of mass, few atomic mass units and of size, few nanometers, the momentum relaxation time becomes $\tau_{\mathrm{m}} \sim \frac{m}{6\pi\eta a} \sim 10^{-14}$s, whereas the typical diffusion time becomes, $\tau_{\mathrm{d}} \sim 6\pi\eta a^3/k_{\mathrm{B}}T_0 \sim 10^{-9}$s, which is much larger than τ_{m}. This rough argument ensures that we are in the high friction limit. In the high friction limit, one can neglect the inertial term from Eq. (2.28) [90], which simplifies to

$$\zeta_i \, \dot{\mathbf{r}}_i(t) = -\nabla_{\mathbf{r}_i} \mathcal{H}(\{\mathbf{r}\}) + \boldsymbol{\xi}_i(t) \tag{2.31}$$

Eq. (2.31) can be numerically integrated using the simple Euler's scheme :

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) - \left[\frac{1}{\zeta} \nabla_{\mathbf{r}_i} U(\{\mathbf{r}\}) + \sqrt{\frac{6\, k_{\mathrm{B}} T}{\zeta \Delta t}} \tilde{\boldsymbol{\xi}}_i(t) \right] \Delta t \tag{2.32}$$

Since the beads are identical, $\zeta_i = \zeta, \forall i$. $\tilde{\boldsymbol{\xi}}(t)$ is the random noise with mean 0 and variance 1. The value of Δt should be chosen small as it might create instabilities in the simulations. At the same time, Δt should be greater than the momentum relaxation time τ_{m} as we are dealing with the over-damped situation. We have chosen $\Delta t = 0.045$ ps.

2.6 Summary

In this chapter, the basics of polymers/biopolymers coarse-graining have been described. The coarse-grained models of dsDNA which have been introduced in the literature have been reviewed. At the end of this chapter, we introduced our models of dsDNA which will be used in the following chapters to study the problem of interest. Finally, we gave the simulation details of the Brownian Dynamics simulations in the free draining limit where one neglects hydrodynamics interactions (long range interactions) between the distant parts of the DNA/polymer.

Chapter 3

Denaturation bubble closure and bending

In this chapter, we study the closure dynamics of pre-equilibrated denaturation bubbles in dsDNA at room temperature. We use the ladder model 2.3 defined in Chapter 2. We show that the closure occurs in two stages:

1. a fast zipping driven by the attraction between strands, where the initial bubble eventually reaches a metastable state ;
2. the closure of the metastable state which is limited by the alignment of the two arms *via* rotational diffusion.

We also show that the bending of the bubble plays an important role in the bubble closure. We measured the closure times on the order of μs, larger than the times that are found in one-dimensional breathing models that do not take chain degrees of freedom into account [24, 26, 48, 92].

3.1 Numerical simulations

We used Brownian Dynamics (BD) simulations of two interacting semi-flexible strands. We also used the Kinetic Monte Carlo (KMC) algorithm which simulates the mean semi-flexible chain with an internal Ising spin dynamics corresponding to the bp state (broken or unbroken) [27, 29].

3.1.1 Brownian dynamics simulations

We studied the closure dynamics of a pre-equilibrated bubble in dsDNA using the ladder model 2.3 defined in Chapter 2. We created the bubble by switching off the Morse potential between strands and clamping the first and last three base-pairs to avoid the complete opening of dsDNA. The initial bubble of size $L(0) = N - 6$ was created in the middle of an homopolymer DNA of N bps. Then we equilibrated the bubble for about 2 μs. We simulated the closure of the bubble by switching on the Morse potential at $t = 0$. During the closure the first three base-pairs on each side of the initial bubble are kept closed by applying an inter-strand attractive potential of $100 k_{\mathrm{B}}T$, to avoid complete opening on either side of dsDNA. We also clamped the ends with a smaller value of 10 $k_{\mathrm{B}}T$ (which is slightly higher than the chosen value, 8 $k_{\mathrm{B}}T$). The results are insensitive to changes in this parameter value.

3.1.2 Kinetic Monte Carlo simulations

In addition to Brownian Dynamics simulations, out-of-equilibrium dynamics of the mean DNA chain (the center of mass of the two strands) has also been explored numerically by Kinetic Monte Carlo simulations. We implemented the coupled model defined in Refs. [27, 29] where the mean chain is composed of N identical beads representing the base-pairs (with the adequate friction coefficient). Simulation details are given in Ref. [59] (each bead now represents one base-pair and has the mobility of a pair of beads in BD simulations). At each Monte Carlo step of physical duration $\delta t = 0.019$ ps, a bead is chosen at random and a random move is attempted for this bead. In addition, at each Monte Carlo step, we also attempt to flip the sign of one Ising spin variable σ_i [see Eq. (1.4)], according to a standard Metropolis procedure. However, it might be that, in a real DNA, the frequency of change of internal degrees of freedom is different from this arbitrarily chosen one. To rule out this possibility, we simulated various systems where

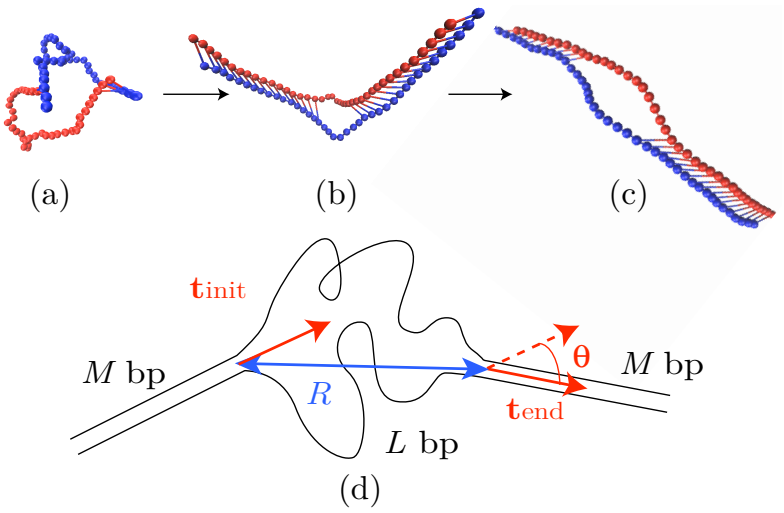

FIGURE 3.1: Snapshots of a typical Brownian dynamics simulation showing (a) the initial equilibrated bubble, (b) the metastable state, and (c) the bubble just before closure ($N = 40$ bp). (d) Sketch of the metastable state, showing two dsDNA arms of size M bp, the bubble of size L bps, mean tangent vectors of arms $\mathbf{t}_{\text{init}}$ and $\mathbf{t}_{\text{end}}$ *and* the angle, θ, between them. R is the end-to-end distance of the bubble.

10^{-3} to 100 spin-flips are attempted per δt. The average closure times then changed by at most $\pm 20\%$ as compared to 1 spin-flip, thus proving that this is not a critical issue.

3.2 Closure dynamics

In Fig. 3.2 the evolution of the bubble size (normalized by the initial bubble size $L(0)$), $L(t)/L(0)$ is plotted against time for a DNA molecule of length $N = 60$ bps. Along with the bubble size $L(t)/L(0)$, the tangent-tangent product $C(t) = \mathbf{t}_{\text{init}} \cdot \mathbf{t}_{\text{end}}$ and the end-to-end distance of the bubble $R(t)$ are plotted. The tangent vectors $\mathbf{t}_{\text{init}}$ and $\mathbf{t}_{\text{end}}$ are the mean tangent vectors of the two arms as shown in Fig. 3.1. The mean tangent vectors are constructed by taking the first two closed base-pairs before the beginning of the bubble and after the bubble end. For example, if the bubble is existing in between the 20^{th} and the 30^{th} base-pairs, we take the 18^{th} and the 19^{th} base-pairs and compute their mid-points in constructing $\mathbf{t}_{\text{init}}$. Similarly, we take the 31^{th} and the 32^{th} base-pairs in constructing $\mathbf{t}_{\text{end}}$. The end-to-end distance of the bubble is normalized by the contour length of the bubble $aL(t)$. From Fig. 3.2, it is clear that there are 2 regimes:

1. the *fast zipping* regime where the bubble size $L(t)$ decreases rapidly until it reaches an almost constant bubble size of average value denoted by $\bar{L}$

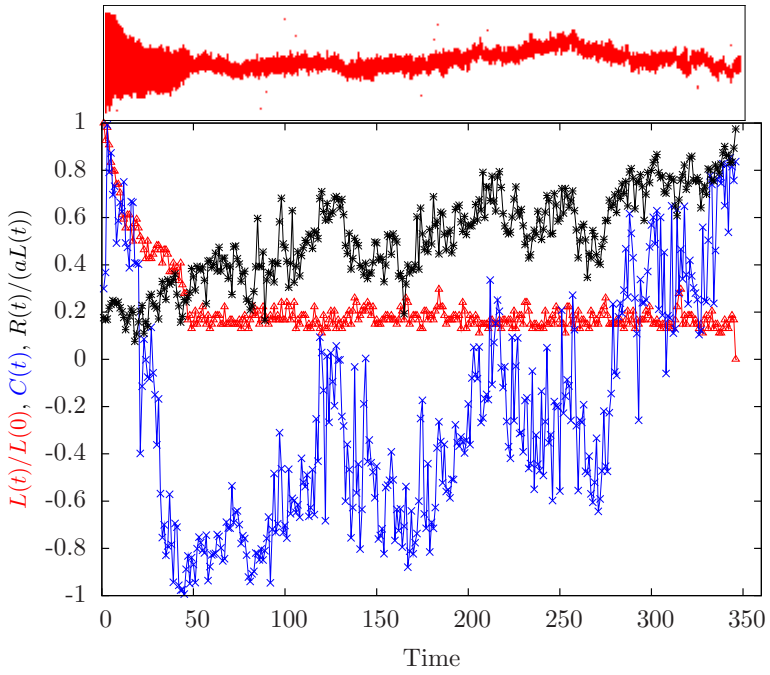

FIGURE 3.2: Top: melting map (position of open base-pairs in DNA as a function of time). Bottom: time evolution of the bubble length, $L(t)/L(0)$, in red, the bubble end-to-end distance, $R(t)/(aL(t))$, in black, and the tangent-tangent product $C(t)$ in blue ($N = 60$, time is in units of 10^4 BD steps = 0.45 ns).

2. the *metastable state* where the bubble of an almost constant size $\bar{L} \approx 10$ bps diffuses slowly along the DNA chain and eventually closes.

The zipping time, τ_{zip}, defined as the time required for the bubble to be zipped until it reaches the metastable state is faster than the diffusion time of the small arms (zipping time is around 5 orders of magnitude faster than diffusion time) so that the distance between both DNA extremities remains almost constant while the bubble "pushes" in the direction parallel to the arms such that $C(\tau_{\text{zip}}) \simeq -1$. This forms a "hair-pin" configuration just after zipping as shown in the Fig. 3.1b. The bubble in the metastable state is in a highly bent state with an approximate bending energy of 2 $k_{\text{B}}T$ (as observed from simulations) per base-pair. The final closure of bubble occurs when the bubble relaxes from this stress associated with the bending energy. From Fig. 3.2, it is clear that the metastable bubble closes when both arms are aligned ($C(\tau_{\text{closure}}) \simeq 1$) by rotational diffusion and the end-to-end distance of the bubble equals to the contour length of the same ($R(t)/aL(t) \simeq 1$).

Closure times simulated using both Brownian Dynamics and Kinetic Monte Carlo are shown in Fig. 3.3 as a function of the length of DNA, N. For one, the two types of

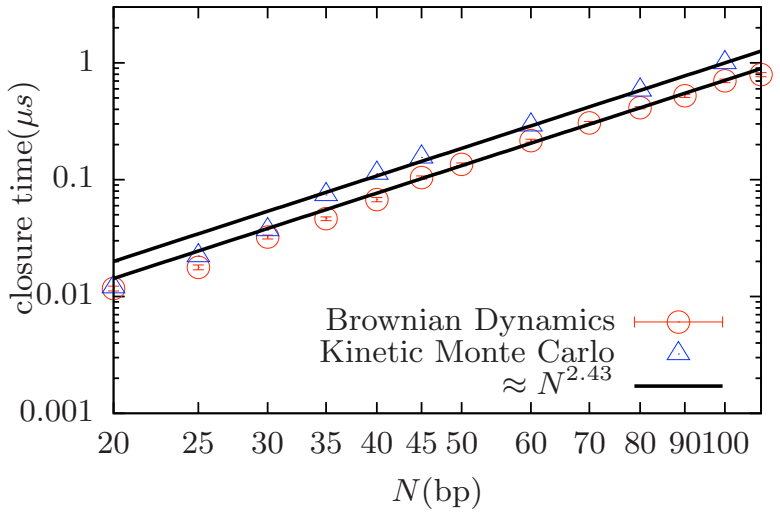

FIGURE 3.3: Log-log plot of the total closure time, τ_{closure}, *vs* total DNA length N. Circles correspond to BD simulations results and triangles to KMC ones. The solid line is a fit between $N = 50$ and 100 yielding $\tau_{\text{closure}} \sim N^{2.43}$.

simulations yield almost the same closure times and variation with N, showing that this quantity does not depend on the specific model. Since in KMC simulations the two strands are not simulated explicitly, this indicates that closure dynamics is dominated by the whole chain dynamics, which is the slowest process. For another, the fit of numerical data yields,

$$\tau_{\text{closure}} \sim \begin{cases} N^{2.4\pm0.1} & \text{for } 50 < N < 120 \\ N^3 & \text{for } N < 50 \end{cases} \tag{3.1}$$

Note that as we increase N further such that $N \sim 2\,\ell_{\text{ds}}$, τ_{closure} saturates. This will be clear when we discuss about the metastable residence time. In the next sections, we elaborately study the two regimes, zipping and metastable regimes, for different values of the Morse potential depth A.

3.2.1 Fast zipping process

In Fig. 3.4, averaged bubble size against time is plotted for $N = 100$ bps for three different Morse potential depths, $\beta A = 8$, $\beta A = 9$ and $\beta A = 10$. Averaged data are computed over 200 realizations. The data are plotted until the bubble size saturates (metastable regime). This is clearly seen in log-log plot shown in Fig. 3.4. Data for all the three values of A saturate for the same bubble size, $\bar{L}$, and are fitted with power law by taking a suitable cut-off value. Part of the data which start saturating falls outside the cut-off, thus is not

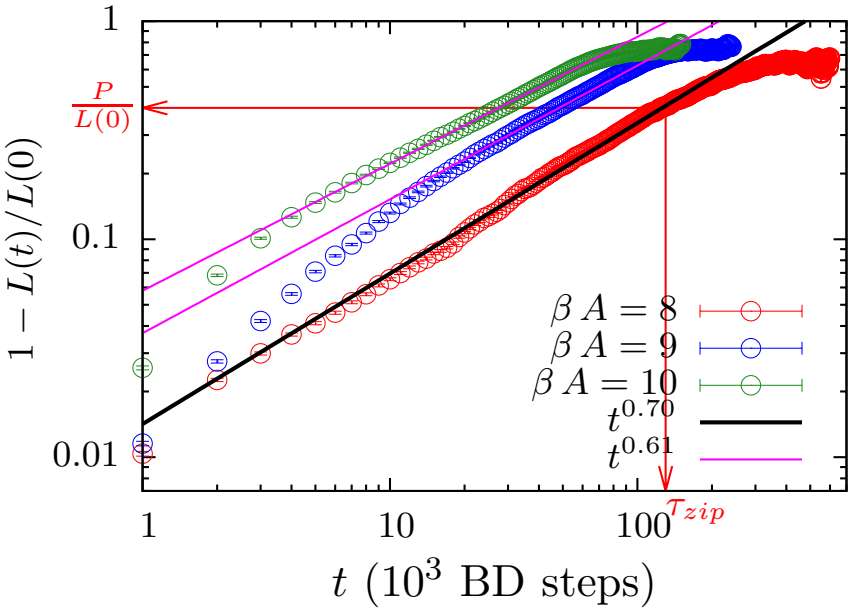

FIGURE 3.4: The averaged bubble size (normalized), $L(t)/L(0)$ for $N = 100$ bps and for different Morse potential depths is shown with time. Each data is averaged for 200 independent trajectories. Both P and τ_{zip} are shown for $N = 100$ bps and $\beta A = 8$.

considered in the fitting. Data for Morse potential depth $\beta A = 8$ scales with an anomalous exponent of 0.70, $1 - L(t)/L(0) \simeq t^{0.70}$. For higher values of Morse potential depths $\beta A = 9$ and $\beta A = 10$, a different scaling exponent of 0.61 is found. ATsimilar crossover in the exponents is also found in polymer adsorption [93]. Polymer adsorption is adsorption of a polymer on a solid surface with attractive potential acting on the polymer as shown in Fig. 3.5. The unzipped part of the bubble is compared to the unadsorbed part of the polymer. In [93] the authors found two different exponents for weak and strong adsorption energies: the adsorption time scales as $t \simeq N^{(1+2\nu)/(1+\nu)} \sim N^{1.4}$ and $t \simeq N^{1+\nu} \sim N^{1.6}$ respectively, where $\nu = \frac{3}{5}$ is the Flory's exponent. We recovered the same exponents, for higher Morse potential depths $1 - L(t)/L(0) \simeq t^{0.61} \simeq t^{1/1.6}$ and for lower Morse potential depths $1 - L(t)/L(0) \simeq t^{0.70} \simeq t^{1/1.4}$. The physical mechanism behind these two different problems appears to be same. The energy scale that differentiates the two regimes of different exponents is contributed by non-equilibrium effects [93]. Understanding the quantitative link between both mechanisms will deserve further investigation in the future. As the bubble size starts saturating at higher timescales, we have set the cut-off value as $P = \frac{3}{5}\left(L(0) - \bar{L}\right)$, to define the zipping time [94]. The zipping time, τ_{zip}, is defined as the time taken for P base-pairs to be zipped as shown in Fig. 3.4. In Fig. 3.6, we plotted

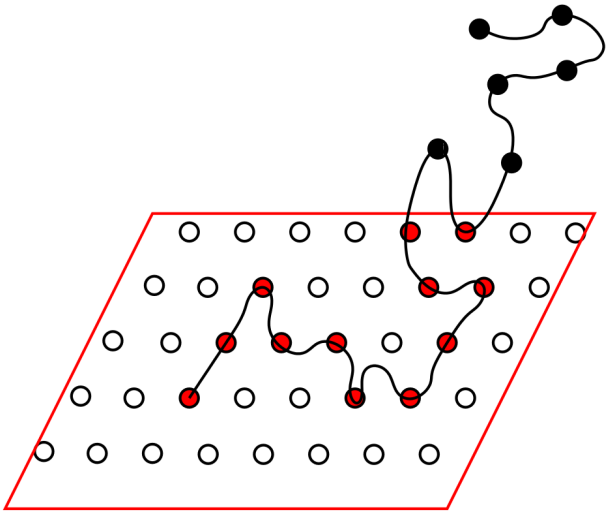

FIGURE 3.5: Sketch of polymer adsorption on a solid surface, black beads are unadsorbed part of polymer (of length $L(t)$) and red beads are adsorbed part of DNA. Taken from [93].

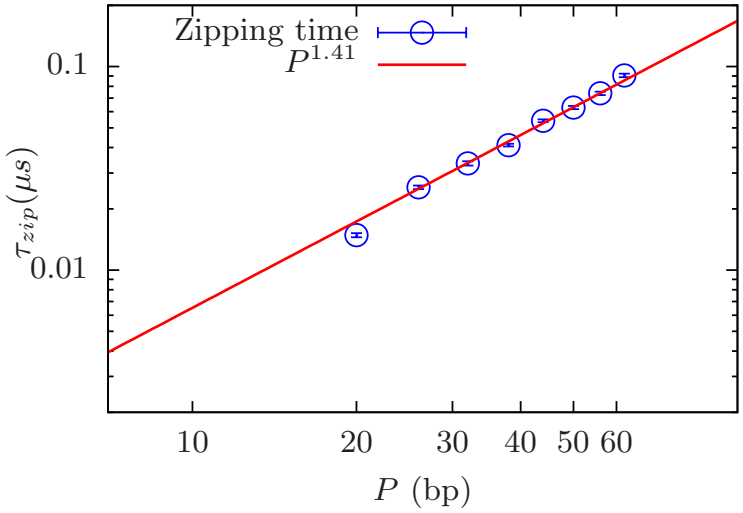

FIGURE 3.6: Log-log plot of the zipping time *vs* the zipped ssDNA length $P = 3(L(0) - \bar{L})/5$.

the zipping time, τ_{zip}, against P. The zipping time is also found to scale with P as:

$$\tau_{\text{zip}} \sim P^{1.4} \tag{3.2}$$

This anomalous scaling exponent is already observed in polymer translocation [94, 95] as explained below. The same exponent is also recovered in the case of DNA renaturation or hybridization, where two single strands of DNA form a dsDNA. *In vitro*, DNA hybridization is studied by placing the target ssDNAs in the pool of its complementary

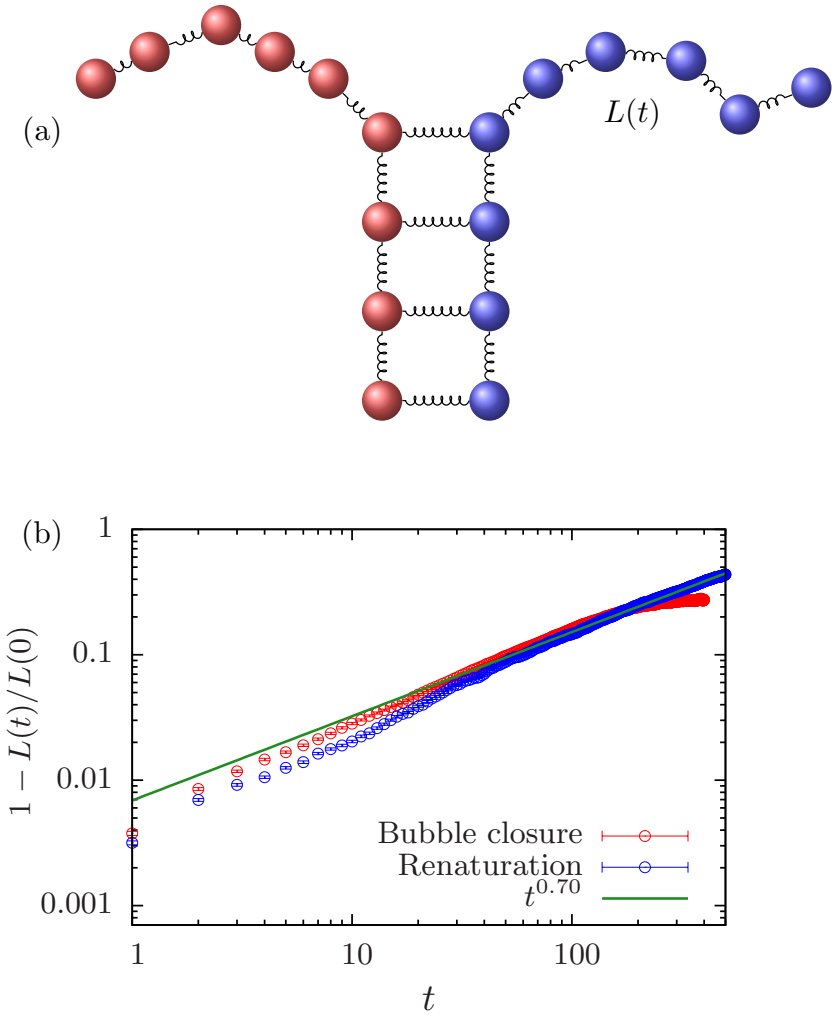

FIGURE 3.7: Zipping in DNA renaturation (a) Sketch of the model DNA going through renaturation, defining $L(t)$ as the length of denatured part of DNA. (b) Log-log plot of averaged bubble size *vs* time for both bubble closure and renaturation.

ssDNAs [96]. To understand the dynamics of hybridization with the current model, we started with an equilibrated configuration where one end is completely opened and the other is closed as sketched in Fig. 3.7a. In Fig. 3.7b, the averaged size of the denatured part of DNA in renaturation and the bubble size in bubble closure are plotted and we found the same scaling exponents. For comparison purpose, we have chosen the same $L(0)$ for both bubble closure and renaturation in normalization. As one can see from Fig. 3.7, bubble closure data saturates and renaturation data does not saturate. It implies that there is no metastable state during renaturation, the reason will become clear in the following sections.

The out-of-equilibrium dynamics of zipping can be understood thanks to an analogy with polymer translocation or polymer adsorption. The driving force (see Fig. 3.8) acting at

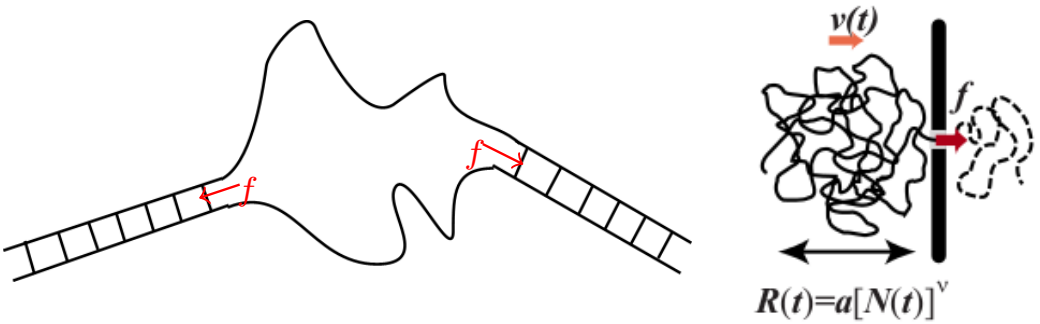

FIGURE 3.8: Sketch of (a) a bubble in the zipping regime, where f is the driving force acting at the bubble ends. (b) A polymer going under the translocation process (Taken from [97]). f is the pulling force, $v(t)$ is the velocity and $R(t)$ is the polymer segment involved in the friction.

the ends of the bubble in zipping can be compared with the pulling force acting on the polymer end at the tiny pore in polymer translocation [97] or the adhesive force acting on the starting monomer of the unadsorbed polymer in polymer adsorption. The driving force, f, is an effective force mainly resulting from the energy gain when the base-pair closes on each end of the bubble. The dynamics of polymer in the zipping stage can be written as

$$\eta\, R(t)\frac{\mathrm{d}(aL)}{\mathrm{d}t} = -f - \nabla E_{\mathrm{bend}} + \xi(t) \tag{3.3}$$

where $R(t)$ is the segment of the polymer size involved in friction and η is the viscosity of water. The force terms in Eq. (3.3) are effective mean forces only. The main assumption behind the Eq. (3.3) is that the end-to-end distance is in instantaneous equilibrium with $L(t)$ during the initial stage of zipping which is qualitatively valid as the zipping time is much faster than the Rouse time of ssDNAs of the bubble during the initial stage of zipping. For a flexible bubble, $R(t) \simeq a[L(t)]^{\nu}$ [97, 98, 99]. The first term on the rhs of Eq. (3.3) is the driving force, $f \simeq A/a$ acting at the ends of the bubble, the second term is the force arising from the bending energy of the bubble and the third term, $\xi(t)$ is the random force. As the size of the bubble reaches the order of the persistence length of ssDNA, ℓ_{ss}, the force arising from the bending energy stored in the bubble should also be taken into account unlike in the polymer translocation or polymer adsorption. This bending energy is, in the continuous limit:

$$E_{\mathrm{bend}}\,(L) = 2\int_{s_0}^{s_0+aL(t)} \frac{\kappa_{\mathrm{ss}}}{2}\left(\frac{\partial \hat{t}}{\partial s}\right)^2 \mathrm{d}s \tag{3.4}$$

At short times, the bubble is flexible and the contribution from bending energy is on the order of thermal noise and can be neglected

$$\eta\, R(t)\frac{\mathrm{d}(aL)}{\mathrm{d}t} = -f \tag{3.5}$$

From Eq. (3.5), using the fact $R(t) \simeq a[L(t)]^{\nu}$ and a simple integration [97] yields for short times,

$$\tau_{\mathrm{zip}} \simeq P^{1+\nu}/f \tag{3.6}$$

The exponent $1+\nu \simeq 1.6$ is slightly larger than 1.4. Eq. (3.6) is indeed valid in the limit of weak forcing only, otherwise it depends on the non-equilibrium response of pulling force in polymer translocation [97] or driving force in zipping.

As the bubble size during zipping reaches the order of the persistence length of ssDNA, ℓ_{ss}, the bending force starts competing with the driving force. When the bubble size becomes around 10 bps$\simeq 2\,\ell_{\mathrm{ss}}$, the bubble stops closing. By taking the second term in Eq. (3.3) into account

$$\eta\, R(t)\frac{\mathrm{d}(aL)}{\mathrm{d}t} = -f - \frac{\partial E_{\mathrm{bend}}}{\partial(aL)} \tag{3.7}$$

The steady state solution of Eq. (3.7) gives the metastable bubble size, which is given by $E_{\mathrm{bend}} \sim fa\bar{L}$. In a simple case of a circularly bent bubble, $E_{\mathrm{bend}} = \kappa_{\mathrm{ss}} 4\pi^2/4\bar{L} = \pi^2\kappa_{\mathrm{ss}}/\bar{L}$. The size of the metastable bubble becomes, $\bar{L} = \sqrt{\pi^2\kappa_{\mathrm{ss}}/A}$. For $\beta\kappa_{\mathrm{ss}} = 3$ and $\beta A = 8$, the metastable bubble size becomes $\bar{L} \approx 2.5$ bps. This predicted value is reasonable even if slightly lesser than the observed value. The reason could be that the bending constant of bases at the ends of the bubble are higher than the chosen value of κ_{ss} (physically because of the effect of stacking). So, the effective bending constant is slightly higher than the chosen value, $\beta\kappa_{\mathrm{ss}} = 3$. Thus the metastable state is evident as the cost of closing one more base-pair becomes much more than the driving force. Above arguments should also imply that the metastable bubble size should increase with $\sqrt{\ell_{\mathrm{ss}}}$ which is observed in the simulations. From simulations, the metastable bubble size is found to be 8 bps, 10 bps and 13 bps for $\beta\kappa_{\mathrm{ss}} = 1$, $\beta\kappa_{\mathrm{ss}} = 3$ and $\beta\kappa_{\mathrm{ss}} = 7$ respectively and the metastable bubble size $\bar{L}$ is found to scale with κ_{ss}, $\bar{L} \sim \kappa_{\mathrm{ss}}^{0.25\pm0.03}$. The exponent deviates from 1/2. It might be that the expected scaling is found for a circularly bent bubble which is not exactly the case in practice. Note that the effect of random forces is not studied here as we are interested in understanding the role of mean deterministic forces.

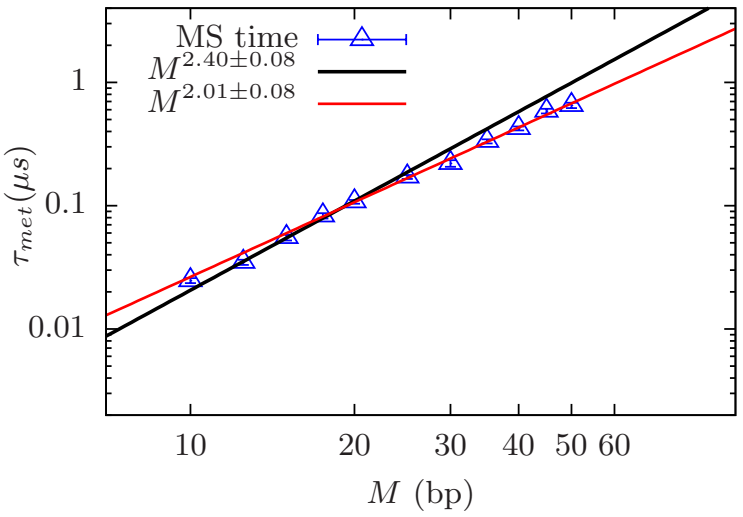

FIGURE 3.9: Log-log plot of metastable (MS) residence time *vs* the arm's length M. Fits correspond to $\tau_{\rm met} \sim M^2$ (Red line) and $M^{2.4}$ (Black line).

3.2.2 Metastable state

The metastable residence time or dwell time dominates over the zipping time, thereby giving the large timescales in the bubble closure. The dwell time, $\tau_{\rm met}$ is plotted against arm's size, M in Fig. 3.9. The dwell time is found to scale with M, $\tau_{\rm met} \sim M^2$ for the longest arms and $\tau_{\rm met} \sim M^{2.4}$ for shorter arms.

For short arm sizes, we recovered an exponent of 2.4 and for long arm sizes, an exponent of 2. As the arms size, $M < \ell_{\rm ds}$ $(M/\beta\kappa_{\rm ds} \sim 0.2 - 0.4)$, one would expect that the arms are stiff. For stiffer arms connected by a flexible joint, the rotational diffusion time of one arm is given by [82]:

$$\tau_{\rm R} \simeq D_{\rm R}^{-1} \simeq \frac{\eta(aM)^3}{k_{\rm B}T} \tag{3.8}$$

From the Fig. 3.1, the closure of metastable bubble occurs when both the arms are aligned for the first time from the hairpin state. The rotational diffusion of the two arms can be seen as the rotational diffusion of one arm with respect to the other. So, the metastable residence time is proportional to the rotational diffusion timescale, $\tau_{\rm R}$. Instead of an exponent of 3, we observed an exponent in the range of $2 - 2.4$. To understand this exponent of 3, we did simulations of two stiff polymers connected by a flexible polymer, see Fig. 3.10. We have chosen the bending constant ($\beta\kappa_{\rm ds} = 150$) large enough such that the arms are stiff and we have chosen a small bending constant $0 < \beta\kappa_{\rm ss} < 6$ for the connecting polymer such that it is flexible enough. The length of the flexible polymer is chosen to be 10 bps irrespective of length, N. By computing the orientational correlation

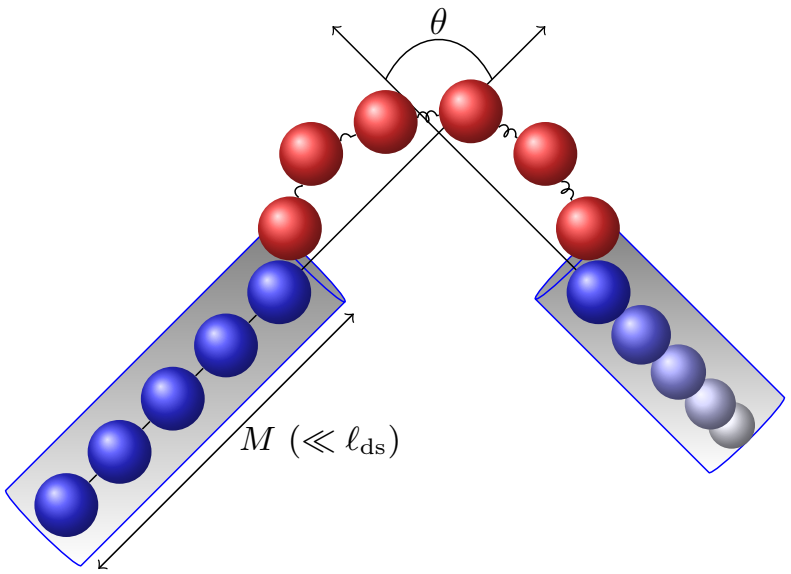

FIGURE 3.10: Sketch of two stiff rods connected by a flexible polymer.

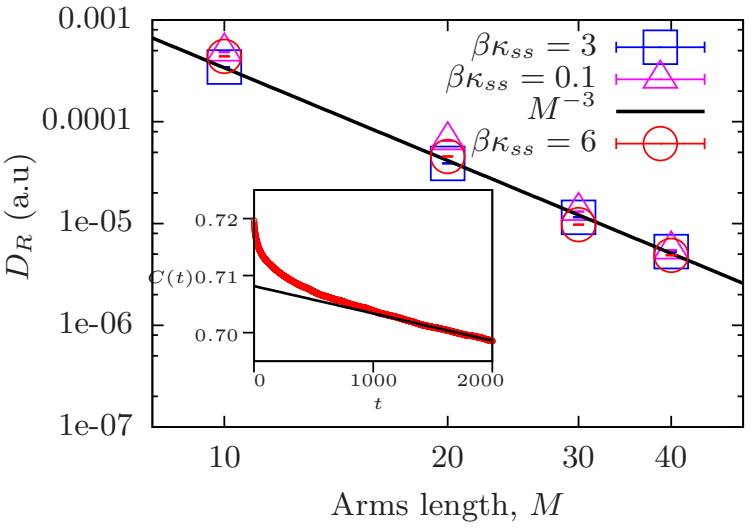

FIGURE 3.11: Rotational diffusion constant $D_{\rm R}$ plotted against arm size, M, for $\beta\kappa_{\rm ds} = 150$ and for different values of $\kappa_{\rm ss}$. Inset: The Correlation function (in red) is fitted at timescales larger than the transient regime in order to recover the diffusion constant.

function, $C(t) = \langle \cos(\theta(t)) \cdot \cos(\theta(0)) \rangle \simeq \exp(-2D_{\rm R}t)$, we measured the rotational diffusion constant $D_{\rm R}$ by fitting $C(t)$ at timescales larger than the transient regime, as shown in Fig. 3.11 (inset). For different values of $\kappa_{\rm ss}$, we computed the rotational diffusion constant and indeed recovered an exponent of 3.

To check that this mechanism holds in bubble closure dynamics, we did Brownian dynamic simulations of the ladder model for large value of $\ell_{\rm ds}$, $\beta\kappa_{\rm ds} = 400$, to decrease the

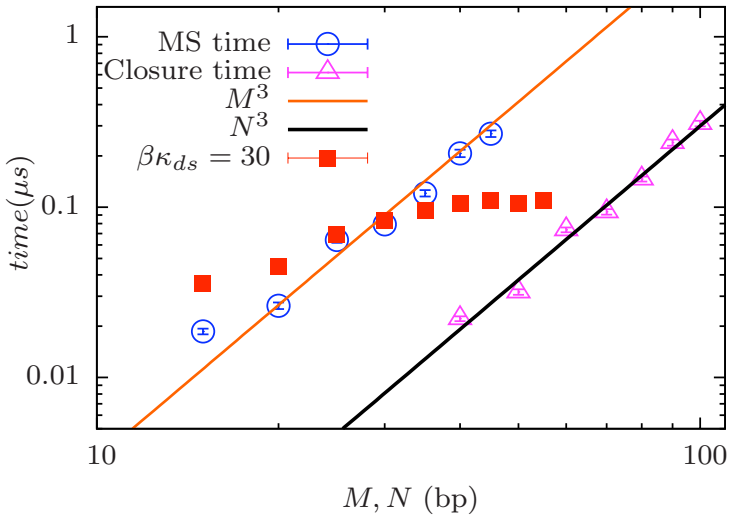

FIGURE 3.12: Log-log plot of (i) the MS dwell time and the closure time *vs* the arm's length M (circles) and chain length N (triangles) respectively for $\beta\kappa_{\rm ds} = 400$. Fits lead to $\tau_{\rm met} \simeq M^3$ and $\tau_{\rm closure} \sim N^3$. (ii) The dwell time *vs* M for $\beta\kappa_{\rm ds} = 30$ (squares) showing a saturation for $M > 2\beta\kappa_{\rm ds}$.

ratio $M/\ell_{\rm ds}$. We plotted both the dwell time and the total closure time with M and N respectively as shown in Fig. 3.12. Both the times $\tau_{\rm met}$ and $\tau_{\rm closure}$ scales with an exponent of 3 as expected.

The deviation of the exponent from 3 to 2 can be due to the following reasons:

1. The arms are semiflexible so that the relaxation of the bending modes might accelerate the dynamics.

2. The bubble is also semiflexible ($\bar{L} \sim 2\ell_{\rm ss}$) which might introduce a restoring elastic torque between the arms.

3. Diffusion law works only at long timescales and the closure might occur before the diffsion regime is reached.

So far, we were limited to finite range in M. To understand the effect of long arms ($M \geq \ell_{\rm ds}$), we did simulations for smaller $\ell_{\rm ds}$ (or $\kappa_{\rm ds}$), specifically $\beta\kappa_{\rm ds} = 30$. As one can see from Fig. 3.12 (denoted by ■), the dwell time $\tau_{\rm met}$ saturates when the arm size reaches $\ell_{\rm ds}$. For longer arms greater than $\ell_{\rm ds}$, $\tau_{\rm met}$ depends only on $\ell_{\rm ds}$. One can estimate the maximum dwell time $\tau_{\rm met}^{\rm max} \simeq \eta\beta\,(2a\ell_{\rm ds})^3 \simeq 20\mu s$ when M reaches $\ell_{\rm ds} = 150$ bps.

In this section, we described the metastable regime which is a dominant timescale in the bubble closure. We saw that the metastable state is limited by the rotational diffusion of the arms. We have seen how the dwell time depends on the arm size M and the persistence length of dsDNA $\ell_{\rm ds}$. For long DNAs of length $N \gg \ell_{\rm ds}$, the dwell time saturates and the maximum dwell time is about 20 μs for $\ell_{\rm ds} = 50$ nm.

Mean First Passage Time approach

From Fig. 3.2, we understand that the final closure occurs when both the arms are almost aligned ($\cos(\theta) \simeq 1$) and the end-to-end distance of the bubble almost equals to the contour length of the bubble ($R(t)/aL(t) \simeq 1$). So, the dwell time is the mean first passage time, $\tau_{\rm met}(\mathbf{r}|\mathbf{r}_0)$, for going from $\mathbf{r}_0$ to $\mathbf{r}$, where $\mathbf{r}_0 = (\theta_0, R_0)$ and $\mathbf{r} = (\theta, R)$ namely θ is the angle between the arms and R is the end-to-end distance of the bubble as schematically represented in Fig. 3.1d. The Mean First Passage Time (MFPT) is the mean time of a diffusing particle to hit a given target for the first time, and is the solution of the backward Smoluchowski equation [90]:

$$-\beta\nabla U \cdot \nabla[D\tau_{\rm met}(\mathbf{r})] + \Delta[D\tau_{\rm met}(\mathbf{r})] = -1 \tag{3.9}$$

where $U(\mathbf{r})$ is the potential in the coordinate $\mathbf{r}$, D is the diffusion constant. To begin with, we look at the case where two stiff rods are connected by a flexible hinge. The rotational diffusion of the arms can be seen as a particle diffusing over a spherical surface. Now we compute the mean first passage time for the rods starting at an angle θ and to reach a fixed angle $\theta_{\rm c}$ at the north pole, R being fixed, as shown in Fig. 3.13. The Eq. (3.9) becomes

$$\Delta\tau_{\rm met}(\theta,\phi) = \frac{1}{\sin(\theta)}\frac{\partial}{\partial\theta}\left(\sin(\theta)\frac{\partial\tau_{\rm met}(\theta,\phi)}{\partial\theta}\right) + \frac{1}{\sin^2(\theta)}\frac{\partial^2\tau_{\rm met}(\theta,\phi)}{\partial\phi^2} = -\frac{\mathcal{R}^2}{D_{\rm R}} \tag{3.10}$$

where $\mathcal{R}$ is the radius of the sphere. Since the problem has the azimuthal symmetry, one can integrate out over ϕ. We get

$$\frac{1}{\sin(\theta)}\frac{\partial}{\partial\theta}\left(\sin(\theta)\frac{\partial\tau_{\rm met}(\theta)}{\partial\theta}\right) = -\frac{\mathcal{R}^2}{D_{\rm R}} \tag{3.11}$$

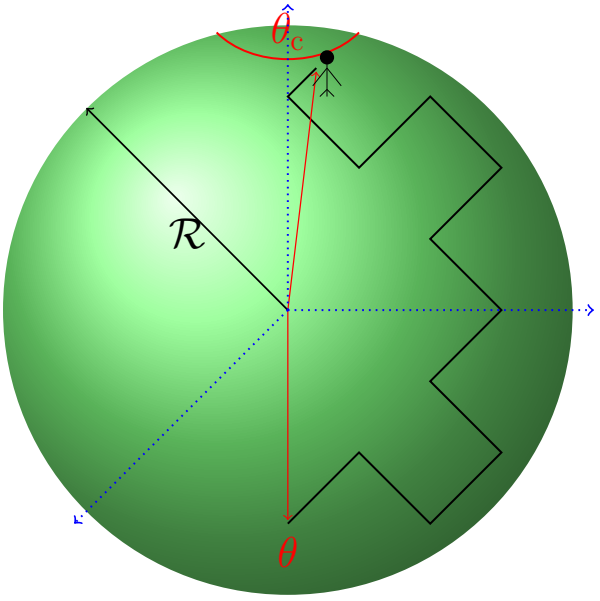

FIGURE 3.13: Sketch of a particle diffusing over a spherical surface of radius $\mathcal{R}$, starting at θ (polar angle) and reaching θ_c.

The above Eq. (3.11) can be solved either by direct integration or using Green function approach with the bounding conditions

$$\tau_{\text{met}}(\theta_c) = 0 \tag{3.12}$$

$$\left.\frac{d\tau_{\text{met}}(\theta)}{d\theta}\right|_{\theta=\pi} = 0 \tag{3.13}$$

The boundary condition Eq. (3.12) is the absorbing boundary condition and Eq. (3.13) is for the differentiability in θ. The solution is,

$$\tau_{\text{met}}(\theta) = \frac{\mathcal{R}^2}{D_{\text{R}}} \ln\left[\frac{1-\cos(\theta)}{1-\cos(\theta_c)}\right] \tag{3.14}$$

where D_{R} is the diffusion constant of the rod undergoing rotational diffusion, given in Eq. (3.8).

Next, we take the semiflexible nature of the bubble into account. The MFPT will slightly decrease as the bending rigidity of the bubble will favor $\theta = 0$ (aligned) state with the effective potential [100]

$$U(\theta) = \kappa_{\text{ss}}(1-\cos\theta)/\bar{L} \tag{3.15}$$

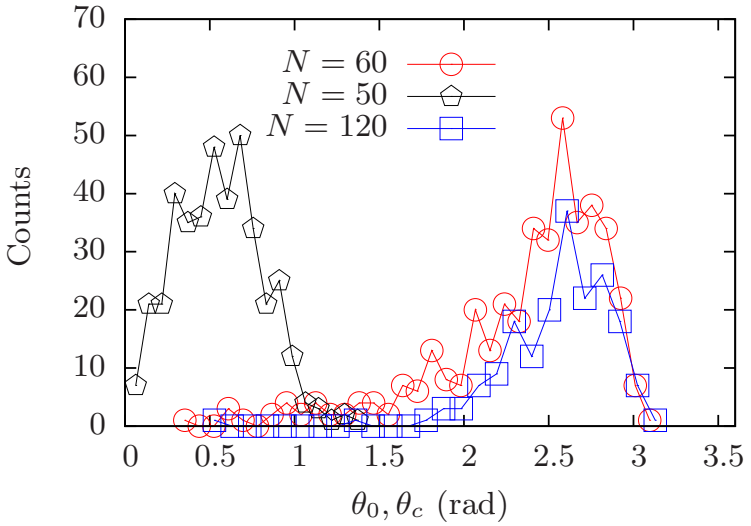

FIGURE 3.14: Distributions of angles, θ_0 and θ_c are shown. The red and blue lines belong to the θ_0 distribution, whereas the black line belongs to the θ_c distribution.

The Eq. (3.9) can be solved with the above potential Eq. (3.15). The solution is

$$\tau_{\rm met}(\theta_c|\theta) = \frac{\bar{L}}{2\ell_{\rm ss} D_{\rm R}} \left[t\left(\cos\theta, \frac{\ell_{\rm ss}}{\bar{L}}\right) - t\left(\cos\theta_c, \frac{\ell_{\rm ss}}{\bar{L}}\right) \right] \tag{3.16}$$

where

$$t(x,y) = \ln\left(\frac{1-x}{1+x}\right) + {\rm Ei}[-y(1+x)] - e^{-2y}{\rm Ei}[y(1-x)] \tag{3.17}$$

and ${\rm Ei}[z] = -\int_z^\infty \frac{e^{-t}}{t}{\rm d}t$ is the Exponential integral. When $\bar{L}/\ell_{\rm ss} \to \infty$, Eq. (3.17) simplifies to Eq. (3.14) as desired. In Fig. 3.14, we have shown the distribution of angles, θ_0 and θ_c extracted from the simulations, where θ_0 is the angle between arms at the end of zipping. θ_0 has a maximum around 2.5 rad, whereas θ_c has a maximum at smaller value around $0.3 - 0.7$ rad, depending on N. In Fig. 3.15, the dwell time distributions for different $\kappa_{\rm ss}$ are plotted. The dwell time distribution becomes wider and the mean value increases for smaller $\kappa_{\rm ss}$ values. The numerical mean dwell time is plotted along with Eq. (3.16) in the inset of Fig. 3.15 for $\theta_c = \pi/10 \simeq 0.3$ rad. The Eq. (3.16) reproduces the simulation data qualitatively, but not quantitatively. The reason could be that both diffusive processes in θ and R are in fact entangled. Solving the full equation is out of reach, but approximation could be used in future works.

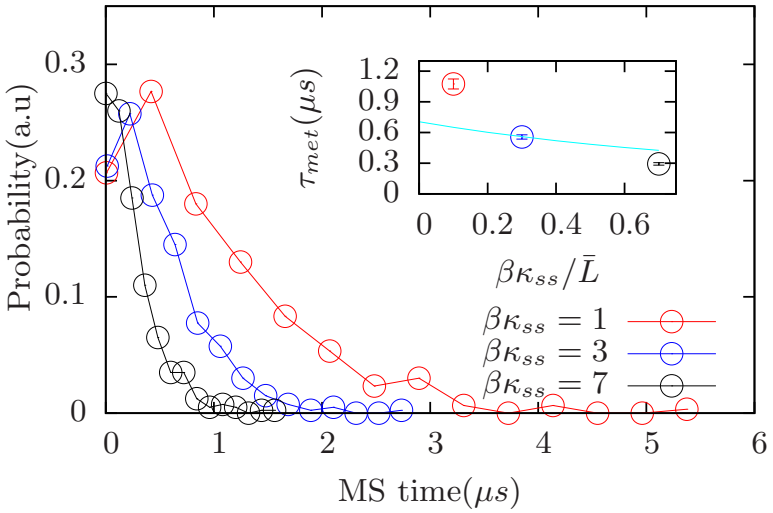

FIGURE 3.15: Closure time probability for $\beta\kappa_{\rm ss} = 1, 3$ and 7 (300, 600 and 400 samples respectively) and $N = 100$. The average value, $\tau_{\rm met}$, is plotted in the inset and clearly decreases with increasing $\kappa_{\rm ss}$. The solid line corresponds to Eq. (3.17).

3.3 Discussion

In this chapter, we studied the bubble closure dynamics in dsDNA using a simple coarse-grained ladder model. We showed that the closure of bubbles occurs *via* 2 consecutive steps, fast zipping followed by a slow closure of the metastable bubble which is further limited by the rotational diffusion of arms. We have also found the zipping rate to be $\approx 10^7$ bp/s. Imino proton exchange experiments on DNA have measured the zipping rate $\approx 10^8$ bp/s [101]. Note that in our case the zipping is not linear as the bubble finally falls in metastable stage which dominates the total closure time. We found closure times of 0.1 μs to 4 μs, which scale with the length of DNA as $N^{2.4}$. We have also shown the strong dynamic coupling of base-pair degrees of freedom with the elastic properties of the DNA chain.

Several models have been used to study the bubble dynamics or breathing dynamics of dsDNA in the literature. One of the extensively used models is the Poland-Scheraga model (PS model) [17, 18]. PS model is a one-dimensional Ising model with entropic penalty for creating a flexible loop (bubble) characterized by an exponent c. The further generalization of the model is done by including excluded volume interactions [19, 20, 21, 22]. The bubble breathing dynamics are studied using the Poland-Scheraga free energy either in discrete version [25] or continuous version [24, 48, 102]

using the Fokker-Plank equation approach. Another model used in breathing dynamics is the Peyrard-Bishop model which is a one dimensional non-linear model consisting of two complementary chains connected by a Morse potential representing hydrogen bonds [31]. The bubbles are created by the energy localization because of non-linear effects. The lifetime of bubbles in this model is on the order of few picoseconds [103]. These models only describe short breathing events, fast opening and closing of base-pairs. The above models did not consider the large bubbles and equilibrated bubbles in dsDNA. Above models also did not consider the conformational degrees of freedom which we have shown to play a crucial role in the closure of equilibrated bubbles.

Altan-Bonnet et *al* measured the timescales of denaturation bubbles of length 18 bps using florescence correlation spectroscopy [13] and found very long timescales, 20 μs to 100 μs. Bubble lifetimes of about 1 μs are also found the NMR measurements on short oligomers of DNA [54]. Even though we investigated the closure of large equilibrated bubbles, the dynamics will be the same (for a given N) for any equilibrated bubble of size greater than the metastable bubble size, $\bar{L}$, as the metastable residence time dominates the zipping time, $\tau_{\rm met} \gg \tau_{\rm zip}$. Even though we recovered large timescales, 0.1 μs to 4 μs, these timescales are smaller than the experimental ones. The gap between these timescales might be due to the missing helicity in the current model. Hence, the effect of twist will be studied in the next chapter and we will show that the twist is indeed the missing ingredient to account for the gap between experimental timescales and our estimates.

Chapter 4

Role of DNA twist in bubble closure

The closure dynamics of long equilibrated bubbles of dsDNA is studied using the helical model described in Chapter 2. The torsional elastic modulus depends on the base-pair (bp) state leading to stiff double helical DNA in double stranded state and 2 flexible chains in single stranded state. For DNA lengths $N = 40$ bps to 100 bps and initial bubble size of $N - 20$ bp, long closure times of 0.1 to 100 μs have been found. The bubble closes *via* two phases, zipping and metastable state such as in the previous ladder model. Moreover, the final closure in the metastable state is limited by mainly three mechanisms depending on the torsional modulus and the length of DNA:

- *Arms Diffusion Limited* closure, limited by the alignment of the two arms through rotational diffusion as in the previous model,
- or *Bubble Diffusion Limited* closure which is limited by the diffusion of the bubble towards one end of the DNA,
- or *Temperature Activated closure* limited by the crossing of an activation barrier.

4.1 Bubble closure dynamics

We followed the same procedure to create the bubble as we did for the ladder model described in Chapter 3. We start with an equilibrated dsDNA of length N bps. To create the bubble, we switch off the Morse potential in the middle of the dsDNA leaving 2 arms of length 10 bps on each side (around one pitch), and then equilibrate the bubble. The final configuration will be a flexible bubble sandwitched between the two stiff dsDNA arms. To study the bubble closure dynamics, we start closing the bubble by applying the Morse potential in the bubble region at $t = 0$. We stop the simulation as soon as the bubble closes for the first time. We have chosen 1.2 nm as a cut-off base-pair distance to define closed and open state. If the distance exceeds the cut-off value, the base-pair will be considered as an open state.

In a typical realization, the evolution of the bubble size $L(t)/L(0)$ (normalized by the initial bubble size $L(0)$), and the tangent-tangent product $\hat{\mathbf{n}}_\mathrm{i} \cdot \hat{\mathbf{n}}_\mathrm{e}$ of the arms where $\hat{\mathbf{n}}_\mathrm{i}$ and $\hat{\mathbf{n}}_\mathrm{e}$ are average tangent vectors of initial and end arms of DNA, are shown in Fig. 4.1 for $\beta\kappa_\phi = 200$ and $N = 60$ bps. In addition, we computed the mean twist per base-pair in the bubble region, because it plays an important role below. The twist angle $\phi_i(t)$ is computed from

$$\sin(\phi_i) = \left(\hat{\boldsymbol{\rho}}_i \times \hat{\boldsymbol{\rho}}_{i-1}\right) \cdot \hat{n} \tag{4.1}$$

where $\hat{n}$ is the helical axis and is computed from the mid points of corresponding base-pairs, and $\hat{\boldsymbol{\rho}}_i$ is the base-pair vector. The mean twist $\Delta\phi(t)$ in the bubble is defined as

$$\Delta\phi(t) = \frac{1}{L(t)} \sum_{i=i_0}^{i_0+L(t)-1} \phi_i(t) \tag{4.2}$$

where i_0 is the first monomer index of the bubble and $i_0 + L(t) - 1$ is the last one. From Fig. 4.1, it is clear that there are also two regimes as in the ladder model: (1) $L(t)/L(0)$ decreases rapidly thus defining the *zipping regime*; (2) $L(t)/L(0)$ remains almost constant with a bubble size, $L(t) \simeq \overline{L} \simeq 10$ bps, corresponding to the *metastable regime*. The mean twist in the bubble increases slowly with the tangent-tangent product. The bubble melting map of the same trajectory is also shown on the top of Fig. 4.1. We observe in Fig. 4.1 that the dsDNA falls into an hairpin configuration just after the zipping regime ($\hat{\mathbf{n}}_\mathrm{i} \cdot \hat{\mathbf{n}}_\mathrm{e} \simeq -1$). Thereby the closure of the metastable bubble is limited by the alignment of the dsDNA arms through rotational diffusion. This mechanism of closure is familiar with the ladder model. We recovered a similar mechanism of closure for the parameter

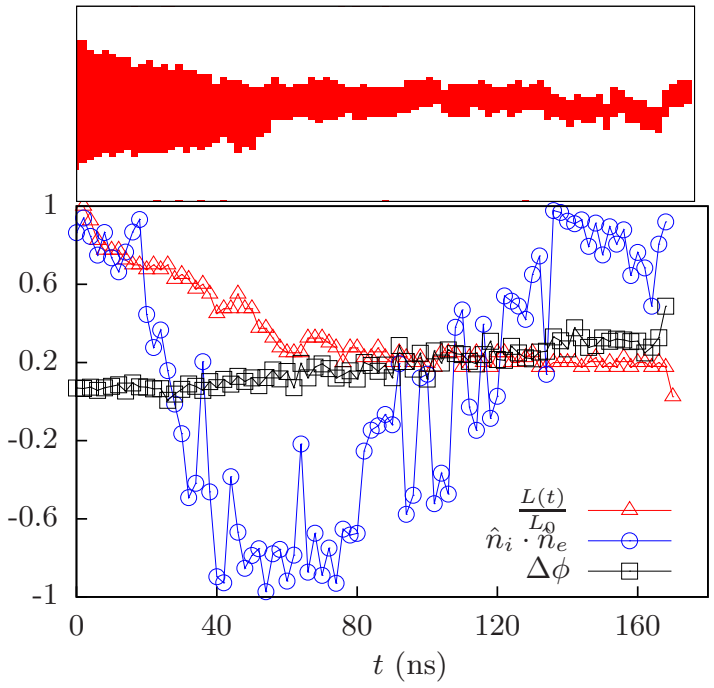

FIGURE 4.1: Evolution of the normalized bubble size, $L(t)/L(0)$ (red), the tangent-tangent product, $\hat{\mathbf{n}}_\mathrm{i} \cdot \hat{\mathbf{n}}_\mathrm{e}$ (blue), and the mean twist in bubble, $\Delta\phi(t)$ (black) for $N = 60$ bps, $\beta\kappa_\phi = 200$. The corresponding melting map is shown on top.

values ($\beta\kappa_\phi = 200$, $N = 60$ bps) even though the present model differs from the ladder model.

We have also shown a typical realization for $\beta\kappa_\phi = 300$ and $N = 60$ bps in Fig. 4.2. The normalized bubble size, the tangent-tangent product and the mean twist of base-pair in the bubble are plotted against time in Fig. 4.2, together with its melting map. The arms of dsDNA are already almost aligned after the zipping stage. The mean twist in the bubble is almost constant during the whole metastable state, $\Delta\phi(t) = \overline{\Delta\phi} \simeq 0.3$ rad. The final closure of the bubble, *i.e* the closure of the metastable bubble, is limited by the diffusion of the metastable bubble towards one end of the DNA. The opening at one of the ends of the DNA relaxes the torsional stress in the bubble which eventually leads to the immediate closure of the metastable bubble. In Fig. 4.2, as the bubble starts reaching one end, the tangent-tangent product undergoes large fluctuations and the size of the corresponding arm decreases, leading to a faster rotational diffusion. Indeed, the tangent-tangent product can not be well-defined once the bubble opens at one end.

We study the two regimes, the zipping regime and the metastable one in more details in the following sections.

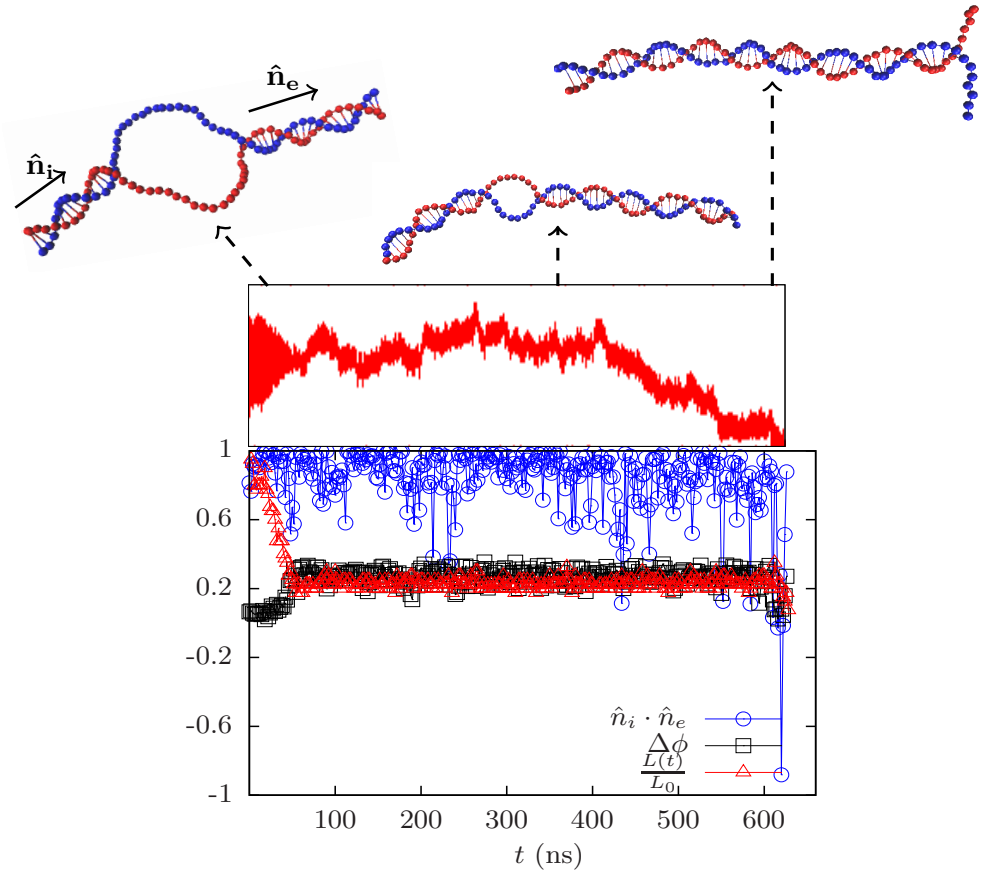

FIGURE 4.2: Same as Fig. 4.1, for $\beta\kappa_\phi = 300$. Snapshots at different times are shown.

4.1.1 Zipping regime

In the zipping regime, the initial flexible bubble starts closing from both DNA ends. In Fig. 4.3, we plotted the normalized bubble size, $L(t)/L(0)$, against time for $N = 70$ bps and $N = 100$ bps. The data are well fitted by a linear law, but at short timescales only. As the bubble size reaches the order of persistence length of ssDNA, $\ell_{\rm ss}$, the bubble size starts saturating and finally stops closing further, leading to the metastable bubble of size $\bar{L}$. As in the ladder model, we have chosen the cut-off length, $P = \frac{3}{5}(L(0) - \bar{L})$ to distinguish the pure zipping regime from the crossover to the metastable state. In Fig. 4.4, we plotted the zipping time *vs* P for three different values of $\beta\kappa_\phi$. The zipping time, $\tau_{\rm zip}$ is found to scale with P with an anomalous scaling exponent, $\tau_{\rm zip} \sim P^\gamma$, where $\gamma = 1.4 - 1.6$. The variation of this exponent could be due to the definition of P as we have chosen P as an arbitrary cut-off. The zipping time, $\tau_{\rm zip}$, also scales with the same exponent as found in polymer translocation [93, 95] and renaturation dynamics [94]. We have also simulated helical DNA renaturation and found that the dynamics is different

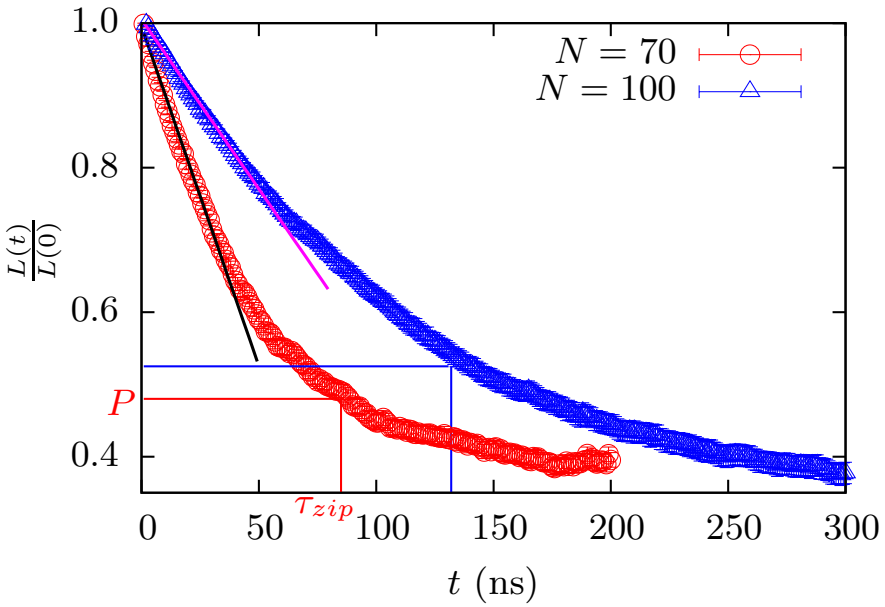

FIGURE 4.3: The bubble size normalized by initial bubble size $L(0)$ against time for $\beta\kappa_\phi = 300$. The black and magenta lines are linear fits at short times.

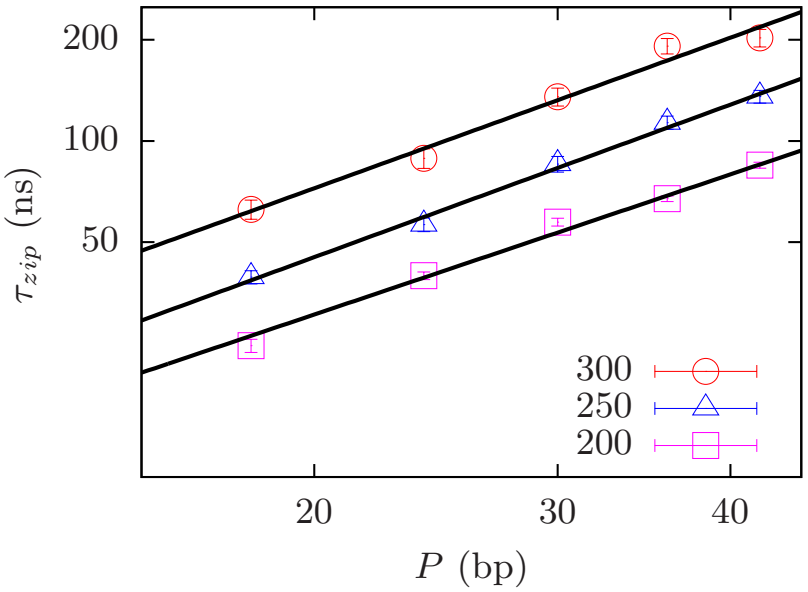

FIGURE 4.4: Log-log plot of zipping time, τ_{zip}, vs P for three different values of $\beta\kappa_\phi$. Data are fitted (black lines) by P^γ, where $\gamma \sim 1.4 - 1.6$.

from zipping dynamics unlike in the ladder model in which zipping and renaturation dynamics were the same (for more details, see Appendix A).

The zipping can also be understood with simple scaling arguments. The initial flexible bubble starts winding from the ends by the driving torque, $\mathcal{T}$, acting on each arms. The driving torque plays the same role as the driving force in the Ladder model, coming from

the energy gain when the base-pair closes which is, $\mathcal{T} \simeq A/\phi_{\rm eq} \simeq 16\ k_{\rm B}T$/rad. The zipping scenario can be compared to wrapping a chain around two cylinders (provided the arms are stiff which is the case here) as shown in Fig. 4.7. The arms can be considered as 2 growing and spinning cylinders. The evolution of the total twist, $\phi(t)$ can be written as

$$2\zeta(t)\frac{{\rm d}\phi(t)}{{\rm d}t} = 2\mathcal{T} - \kappa_\phi\phi_{\rm el} - \mathcal{T}_{\rm bend}(t) + \mathcal{R}(t) \tag{4.3}$$

where $\phi(t) = \phi_1(t) - \phi_2(t)$, where $\phi_1(t)$, $\phi_2(t)$ are the twists of each arm of length $\ell(t)$ and radius ρ_0 and $\zeta(t) = \alpha\pi\rho_0^2\ell(t)$ is the rotational friction of each arm which depends on the time as the size of the arm increases with time. The prefactor α comes from the geometry of the arm, for an ellipsoid $\alpha = \frac{8}{3}$, for sphere $\alpha = 8$ and for this case (almost rigid cylinder) α is fitted below. The torque $\mathcal{T}_{\rm bend}(t)$ comes from the bending energy of two ssDNA in the bubble and plays a significant role only when the two ssDNAs are stiff enough. The second term in Eq. (4.3) is the elastic restoring torque which comes from the fact that the κ_ϕ profile is non-zero at the ends of the bubble (unlike in the middle of the bubble where it is zero, see Fig. 4.13) and thus depends on κ_ϕ. The small but non-zero twist, $\phi_{\rm el}$, is the average twist stored inside the bubble (the bubble is semi-flexible during the zipping regime also). The last term in Eq. (4.3), $\mathcal{R}(t)$ is the random torque. As the bubble size decreases, the length of the arms and the total twist increases according to $\phi(t) = \phi(0) + \frac{2\pi}{p}\left(L(0) - L(t)\right)$, where p is the pitch of dsDNA and $\phi(0)$ is the initial twist of each arm, taken to be 0 at $t = 0$.

First we look at the evolution of $\phi(t)$ as long as the bubble is flexible enough, in which case the torque $\mathcal{T}_{\rm bend}$ can be neglected. As soon as the bubble size reaches the order of the persistence length of ssDNA, the torque $\mathcal{T}_{\rm bend}$ plays a role. To understand the evolution of $\phi(t)$, we need to estimate two parameters, α, the prefactor of rotational friction and $\phi_{\rm el}$, the small twist angle appearing in the elastic restoring torque. We computed the angular velocity $\frac{{\rm d}\phi(t)}{{\rm d}t}$ at $t = 0$ for three different values of κ_ϕ. The data are fitted with a straight line as shown in Fig. 4.5. One can estimate the angular velocity from Eq. (4.3).

$$\omega \equiv \frac{{\rm d}\phi(t)}{{\rm d}t} \simeq \frac{2\mathcal{T} - \kappa_\phi\phi_{\rm el}}{2\zeta(t)} \tag{4.4}$$

By comparing the fitted linear law in Fig. 4.5 and Eq. (4.4), the values of α and $\phi_{\rm el}$ are found to be 6.15 and 0.064 rad respectively. The velocity of zipping, v becomes

$$v(0) = \frac{p\omega(0)}{2\pi} = \begin{cases} 0.82 \text{ bps/ns}, & \beta\kappa_\phi = 200 \\ 0.51 \text{ bps/ns}, & \beta\kappa_\phi = 300 \end{cases} \tag{4.5}$$

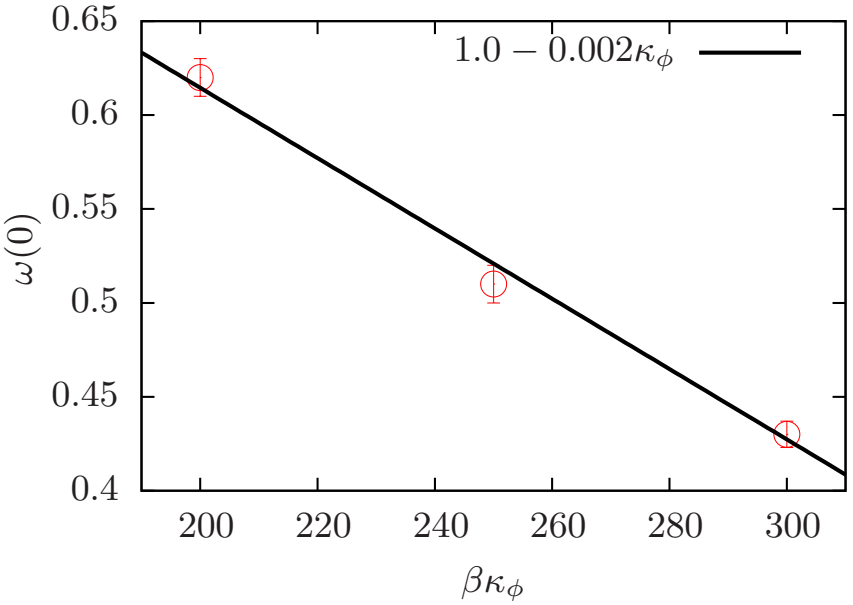

FIGURE 4.5: The angular velocities of $\phi(t)$ are plotted against κ_ϕ. The data are fitted by a straight line, $\omega(\kappa_\phi) = 1.0 - 0.002\kappa_\phi$.

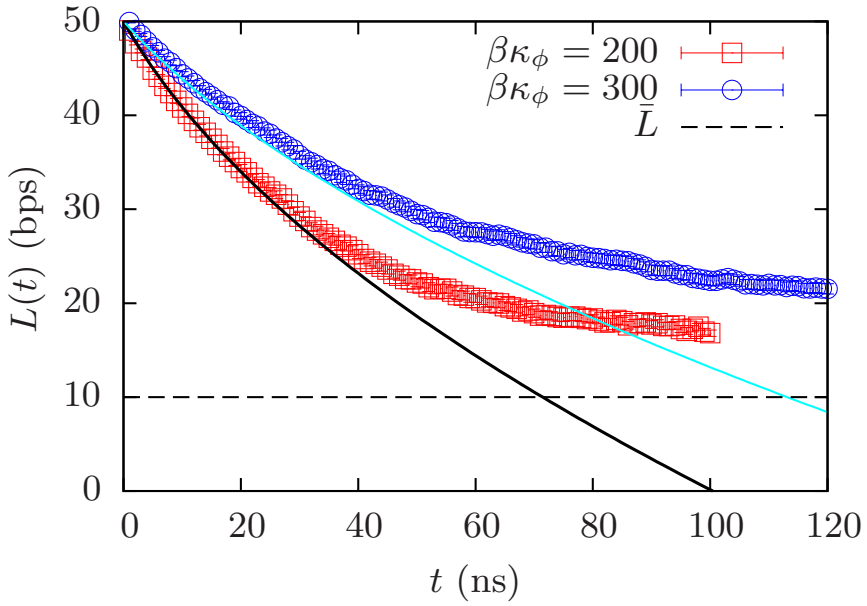

FIGURE 4.6: The bubble size, $L(t)$ against time for $\beta\kappa_\phi = 200$, $\beta\kappa_\phi = 300$ and for $N = 70$ bps. The black and cyan lines are given by Eq. (4.9).

One can rewrite Eq. (4.3) as

$$2\zeta(t)\frac{\mathrm{d}\phi(t)}{\mathrm{d}t} = 2\mathcal{T}_{\mathrm{eff}} \tag{4.6}$$

where $2\mathcal{T}_{\mathrm{eff}} = 2\mathcal{T} - \kappa_\phi\phi_{\mathrm{el}}$. As the bubble size decreases, the arm size increases thereby

increasing the rotational friction of the arms. The evolution of $\phi(t)$ (or $\ell(t)$) is

$$b\frac{\mathrm{d}\ell(t)}{\mathrm{d}t} = \frac{2p\mathcal{T}_{\mathrm{eff}}}{2\pi\ell(t)} \tag{4.7}$$

where $b = \alpha\pi\eta\rho_0^2$. After integration

$$\frac{\ell(t)^2}{2} = \frac{\ell(0)^2}{2} + \frac{2p\mathcal{T}_{\mathrm{eff}}}{2\pi b}t \tag{4.8}$$

Since $\ell(t) = \ell(0) + (L(0) - L(t))$, the equation becomes

$$L(t) = L(0) + \ell(0) - \sqrt{\frac{2p\mathcal{T}_{\mathrm{eff}}}{b\pi}t + \ell(0)^2} \tag{4.9}$$

It is linear in t at short times which is contrary to the case of Ladder model. In Fig. 4.6, Eq. (4.9) is plotted along with the bubble size for $\beta\kappa_\phi = 200$ and $N = 70$ bps and for $\beta\kappa_\phi = 300$. Eq. (4.9) fits well for $t \leq 50$ ns. As soon as the bubble size starts reaching the order of ℓ_{ss}, the torque $\mathcal{T}_{\mathrm{bend}}$ starts competing with the effective driving torque, $\mathcal{T}_{\mathrm{eff}}$, thus saturating at a metastable bubble size, $\bar{L} \simeq 10$ bps. The torque, $\mathcal{T}_{\mathrm{bend}}$, is solely coming from the two ssDNAs. The bending energy of these two ssDNAs gives rise to a small, but non-zero torsional coefficient, κ_{ss}. This small, but non-zero value of κ_{ss} starts competing with other torques in Eq. (4.3) when the length of two ssDNAs reaches the order of their persistence lengths, ℓ_{ss}. The steady state solution of Eq. (4.3) gives the metastable bubble size, thus giving rise to a highly bent and under-twisted metastable bubble of size, $\bar{L} \simeq 10$ bps and $\overline{\Delta\phi} \simeq 0.3$. Finally, we have not studied the effect of random torque, $\mathcal{R}(t)$, as we are only interested in the mean deterministic behaviour of the system.

4.1.2 Metastable regime

The final closure of the bubble is defined as the end of the metastable regime. The size of the bubble in the metastable state is observed to be around 10 bps. Unlike in the earlier ladder model, rotational diffusion of the arms is not the only limiting step for closure in the current model. The tangent-tangent product between the arms just after the zipping regime is computed as shown in the Fig. 4.2. We recall that the average tangent vectors, $\hat{\boldsymbol{n}}_{\mathrm{i}}$ and $\hat{\boldsymbol{n}}_{\mathrm{e}}$ of both the arms are computed from the base-pair centers of mass. For small ratios of $L(0)/\ell_{\mathrm{ss}}$, around $3-4$, the angle between the arms after zipping is either in the hairpin configuration ($\cos\theta \approx -1$) or in the aligned configuration ($\cos\theta \approx 1$) irrespective

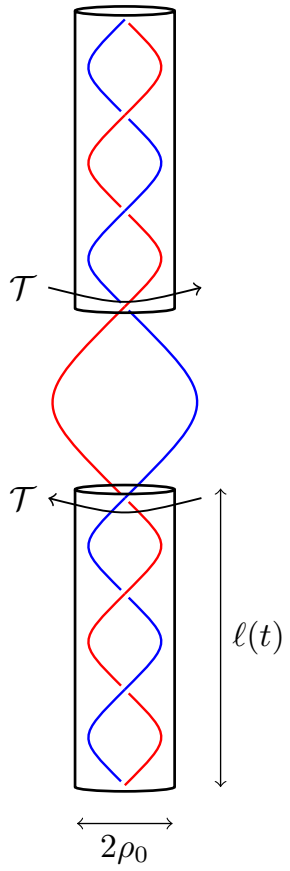

FIGURE 4.7: Sketch of the model DNA going through zipping with driving torque, $\mathcal{T}$ acting at each end of the bubble in opposite directions. The length of the arm is $\ell(t)$ and $2\rho_0$ is the diameter of the arm modeled by a cylinder.

of κ_ϕ. As the ratio, $L(0)/\ell_{\rm ss}$ increases, the distribution of angles between the arms just after the zipping stage shifts towards an hairpin configuration irrespective of $\beta\kappa_\phi$ value. It means that for size DNAs, both the configurations (hairpin and aligned ones) exist whereas for logn DNAs only hairpin configuration exists.

Furthermore, in Fig. 4.8, we have shown the histogram for the final position of the bubble just before closure. In case of $\beta\kappa_\phi = 200$, apart from roughly 25% of the realizations, the bubble closes without reaching one end of the DNA. For remaining 75%, the bubble closes in the middle, i.e without reaching one end through diffusion. So, bubble diffusion does not play a dominant role in the case of $\beta\kappa_\phi = 200$. As also shown in Fig. 4.1, the final closure is essentially limited by the rotational diffusion of the arms as in the ladder model. We call this mechanism **Arms Diffusion Limited** (ADL) closure. We have plotted the mean squared displacement (M.S.D) of the bubble for $\beta\kappa_\phi = 200$ in Fig. 4.9. We also plotted in the same figure the M.S.D of the bubble in the case of the ladder model for $N = 70$ bps. We see that both M.S.Ds are comparable at short times. For $\beta\kappa_\phi = 300$, the alignment of the arms is not the limiting step of the final closure of the bubble. The

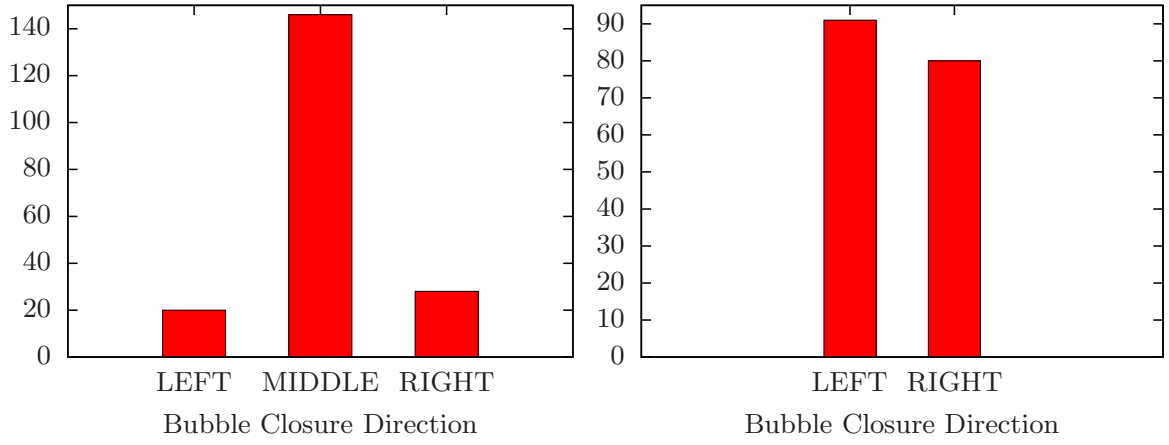

FIGURE 4.8: Histogram of the position of the bubble at the final closure for $\beta\kappa_\phi = 200$ (left) and $\beta\kappa_\phi = 300$ (right) and $N = 60$ bps.

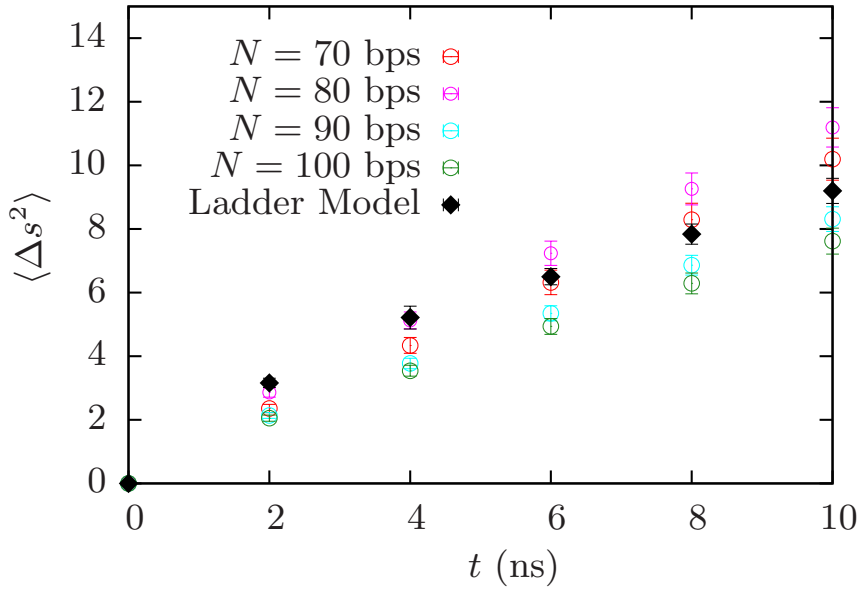

FIGURE 4.9: Bubble mean square displacement along the DNA vs time for $\beta\kappa_\phi = 200$ and for the ladder model, $N = 70$ bps (black diamonds).

limiting step is the diffusion of the bubble towards one end of the DNA. Thereby the bubble opens up one end of the DNA to relax the torsional stress, and closes as shown in Fig. 4.2. This limiting step does not depend on whether the DNA falls to an hairpin configuration after zipping or not. So, the final bubble closure in case of $\beta\kappa_\phi = 300$ is limited by the diffusion of the bubble towards one end of the DNA. From Fig. 4.8, it is also clear that the bubble closes systematically by opening one end of the DNA. We call this mechanism of closure the **Bubble Diffusion Limited** (BDL) closure. Note that we did

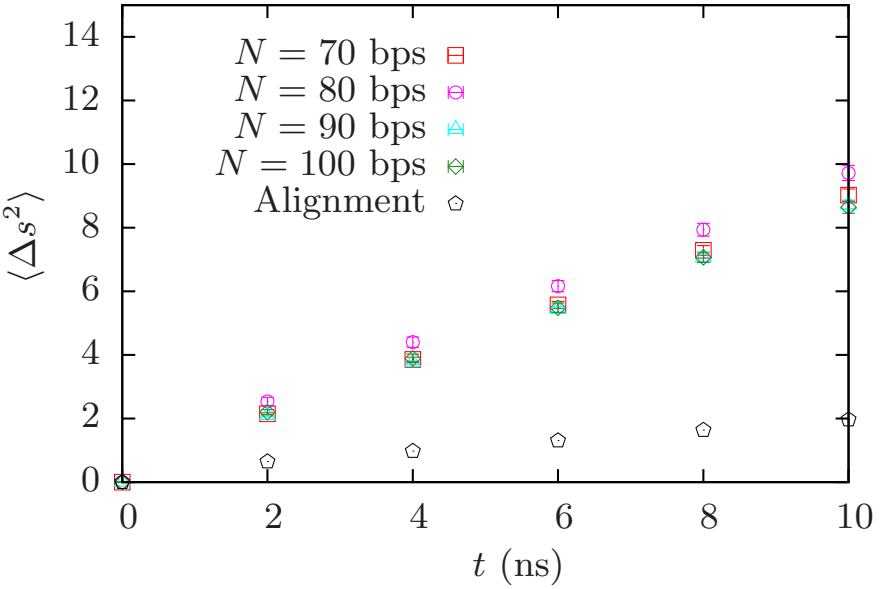

FIGURE 4.10: Bubble mean squared displacement along the DNA for $\beta\kappa_\phi = 300$ and various N values and the same is shown for $N = 100$ bps during the alignment of the arms.

not observe this mechanism for Ladder model for chosen parameter values and the reason could be that the ends of model DNA in Ladder model are clamped which prevents the bubble from opening at one end of the DNA. In Fig. 4.10, the mean squared displacement (M.S.D) of the bubble along the DNA is shown for different N values and for $\beta\kappa_\phi = 300$. The data show that the M.S.D of the bubble does not depend significantly on N. This is due to the fact that the diffusion of the bubble along the DNA is only a local phenomenon. However, the diffusion of the bubble along the DNA is limited by the alignment of arms. In Fig. 4.10, we have also shown the M.S.D of the bubble for $N = 100$ bps and $\beta\kappa_\phi = 300$ for two regimes: (1) when the DNA is going from hairpin configuration to aligned one (⬠). (2) From the aligned configuration to the final closure(◇). It shows that the bubble diffusion is much faster when the DNA is aligned than when it is in hairpin configuration. The geometry of the hairpin configuration tends to localize the bubble in DNA. It means that diffusion of the bubble is constrained by the geometry of the whole DNA.

We now focus on the twist angle, ϕ. We found $\Delta\phi \simeq 0.30$ rad for the metastable bubble whatever the $\beta\kappa_\phi$ value. In Fig. 4.11, we presented the 2-dimensional histogram of the angle between the arms, $\cos(\theta_{\rm bc})$, and the mean twist of the bubble $\Delta\phi_{\rm bc}$ just before the final closure, for $\beta\kappa_\phi = 200$. As one can notice, the alignment of the arms and the non-zero value of twist (around 0.3 rad) is the necessary condition for the final closure of the bubble. Note that for $\beta\kappa_\phi = 300$, both $\Delta\phi_{\rm bc}$ and $\cos(\theta_{\rm bc})$ are ill-defined as BDL (opening at either end) is the dominating mechanism.

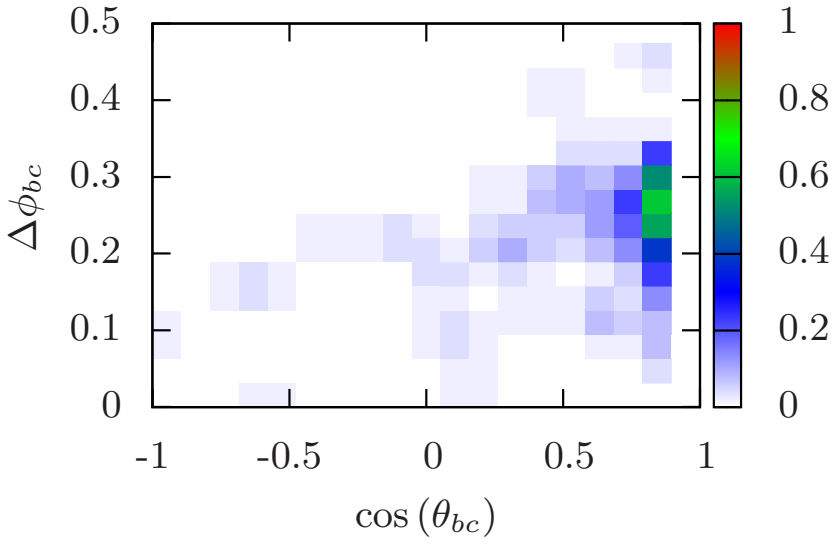

FIGURE 4.11: Two dimensional histogram of tangent-tangent product just before closure $\cos(\theta_{\rm bc})$ and mean twist in the bubble just before closure $\Delta\phi_{\rm bc}$ for $\beta\kappa_\phi = 200$. The colorscale gives the density of the realizations falling in particular range of values. The total number of realizations is 200. $\Delta\phi_{\rm bc}$ is given in radians.

The bubble in the metastable state has a finite torsional energy even though we have chosen explicitly the torsional modulus in the ssDNA state, $\beta\kappa_{\phi,ss} = 0$. This non-zero twist arises from the bending of the two single strands and the constraints at each end. We investigated the dependence of this finite non-zero twist, $\Delta\phi$ with elastic properties of the single strands. The only parameters of ssDNA that enters in the simulations are $\theta_{\rm eq}$ which alters the persistence length of ssDNA, and the stretching constant κ_s. We performed additional simulations to check the dependence of $\Delta\phi$ on these parameters θ_0 and κ_s. We present our results before discussing them. We took a snapshot of a metastable bubble for $N = 60$ bps and $\beta\kappa_\phi = 300$. First we switched off the Morse potential inside the bubble of size 10 bps. We varied the temperature from $\Theta = 1.0$ to $\Theta = 0.0$, where $\Theta = T/T_0$ is the dimensionless temperature and $T_0 = 300$ K. Every 30 ns, we decreased the temperature by 0.06. The value of $\Delta\phi$ did not change from 0.3. To know the dependence of $\Delta\phi$ on θ_0, we changed the value of θ_0 for the above snapshot keeping the temperature 0 inside the bubble and 1 for the arms. As already mentioned earlier, θ_0 changes the persistence length. We varied θ_0 from 0.3 rad to 0.55 rad and found a linear law, $\Delta\phi = 0.53\theta_0 + 0.06$. We also changed $\beta\kappa_{\rm s}$ to 40 (for $\theta_0 = 0.41$ rad) to see the variation of $\Delta\phi$ in the bubble. The mean twist was found to be $\Delta\phi = 0.35$. This means that κ_s also changes $\Delta\phi$. This might be because of the coupling between

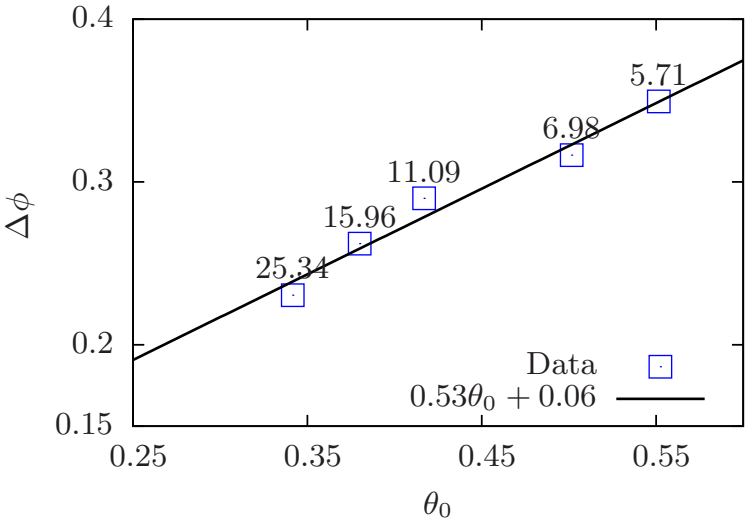

FIGURE 4.12: Variation of mean twist in the bubble $\Delta\phi$ vs bending angle θ_0. Labels for each point represent corresponding values of $\ell_{\rm ss}$. Both θ_0 and $\Delta\phi$ are given in radians.

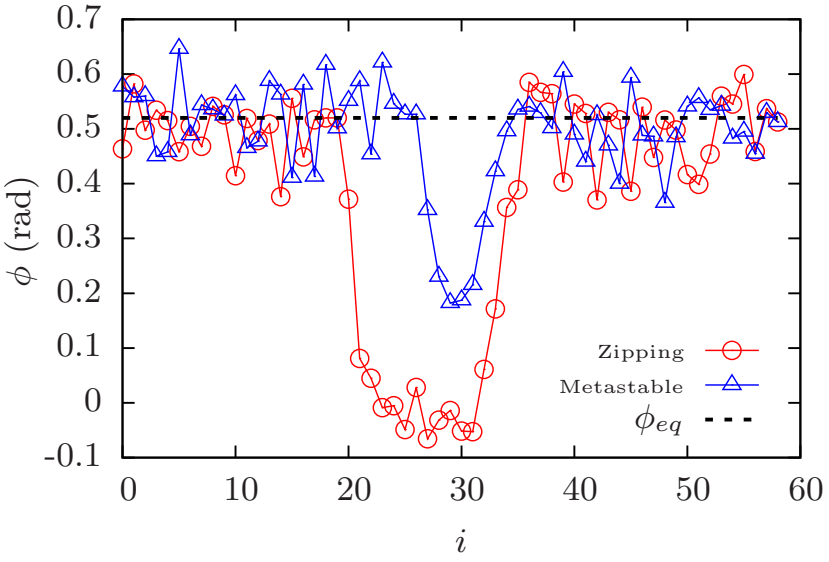

FIGURE 4.13: Twist profiles inside the DNA for $N = 60$ bps during the zipping regime and in the metastable regime for realization as in Fig. 4.2.

stretching and twisting [104]. It is also worth noting that the twist angles in the bubble do not have a uniform distribution along the bubble. In Fig. 4.13, the twist profile in the zipping regime and in the metastable one is shown for the typical realization of Fig. 4.2. The twist profile in the zipping regime has $\phi_i = 0$ in the middle and non-zero twist at the ends of the bubble, whereas the twist profile in the metastable regime has non-zero twist everywhere. Here note that the twist ϕ_i of $i^{\rm th}$ base-pair in the middle of the bubble is non-zero even though $\beta\kappa_{\phi_i}$ is zero. The twist profile in the zipping regime makes the rotation

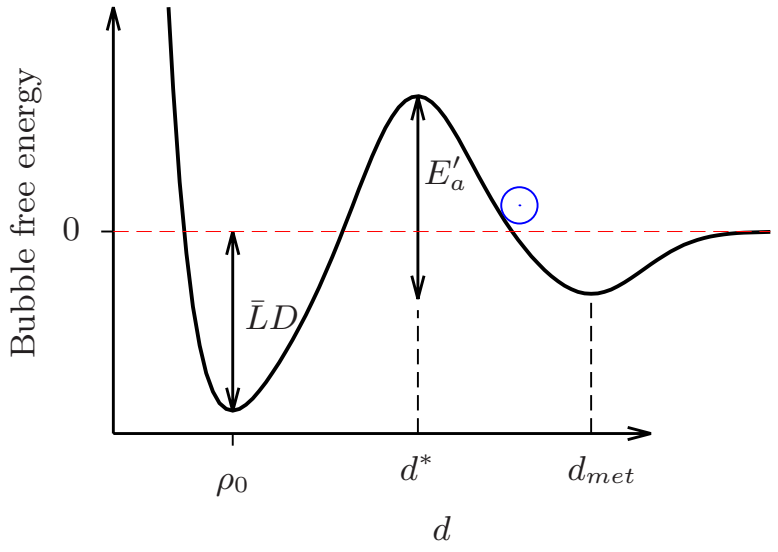

FIGURE 4.14: Sketch of a particle (in blue) crossing an activation energy barrier of height E'_a and $\bar{L}D$ is the energy gain. The reaction coordinate, d, is the distance between the strands. The local minima having reaction coordinate, $d_{\rm met}$, corresponds to metastable state and d^* corresponds to energy barrier.

of the two arms completely uncorrelated, whereas in the metastable regime, the further rotation of the arms (thus, closure of the bubble) becomes correlated. This cooperative nature of the bubble has two contributions:

- Free energy change associated with the two single strands.
- Torsional energy stored in the bubble.

This cooperative nature in further closure of the bubble is the reason for the metastable state. Closure of the bubble occurs when the bubble relieves this energy barrier by opening one end in the BDL mechanism or when the bubble escapes locally from this energy barrier which is analogous to the escape of a particle from an energy wall as sketched in Fig. 4.14. In Fig. 4.14, we show a diffusive particle crossing an activation energy barrier of height $E'_{\rm a}$ against the reaction coordinate, d, the distance between the strands, which is analogous to the activation energy given by Eq. (4.10), and $\bar{L}D$ is the gain in energy after crossing the barrier which is analogous to energy gain for the bubble in a Morse potential where $\bar{L}$ is the metastable bubble size. Closure of the bubble by crossing this energy barrier is another type of closure mechanism that we observed in simulations. We call this mechanism of closure **Temperature Activated** (TA) closure which is dominant roughly in the range of $220 < \beta\kappa_\phi < 260$.

We write the activation energy barrier as:

$$E'_{\rm a} = 2\Delta F_{\rm ss} + E_{\rm tor} \tag{4.10}$$

$$E'_{a} = 2\Delta F_{ss} - \bar{L}|V_{\rm morse}(d^*)| + \frac{1}{2}\sum_{bubble} \kappa(d^*)(\phi_i(d^*) - \phi_0)^2 \tag{4.11}$$

In Eq. (4.10), the first term is the free energy (stretching and bending) change associated with two single strands, the second term is the energy associated with Morse potential and third term is the torsional contribution associated with the twist, where $\kappa(d^*)$ is proportional to κ_ϕ (but smaller) and $\phi_i(d^*)$ is the average twist angle at the barrier (d^* is the average distance between strands at the barrier).
The energy barrier in the case of $\beta\kappa_\phi = 300$ ($E_{\rm a} \simeq 12\ k_{\rm B}T_0$ [1]) becomes higher than the gain $8\ k_{\rm B}T_0$ (by hydrogen bonding) whereas in the case, $\beta\kappa_\phi = 200$ ($E_{\rm a} \simeq 3\ k_{\rm B}T_0$), it is smaller than the gain. Hence for $\beta\kappa_\phi = 300$, the bubble relaxes the torsional stress and bypasses the barrier by opening one end of the DNA, whereas for $\beta\kappa_\phi = 200$, the bubble can close without reaching one end of DNA as the torsional energy contribution to the energy barrier ($E_{\rm tor}$) is almost negligible.

4.2 Three different closure mechanisms and "Phase diagram"

So far, we have identified three mechanisms of bubble closure. All three mechanisms of closure exist, but the fastest one, for given parameters, N and κ_ϕ, and boundary conditions, clamped or unclamped, will dominate. The three closure mechanisms are

1. Arms Diffusion Limited (**ADL**) which occurs when the torsional energy contribution to the energy barrier is negligible, in which case closure is only limited by the rotational diffusion of arms.

2. Bubble Diffusion Limited (**BDL**) which occurs for homopolymer DNAs for which bubble diffusion time along the DNA is shorter than the Temperature Activated one.

3. Temperature Activated closure (**TA**) which occurs for long DNAs or short DNAs for which ends are clamped (heteropolymer DNAs) [13].

[1] $E_{\rm a}$ denotes the measured activation energy barrier, whereas $E'_{\rm a}$ is the theoretical one given in Eq. (4.10).

We describe in further detail their respective features in the following subsections before determining which one dominates for given parameter values and boundary conditions.

4.2.1 Arms Diffusion Limited closure

In Fig. 4.1, by looking at $\hat{\boldsymbol{n}}_{\rm i}\cdot\hat{\boldsymbol{n}}_{\rm e}(t)$ the closure of the bubble for $\beta\kappa_\phi = 200$ and $N = 60$ bps is limited by the rotational diffusion of the arms and occurs as soon as the arms are almost aligned. ADL dwell times for $\beta\kappa_\phi = 200$ are plotted against the length of DNA, N in Fig. 4.15 (◇). The fit leads to the scaling law, $\tau_{\rm met}^{\rm ADL} \sim N^{2.2}$. As we argued in Chapter 3, the scaling estimate depends on the ratio $M/\ell_{\rm ds}$, and when M gets close to the value of $\ell_{\rm ds}$, the time saturates as the bending fluctuations of the arms keep the effective stiff arm size to be on the order of $\ell_{\rm ds}$.

4.2.2 Bubble Diffusion Limited closure

Bubble diffusion limited closure starts dominating for $\beta\kappa_\phi \gtrsim 220$ and DNA lengths up to $N = 100$ bps. The exact range of parameters will be discussed later in greater detail. In Fig. 4.10, the M.S.D of the bubble along the DNA is plotted for various values of N. The diffusion constant of the bubble is found to be $D \simeq 0.85$ bp^2/ns, for any N. As the final closure of the bubble is limited by the diffusion of the bubble to one end of the DNA, the mean first passage time for bubble closure in BDL mechanism is expected to be:

$$\tau_{\rm met}^{\rm BDL} \simeq \frac{(N/2)^2}{2D} \simeq 0.15\, N^2 \text{ ns} \tag{4.12}$$

Total closure times of denaturation bubbles along with metastable residence times (dwell times) for $50 \leq N \leq 100$ are shown in Fig. 4.15 for $\beta\kappa_\phi = 300$. The data for BDL dwell time is fitted with a power law and BDL dwell time scales with the length of DNA, N, as $\tau_{\rm met}^{\rm BDL} \approx 0.06\, N^{2.3}$ ns. The total closure time also scales with the same exponent, $\tau_{\rm closure} \sim N^{2.3}$. The theoretical prediction given in Eq. (4.12) is also shown in Fig. 4.15 which qualitatively explains the simulation data.

The slight disagreement in the prefactors between the simulation data and theoretical prediction is due to the slightly different exponent. The slight change in the exponent is due to the fact that the dwell time is a combination of ADL dwell time and BDL dwell time, $\tau_{\rm met} = \tau_{\rm met}^{\rm ADL} + \tau_{\rm met}^{\rm BDL} \sim \tau_{\rm met}^{\rm BDL}$. Even though the ADL dwell time is smaller in magnitude compared to that of the BDL dwell time, it can explain the different exponents.

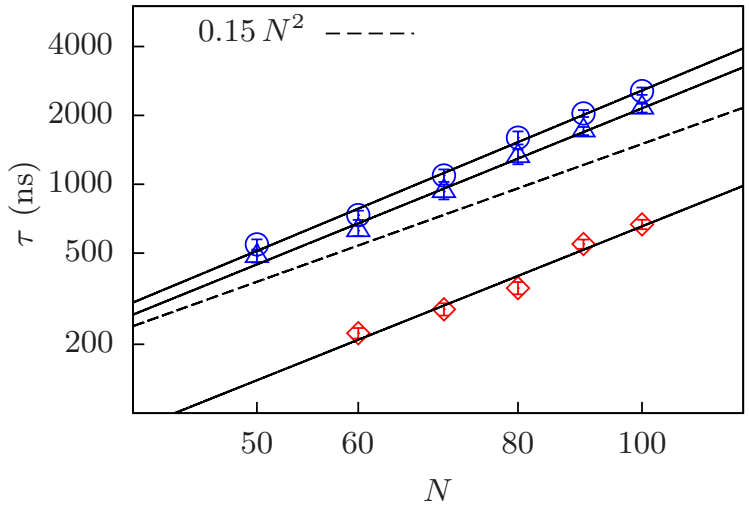

FIGURE 4.15: Log-log plot of BDL dwell (△) and closure times (○) ($\beta\kappa_\phi = 300$), and ADL dwell times (◇) ($\beta\kappa_\phi = 200$) with fitted laws. The dotted line is the theoretical prediction for BDL dwell time.

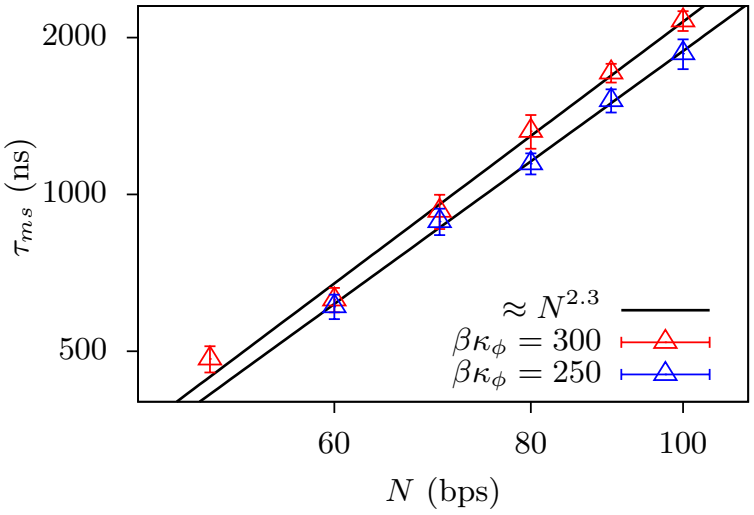

FIGURE 4.16: Log-log plot of BDL times vs the length of DNA, N, for two different torsion constants, $\beta\kappa_\phi = 300$ and $\beta\kappa_\phi = 250$.

The BDL dwell time slightly depends on the value of the torsion constant, κ_ϕ. But, it does not change the scaling law, only the prefactor. In Fig. 4.16, we show the BDL dwell times for two values of $\beta\kappa_\phi = 300$ and $\beta\kappa_\phi = 250$. The BDL dwell time for $\beta\kappa_\phi = 250$ scales with the same exponent as for $\beta\kappa_\phi = 300$, but with slightly higher prefactor, $\tau_{\text{met}}^{\text{BDL}} \simeq 0.075 N^{2.3}$ ns. This could be due to a slight increase in the diffusion constant as the decrement in κ_ϕ might increase the winding and unwinding rates at the bubble extremities, thereby leading to faster diffusion.

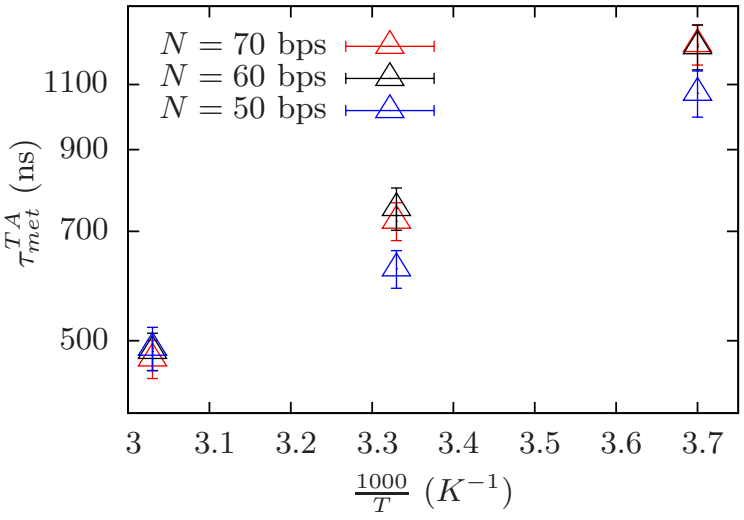

FIGURE 4.17: Arrhenius plot: Linear-log plot of the TA dwell times vs $1/T$ for clamped ends (TA mechanism) for three values of N. The torsion constant is $\beta\kappa_\phi = 220$.

4.2.3 Temperature Activated closure

To study the temperature activated closure in detail, we investigated the bubble closure dynamics for clamped DNA. Note that for clamped DNAs, the BDL mechanism of closure is inhibited as clamping the ends prevents the bubble from relaxing the torsional stress at the ends, and forces the bubble to close in the middle. We clamped the ends (of length 10 bps) with Morse potential of increased depth, $\frac{3A}{2}$. We have chosen the torsion constant slightly higher than $\beta\kappa_\phi = 200$ to reduce the occurrence of realizations belonging to the ADL case. TA dwell times are plotted against inverse temperature, $1/T$ for $\beta\kappa_\phi = 220$ and for various values of N in Fig. 4.17.

This Arrhenius plot allows us to estimate the activation energy, E_a, assuming that the closure occurs when the bubble crosses the activation energy barrier, see section 4.1.2. As the barrier crossing is a local mechanism, TA dwell times should not depend on N, which is supported by Fig. 4.17. TA dwell times $\tau_\mathrm{met}^\mathrm{TA}$, required in crossing the activation energy barrier, are written as:

$$\tau_\mathrm{met}^\mathrm{TA} = \tau_0 \exp\left(-\frac{E_\mathrm{a}}{k_\mathrm{B}T}\right) \tag{4.13}$$

where E_a is the activation energy. In Fig. 4.18, activation energies extracted from Fig. 4.17 are plotted against $\beta\kappa_\phi$. We found that the activation energy, E_a, scales affinely with κ_ϕ, $E_a = -17.0 + 0.10\,\kappa_\phi$. The torsional contribution to activation energy, E_tor, is also linear with κ_ϕ as given by Eq. (4.10). Even though $E_\mathrm{a} \neq E_\mathrm{tor}$, they are roughly equal as the

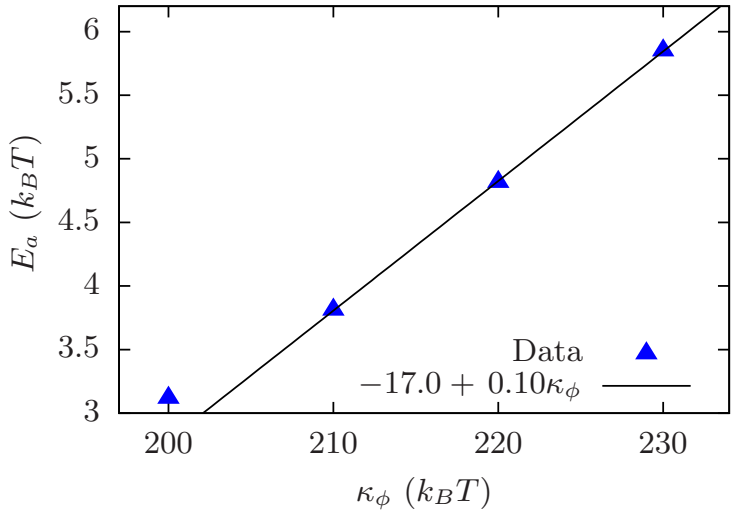

FIGURE 4.18: Activation Energy computed from Arrhenius plots against κ_ϕ for $N =$ 70 bps. The black line is the fitted straight line for $\beta\kappa_\phi \geq 210$.

contribution from the bending energy of two single strands is on the order of a few $k_{\mathrm{B}}T_0$ only. Below the specific value of $\beta\kappa_\phi = 200$, the activation energy, E_{a} saturates as the ADL mechanism, where the rotational diffusion of the arms plays a role, is independent of κ_ϕ and starts competing with the TA mechanism.

We also performed simulations of clamped DNA for $\beta\kappa_\phi = 300$. We used DNA molecules of length 50 bps. Out of 20 realizations, 12 of them did not close up to 100 μs and 8 of them closed in 54, 44, 72, 29, 15, 91, 33 and 82 μs respectively. Thus, we found very long closure times for clamped DNAs for $\beta\kappa_\phi = 300$. The bubble evolution map for a typical realization is shown in Fig. 4.19. For the realizations where the bubble closes, the bubble diffuses back and forth between the clamped arms and closes possibly at some time. *From these simulations, it is clear that DNAs that are clamped at the ends, thus mimicking heteropolymer or long DNAs, take a long time to close on the order of* 10 *to* 100 μs *in agreement with experiments [13].*

4.2.4 Classification of the three regimes

All the simulation data are compiled in Table 4.1. For given κ_ϕ and N, the percentage of realizations belonging to BDL, ADL and TA is given, computed for 200 realizations. Typical snapshots of DNA going through all three types of closure are shown in Fig. 4.20. The ADL dwell time is the time taken by the DNA to go from the hairpin state to the almost aligned state. The BDL dwell time or TA dwell time are the times taken by the DNA for the complete closure of the bubble either by opening at one end of the

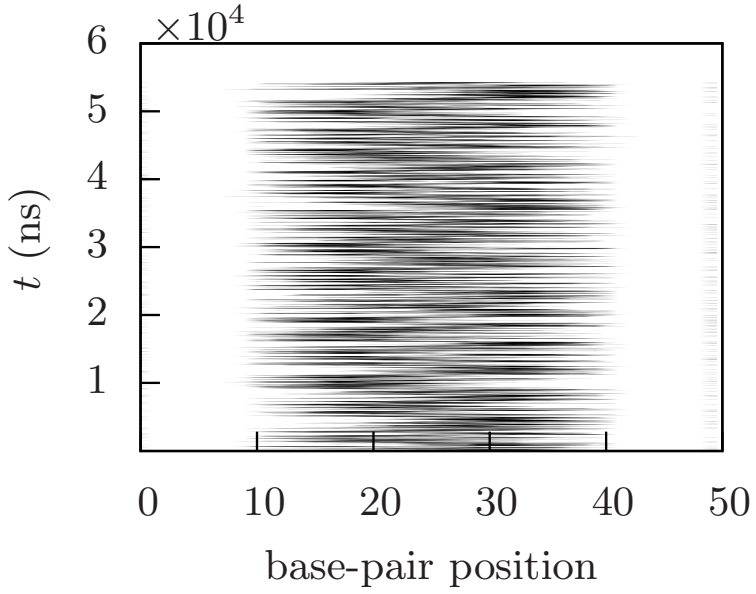

FIGURE 4.19: Typical evolution of the bubble for $N = 50$ bps and $\beta\kappa_\phi = 300$ for clamped DNA. The opened base-pairs are in black and the closed ones are in white.

N	60			70			80			90			100		
κ_ϕ	BDL	TA	ADL	BDL	TA	ADL	BDL	TA	ADL	BDL	TA	ADL	BDL	TA	ADL
200	35.4	49.4	15.1	28.4	40.0	31.5	20.2	49.2	30.5	12.5	33.9	50.5	9.1	31.7	59.1
210	56.4	35.9	7.7	44.4	35.2	20.4	34.7	43.5	21.7	28.3	44.0	27.7	20.2	45.0	34.8
220	70.0	22.7	7.2	56.8	32.9	10.1	54.3	37.5	8.1	42.6	43.6	13.7	34.5	38.0	27.3
240	89.6	9.2	1.0	89.6	9.8	0.5	81.2	16.2	2.5	74.8	20.7	4.3	71.1	23.9	4.9
250	95.2	4.7	0.0	90.5	7.8	1.5	90.6	6.7	2.6	90.1	6.5	3.2	87.8	9.4	2.7
300	100	0	0	100	0	0	100	0	0	100	0	0	100	0	0

TABLE 4.1: Percentage of the bubble trajectories (DNA with free ends) following the three closure mechanisms, ADL, BDL or TA, where κ_ϕ is given in the units of $k_\mathrm{B}T_0$ and N is given in bps.

DNA or crossing the activation energy barrier respectively. The comparison of all the three regimes of closure is based on mean values of time distributions. The distinction between BDL, ADL and TA is done as follows: Comparing BDL to TA or ADL can be easily made because in the BDL mechanism one end of the DNA opens, whereas the distinction between ADL and TA is more subtle. For a given realization, we computed the ADL dwell time, $\tau_{\mathrm{met}}^{\mathrm{ADL}}$, and the TA dwell time, $\tau_{\mathrm{met}}^{\mathrm{TA}}$. If $\tau_{\mathrm{met}}^{\mathrm{ADL}} < \tau_{\mathrm{met}}^{\mathrm{TA}}$, then the realization is considered to be ADL else it is TA. In Fig. 4.21, the ADL and TA dwell time distributions are shown for $\beta\kappa_\phi = 200$ and $N = 100$ bps. The mean values are shown by vertical lines. Even though the mean ADL dwell time is larger than the mean TA one, the time distributions overlap for a large part since the TA starts to become significant for $\beta\kappa_\phi = 200$. Similarly, the TA dwell time distribution (for $\beta\kappa_\phi = 200$) is

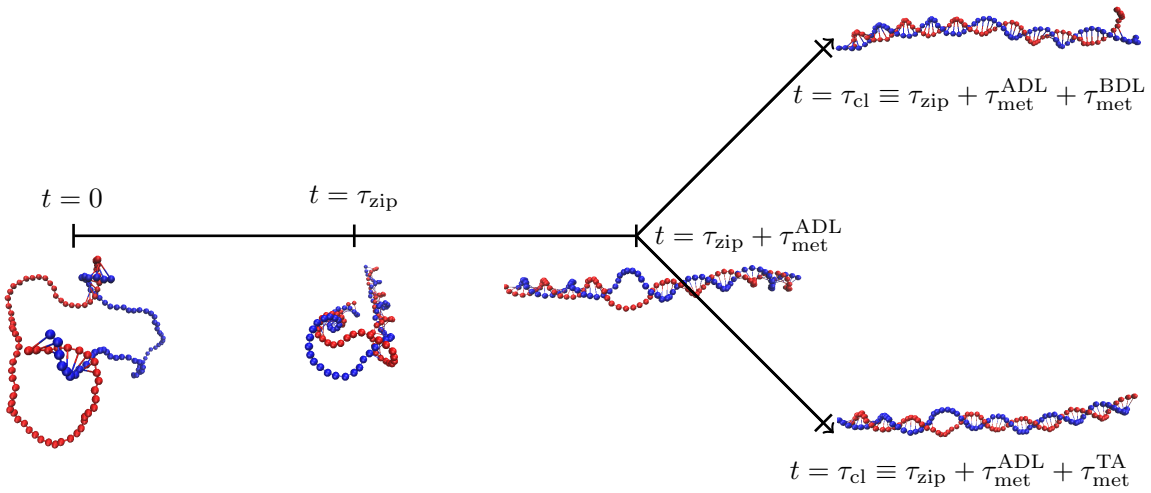

FIGURE 4.20: Sketch of the three dwell times, $\tau_{\rm met}^{\rm ADL}$, $\tau_{\rm met}^{\rm BDL}$, and $\tau_{\rm met}^{\rm TA}$ along with corresponding snapshots for typical realizations. The zipping time, $\tau_{\rm zip}$ and the closure time, $\tau_{\rm cl}$ are also indicated.

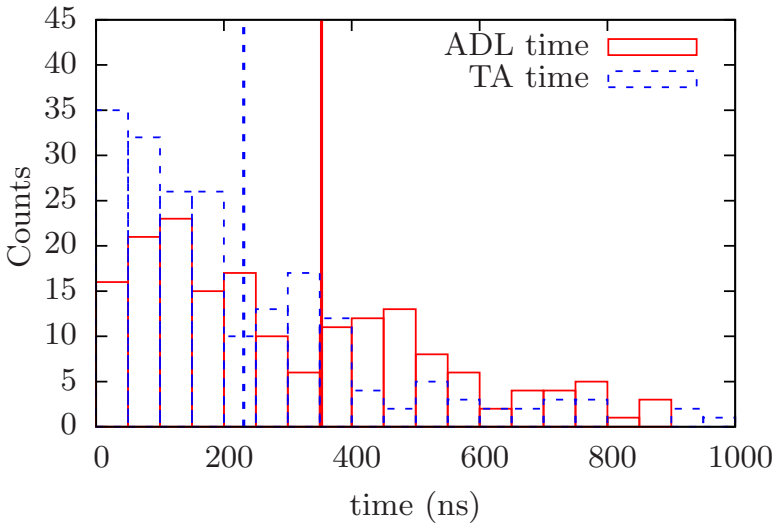

FIGURE 4.21: ADL and TA metastable dwell time distributions (free ends, $\beta\kappa_\phi = 200$, $N = 100$). Mean values are represented by vertical lines.

plotted along with the BDL one (for $\beta\kappa_\phi = 300$ and for three N values) in Fig. 4.22. As N increases, BDL dwell time distributions become wider and the overlap with the TA dwell time distribution becomes smaller. This implies that the number of BDL cases for $\beta\kappa_\phi = 200$ decreases with N, which is also clear from the Table 4.1. Note that we used BDL dwell time distributions of $\beta\kappa_\phi = 300$ to compare with TA dwell time distribution of $\beta\kappa_\phi = 200$. As already mentioned earlier, mean BDL dwell time slightly decreases with decreasing κ_ϕ.

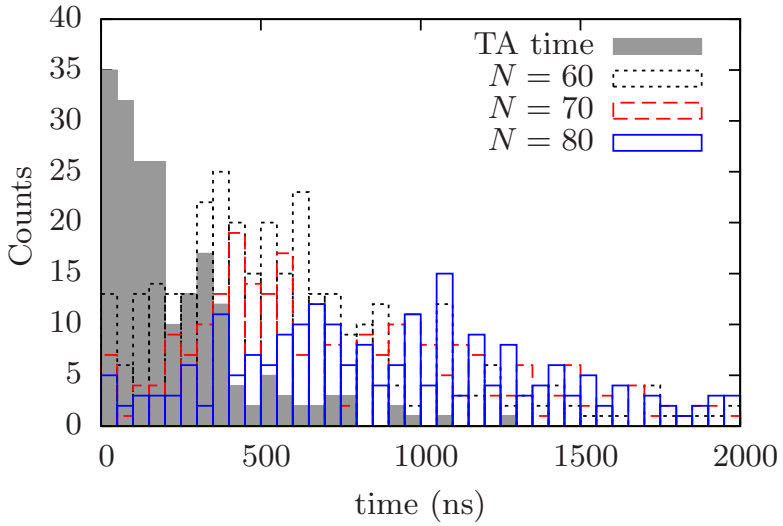

FIGURE 4.22: BDL dwell time $\tau_{\text{met}}^{\text{BDL}}$ distributions for various N and $\beta\kappa_\phi = 300$, together with the TA time $\tau_{\text{met}}^{\text{TA}}$ distribution for free ends and $\beta\kappa_\phi = 200$.

4.2.5 "Phase Diagram"

We construct the generic "phase diagram" in the plane κ_ϕ and N representing the three different regimes of closure. A few assumptions are made in constructing the diagram:

1. The Arms Diffusion Limited time does not depend significantly on the value of κ_ϕ.
2. The Bubble Diffusion Limited time does not depend on the value of κ_ϕ. However, note that there is a slight dependence of BDL dwell time κ_ϕ as shown in Fig. 4.16.
3. The TA dwell time does not depend on N since it is a local mechanism that only depends on the local parameters of DNA, κ_ϕ and A.

To construct the phase diagram, one needs to know the Temperature Activated (TA) times for $\beta\kappa_\phi = 200$ and $\beta\kappa_\phi = 210$ as those values correspond to both ADL times and TA times (as observed in simulations). We find the TA times for both cases by extrapolating the curve, fitted for 220, 230 and 240, to lower $\beta\kappa$ as shown in Fig. 4.23.

We then plot all together the three different dwell times, $\tau_{\text{met}}^{\text{ADL}}$, $\tau_{\text{met}}^{\text{BDL}}$ and $\tau_{\text{met}}^{\text{TA}}$ against N in Fig. 4.24. TA dwell times are shown as horizontal lines since they are assumed not to depend on N. Comparison of these three times allow us to construct the phase diagram. For a given N and κ_ϕ, the least time is expected to be observed. From the Fig. 4.24, we note down the cross-sectional points between ADL, TA and BDL with N and κ_ϕ. These

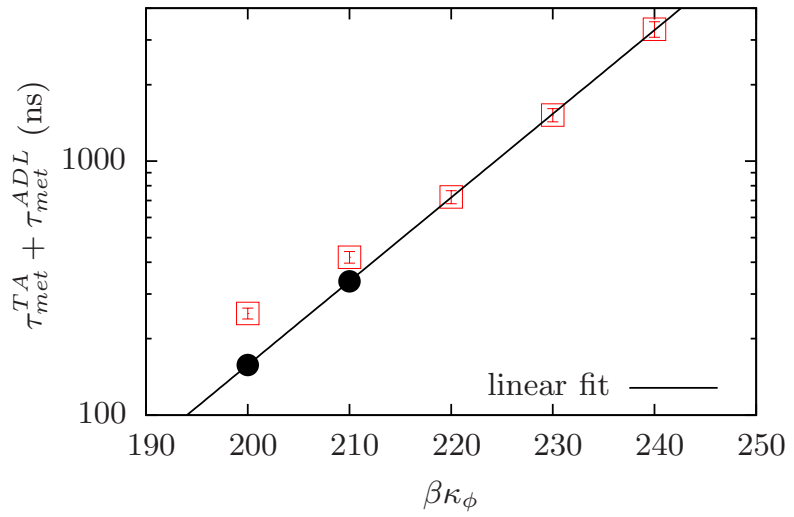

FIGURE 4.23: Linear-log plot of evolution of the TA dwell time with κ_ϕ for clamped ends. The values for $\beta\kappa_\phi = 200$ and 210 (•) are extrapolated (see text).

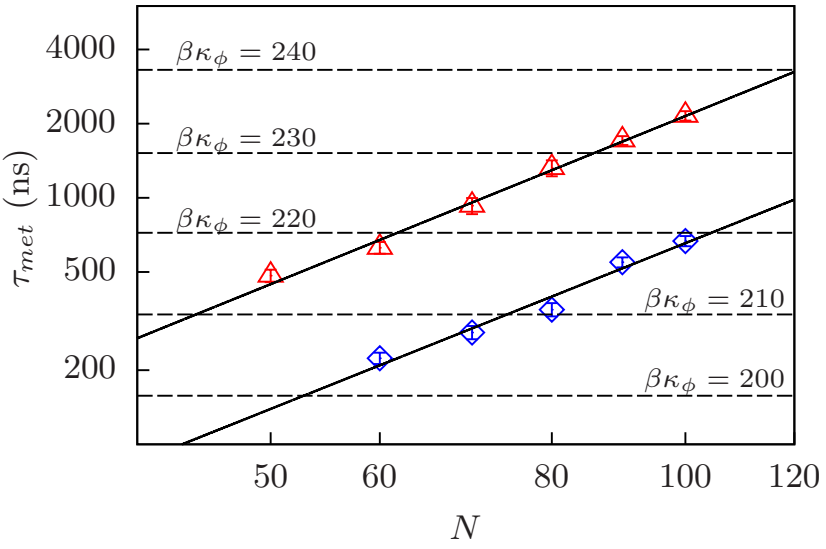

FIGURE 4.24: Log-log plot of dwell times, $\tau_{\text{met}}^{\text{ADL}}$, $\tau_{\text{met}}^{\text{BDL}}$ and $\tau_{\text{met}}^{\text{TA}}$ vs DNA length, N

cross-sectional points will be at the frontiers between all the three regimes. For example, in Fig. 4.24, the intersection between a TA horizontal line for $\beta\kappa_\phi = 200$ and the ADL time gives $N = 52$ bps. It means that above $N = 52$ bps, the metastable bubble closes mainly through ADL mechanism which is clear from the Table 4.1. Similarly, the intersection between a TA horizontal line for $\beta\kappa_\phi = 240$ and the BDL time gives $N = 120$ bps, which states that below $N = 120$ bps the bubble closes mainly by BDL mechanism and above which it closes by TA mechanism. As there is no intersection between ADL and BDL, there will be no frontier between ADL and BDL. The data points extracted in the plane

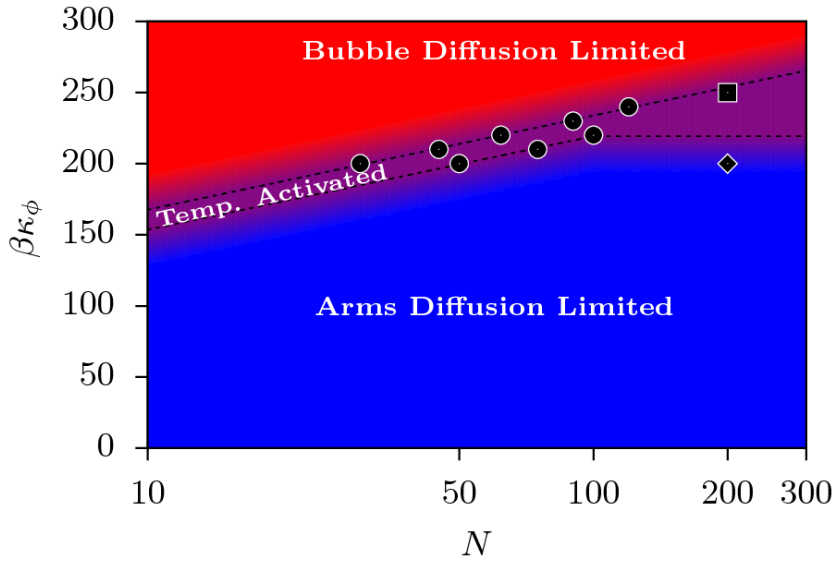

FIGURE 4.25: "Phase Diagram" with three mechanism of closure for DNA with free ends. The extracted data points are shown by black circles ●. In addition, ♦ and ■ are tested points (see Text).

of N and κ_ϕ from Fig. 4.24 are reported in the phase diagram shown in Fig. 4.25.

Since TA dwell times are given by $\tau_{\rm met}^{\rm TA} = \tau_0 \exp\left[\beta E_a(\kappa_\phi)\right]$ and BDL or ADL times are given by $\tau_{\rm met} = \tau_0' N^\alpha$, and equating both laws yields the expected equation of the frontier line between TA and BDL or ADL, $\beta\kappa_\phi = u + v \ln N$. The fitted frontier line for BDL and TA (5 data points) yields, $\beta\kappa_\phi = 88 + 29 \ln N$, and for ADL and TA (3 data points), $\beta\kappa_\phi = 101 + 29 \ln N$.

We have also done a few additional simulations for larger N:

• $\beta\kappa_\phi = 200$ & $N = 200$ (♦ in phase diagram):

Out of 56 realizations, the number of realizations belonging to ADL are 37, to TA are 18 and to BDL are 1. The mean closure time is found to be, $\tau_{\rm closure} = 1.9\ (\pm 0.1)\ \mu s$. This point is thus slightly below the ADL–TA frontier line in the phase diagram as expected, which is shown Fig. 4.25.

• $\beta\kappa_\phi = 250$ & $N = 200$ (■ in phase diagram):

Out of 56 realizations, the number of realizations belonging to TA are 32, to BDL are 23 and to ADL are 1. This point is almost at the BDL-TA frontier in the phase diagram. The mean closure time is found to be, $\tau_{\rm closure} = 4.7\ (\pm 0.4)\ \mu s$.

The phase diagram presented in Fig. 4.25 describes the possible closure mechanism for the denaturation bubble, depending on the range of parameters κ_ϕ and N. The ADL–TA

and BDL–TA frontiers are fuzzy by definition because they are computed by comparing the mean values of wide time distributions. The ADL–TA frontier is made independent of N after $N = 100$ (somewhat arbitrary value) as we already pointed out that the rotational diffusion of the arms starts saturating as the size of the arms reaches the order of persistence length of the DNA, $\ell_{\rm ds}$. Even though persistence length for $\beta\kappa_\phi = 200$ is $\ell_{\rm ds} \simeq 100$ bps, saturation starts well before M reaches $\ell_{\rm ds}$ (refer to Fig. 3.12 in Chapter 3). From the phase diagram, for shorter DNAs, all the three mechanisms exist if one varies κ_ϕ. For larger DNAs ($N > 300$ bps in the phase diagram) only the TA mechanism exist which points out that in longer DNAs the closure of denaturation bubble is only local and does not involve the whole DNA chain.

So far, we discussed the phase diagram of bubble closure mechanisms of DNA with free ends, which mimics homopolymer DNAs. The phase diagram for a clamped DNA or a long one is similar to that with free ends except BDL regime is replaced by TA one, since, in this limit, the diffusion time to reach one DNA end is always larger than the TA one, as shown in Fig. 4.27. For example in the case of $\beta\kappa_\phi = 300$, after $N > 780$ bps, $\tau_{\rm met}^{BDL} > \tau_{\rm met}^{\rm TA}$ which implies that only TA mechanism of closure exists. It means that for a clamped DNA, there are only two regimes of bubble closure, ADL and TA.

4.3 Discussion

In this chapter, we studied the closure dynamics of pre-equilibrated denaturation bubbles using a DNA helical model. The closure dynamics consists of a fast zipping regime followed by a metastable regime. The closure of metastable bubble occurs *via* three possible mechanisms: Arms Diffusion Limited, or Bubble Diffusion Limited or Temperature Activated closure depending upon the parameters and boundary conditions. We have also constructed the "phase diagram" in the plane of torsional constant, κ_ϕ and length of the DNA, N. Total closure times up to 100 μs were found.
As explained in the Introduction 1, Altan-Bonnet et *al* [13] measured in florescence correlation spectroscopic experiments, bubble closure times of 20 to 100 μs for dentauration bubbles of 18 bps. An Arrhenius law of activation energy, $E_{\rm a}^{\rm exp} = 7$ kcal/mol $\approx 12\ k_{\rm B}T_0$, is also found in the experiments. In Fig. 4.18, we have shown the activation energy barrier, $E_{\rm a}$ against $\beta\kappa_\phi$. By extrapolating the linear law that we found to the experimentally measured activation energy, $E_{\rm a}^{\rm exp}$, we found the corresponding torsion constant,

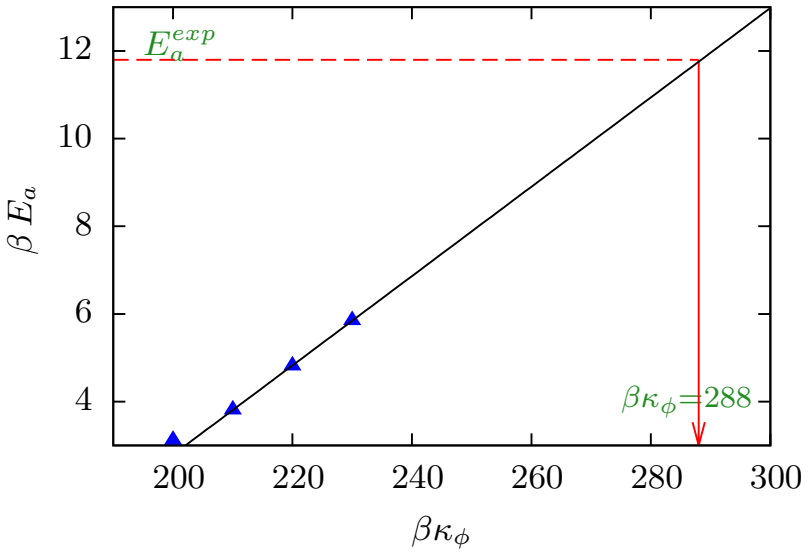

FIGURE 4.26: Activation Energy is plotted against κ_ϕ for $N = 70$ bps. The black line is a linear fitted law and it is extrapolated to experimental activation energy, $E_{\rm a}^{\rm exp} \approx 12\ k_{\rm B}T_0$.

$\beta\kappa_\phi = 288$, as shown in Fig. 4.26. The estimated torsion constant, $\beta\kappa_\phi = 288$ is consistent with the observations [29, 83] and also with single molecule experiments on DNA in which the torsion constant, C is estimated to be 440 ± 40 pN nm^2 ($330 \pm 30\ k_{\rm B}T_0$). In Fig. 4.27, we have indicated the estimated torsion constant in the phase diagram of the two mechanisms of closure. Since the experiments are done in the clamped DNA case, the BDL is replaced by TA in phase diagram. The estimated torsion constant falls in the TA mechanism of closure. It implies that, in clamped or sufficiently long DNAs, the closure of denaturation bubbles does not depend on the DNA chain diffusion.

There exists other DNA models taking torsional degrees of freedom into account. The inclusion of twisting degrees of freedom to the original Peyrard-Bishop (PB) model was done by Barbi et *al.* [40, 41]. Apart from the radial coordinate, as in the PB model, an angular coordinate is also included in this model. The bubbles or breathers in this model appear as localized excitations coming from a wave type equation [35, 44]. Unlike in the PB model, the breathers solutions are associated with the degree of untwisting in angular direction. However, this model lacks the bending fluctuations of dsDNA, as the helical axis of model DNA is assumed straight, and also the bending of single stranded DNA. Here we have shown that chain bending plays a role in both $\Delta\phi$ and $E_{\rm a}$, which are at the origin of the long timescales that we observed.

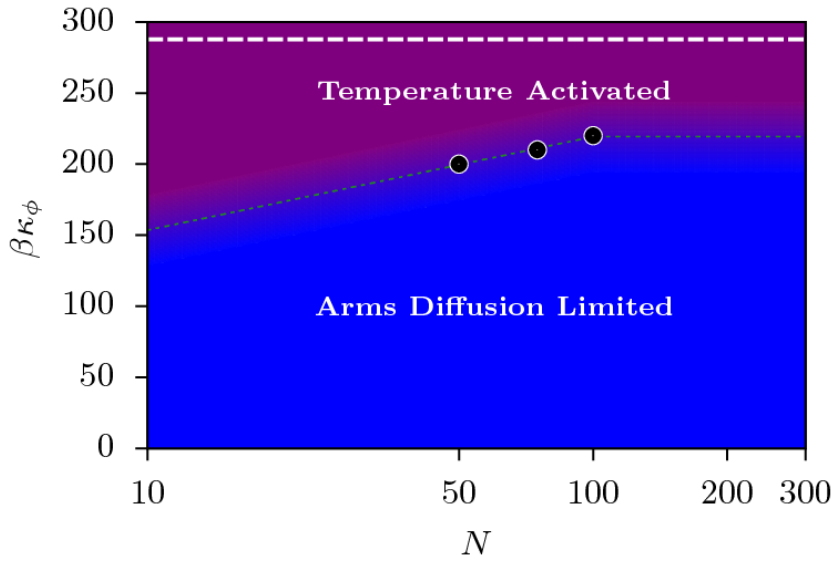

FIGURE 4.27: "Phase diagram" of the closure mechanisms in the case of clamped ends. The white line corresponds to the estimated torsion constant, $E_a^{exp} \approx 12\ k_B T_0$.

The present model did not consider the effects of sequence specificity. The sequence specificity in the DNA mainly influences two interactions, hydrogen bonding interactions between complementary bases and stacking interactions between base-pairs [105]. The sequence effects for hydrogen bonding could be easily considered in our model. Inclusion of sequence effects could modify the BDL mechanism of bubble closure as the bubble diffusion depends on the winding and unwinding rates of base-pairs, whereas ADL and TA might not be affected. Finally, effects of hydrodynamic interactions on the bubble closure dynamics are currently under investigation.

Chapter 5

Conclusions and Perspectives

5.1 Conclusions

In this thesis, we studied the dynamics of the denaturation bubbles in dsDNA. One of the main goals was to understand the physical mechanism behind the large timescales, typically $20 - 100\ \mu s$, found in fluorescence correlation spectroscopy (FCS) measurements by Altan-Bonnet et *al.* [13] on DNA constructs of length $N = 30$ bps. To tackle this problem, we used simple coarse-grained models of dsDNA developed by us.
First, we tackled the problem using the ladder model which is explained in Chapter 3. We showed that the closure of pre-equilibrated bubbles in model DNA occurs via two steps: (1) a fast zipping stage and (2) a metastable regime when the bubble size remains almost constant, $\bar{L} \simeq 10$ bps. The zipping rate is found to be $\simeq 10^7$ bp/s which is comparable with the experimental zipping rate, $\simeq 10^8$ bp/s found in Imino proton exchange experiments [101], eventhough we find a linear zipping only at short times. The closure of the metastable bubble is shown to be limited by the alignment of the dsDNA arms through rotational diffusion as shown in Fig. 5.1.

The origin of the metastable bubble comes from the bending energy of the two ssDNAs inside the bubble. Other one-dimensional models which have been used in studying bubble dynamics or breathing dynamics of dsDNA missed its elastic properties, such as bending of the chain. One such model is the Poland-Scheraga model [17, 18] which is an Ising-like model with entropic penalty for closing the bubble. The breathing dynamics in dsDNA was studied using the Poland-Scheraga model free energy, with stochastic dynamics [25] and with Fokker-Plank equation approach [24, 48, 102]. But, the model is too simple,

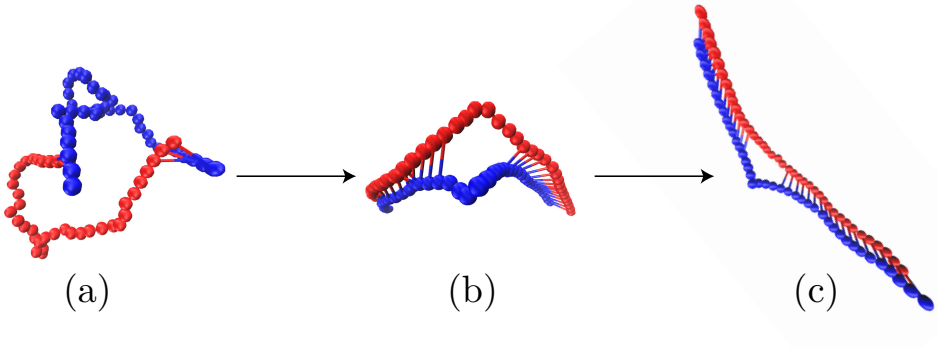

FIGURE 5.1: A typical snapshot showing the two regimes of bubble closure in the ladder model, zipping and metastable state.

and it is based on the equilibrium weights like loop entropy (see Eq. (1.2) in Chapter 1) which is derived from the equilibrium probability distribution of a polymer forming a loop of length, 2ℓ, where ℓ is the bubble size. Another extensively used model is the Peyrard-Bishop model [31] which is also a one-dimensional continuous model as explained in Chapter 1. The bubbles are formed in this model because of energy localization due to non-linear terms and the lifetimes are found to be on the order of pico-seconds [103]. The model assumes only one degree of freedom per base-pair, which is the base-pair distance. These models also lack the chain degrees of freedom and the diffusion of the whole chain in space which were shown to play a crucial role in the closure of denaturation bubbles using the ladder model.

Eventhough we recovered large timescales, $\tau_{\text{closure}} \simeq 0.1 - 4\ \mu s$ for the bubble closure, these timescales are smaller than the experimental timescales found in *in vitro* FCS measurements by Altan-Bonnet et *al.* [13].

To understand the role of twist in the bubble closure, we studied closure of bubbles using the helical model that is described in Chapter 4. After a fast stage of zipping, the bubble closure is now shown to be limited by three distinct mechanisms: (1) Arms Diffusion Limited closure (ADL); (2) Bubble Diffusion Limited (BDL) closure; or (3) Temperature Activated (TA) closure. These three mechanisms depend on the torsional constant, κ_ϕ, the DNA length, N, and the boundary conditions. The first mechanism of closure, ADL, is already familiar to us, as we observed the same mechanism in the ladder model. The fact that this mechanism is a limiting mechanism relies on the fact that DNA falls into an hairpin configuration after the stage of zipping. It becomes the dominating mechanism of closure for small $\beta\kappa_\phi$ values ($\beta\kappa_\phi \leq 220$), as the activation energy barrier goes down to few $k_{\text{B}}T_0$. For larger $\beta\kappa_\phi$ values, both BDL and TA are observed, which again depends on diffusion timescale (of the bubble to reach one end of the DNA) and first passage time

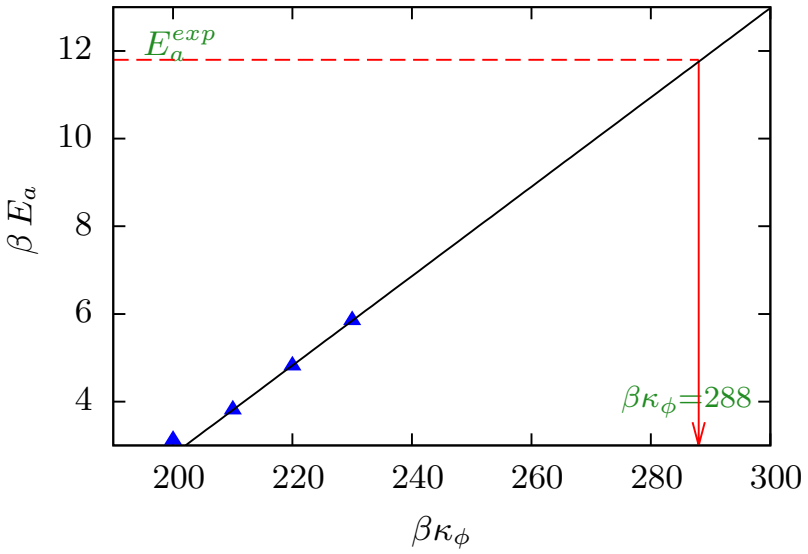

FIGURE 5.2: Activation Energy is plotted against κ_ϕ for $N = 70$ bps. The black line is a linear fitted law and it is extrapolated to experimental activation energy, $E_{\rm a}^{\rm exp} \approx 12\ k_{\rm B}T_0$.

to escape through the energy barrier (Kramer's problem). Overall, total closure times of up to 100 μs are found, which are comparable with the experimental timescales. In FCS experiments, an activation energy of $E_{\rm a}^{\rm exp} = 7$ kcal/mol $\approx 12\ k_{\rm B}T_0$ is also found. From the activation energy barrier *vs* $\beta\kappa_\phi$ plot (see Fig. 4.18 in Chapter 4), we estimated the torsion constant corresponding to the experimental activation energy, $\beta\, E_{\rm a}^{\rm exp} \simeq 12$, which is again shown in Fig. 5.2. The estimated torsion constant, $\beta\kappa_\phi^{\rm exp} = 288$, is consistent with the observations [83]. The estimated torsion constant value falls in the TA regime in the "phase diagram", for which the numerical mean closure time is $\overline{\tau_{\rm met}^{\rm TA}} \simeq 70\ \mu s$, also in agreement with experiments [13]. Note that BDL regime is replaced by TA regime for clamped DNAs that are used in FCS experiments. It also implies that, for clamped DNAs or long DNAs, the metastable bubble closes *via* Temperature Activated regime. As the TA mechanism is local, the final closure of metastable bubble does not depend on the whole chain dynamics.

There are other models having torsional degrees of freedom in the literature. One such model was developed by Barbi et *al.* [40, 41]. It is an extension of the Peyrard-Bishop model [31], by adding torsional degrees of freedom to the PB model. Even though, geometrically this model is richer than the original PB model, the main conclusions are similar. The denaturation bubbles appear as localized excitations coming from nonlinear Schrödinger-type equation [35, 44]. However, as already discussed earlier, this model misses the flexibility of the whole DNA and single stranded DNA, which we have shown to play an important role in the activation energy barrier. Finally, we have shown

that chain bending and twisting, which are elastic properties, are at the origin of long timescales measured in FCS [13] or NMR experiments [54].

To sum up, we claim that DNA denaturation bubbles of size larger than 10 bps behave differently from "DNA breathing", which is widely explored in the literature. DNA breathing typically comes from the non-linearity of hydrogen bonding interaction potential such as the Morse potential. But, the dynamics of closure of denaturation bubbles of size larger than 10 bps couple with the elastic properties such as bending and twisting of both ssDNA and dsDNA, which justifies the current Thesis title.

To this respect, equilibrating the chain *after* opening the bubble and *before* reclosing it is essential. Indeed, just after opening, the ssDNAs inside the bubble keep their straight configuration from their former dsDNA form. If the just opened bubble is not equilibrated, the bubble can close immediately because the remaining dsDNA arms are already aligned. This gives rise to apparently very short bubble lifetimes, in the nano-second range, as in [45]. This remark stresses the important difference that we make between "transient bubbles" (opening rapidly followed by renaturation) and "equilibrated bubbles". Interestingly, also using imino proton exchange NMR experiments, the authors in [54] found two distinct timescales associated with bubble lifetimes. The first one dwells in the micro-second range, as in [101], but the experiments indicate the existence of shorter open states, with nano-second lifetimes. We postulate that these shorter lifetimes correspond to transient bubbles, whereas the longer ones are associated with equilibrated bubbles as studied here.

Up to now, most of the models only considered the base-pairing degrees of freedom [17, 31] to explain the dynamics of DNA denaturation bubbles, ignoring the chain degrees of freedom. It also has been shown by Palmeri et *al.* [27, 28, 29] that the elasticity of DNA plays an important role in the DNA denaturation at an equilibrium level. These different remarks suggest that both elastic degrees of freedom and base-pair degrees of freedom are coupled, and cannot be tackled independently.

5.2 Perspectives

In this thesis, we understood the physical mechanism behind the long timescales involved in the closure of denaturation bubbles of dsDNA measured in experiments [13, 54] using simple coarse-grained models, explained in Chapter 3 and Chapter 4. As the coarse-grained models are simulated using Brownian dynamics in the free draining limit, one

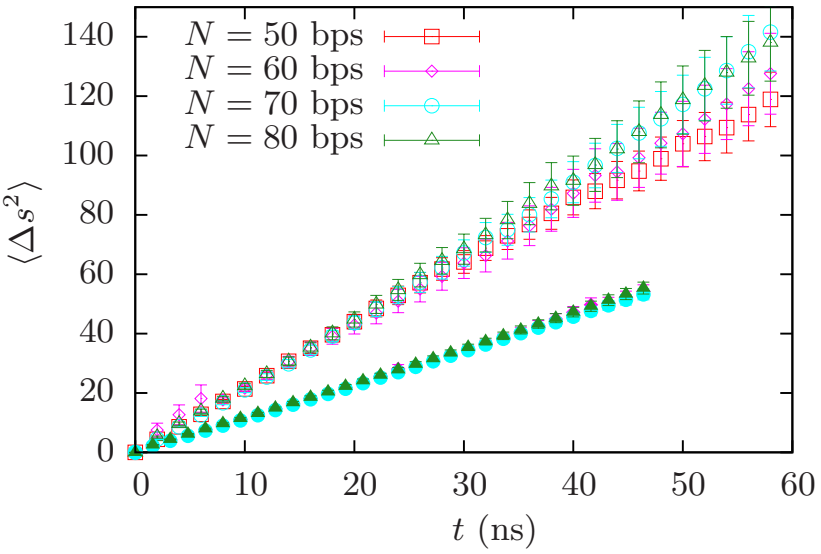

FIGURE 5.3: The mean square deviation of the bubble along DNA, shown in both **presence of Hydrodynamics** (open symbols) and **absence of Hydrodynamics** (closed symbols).

can extend them and study the effects of hydrodynamics interactions which are long-range interactions. The hydrodynamics interactions can be included in Brownian dynamics using the Ermak-McCammon algorithm [106]. The diffusion tensor, $\mathbf{D}$, becomes non-diagonal unlike in the free draining limit. The diffusion tensor is approximated by the Rotne-Prager-Yamakawa tensor [107, 108]. This work is currently under investigation. As the inclusion of hydrodynamics interactions induces long-range effects in the system, the dynamics of bubble closure might be accelerated. For example, inclusion of hydrodynamics interactions to the helical model seems to reduce drastically the number of ADL closures, even though these preliminary results have to be confirmed. As already explained in the Chapter 4, the dsDNA falls into an hairpin configuration just after the zipping stage, which further leads to ADL mechanism for low $\beta\kappa_\phi$ value. By contrast, when hydrodynamics interactions are switched on, the dsDNA is almost in the aligned state after the zipping stage. Another example illustrating the role of hydrodynamic drag forces in bubble dynamics concerns the BDL mechanism of closure, where the bubble diffuses to one end of the DNA and eventually closes. After the inclusion of hydrodynamics interactions, the diffusion of the bubble is increased by almost a factor of 2, as shown in Fig. 5.3, which essentially decreases the BDL dwell time. These preliminary results should affect the "phase diagram".

Another study could be the inclusion of sequence specificity which is ignored in the present model. Right now, we are only able to study homopolymer DNAs. The effect of sequence modifies the main contributions of stability of dsDNA, which are hydrogen

bonding and stacking interactions [105, 109, 110]. The sequence specificity could be incorporated in hydrogen bonding potential by changing the width and depth of the potential. The stacking interactions are included by changing the torsion constant, $\kappa_{\phi_i}(\rho_i, \rho_{i-1})$, depending on the intra base-pair distances. The sequence effects in stacking interactions can not be easily included, as one has to change the twist profile according to the sequence. The sequence also modifies slightly the dsDNA bending modulus [111]. The effect of sequence could modify the BDL mechanism of closure as it depends on the rates of winding and unwinding of base-pairs. The ADL mechanism of closure should not be significantly affected as it is limited by rotational diffusion of the arms which depends only on the bending constant of the bubble. Taking the sequence effects into account could modify the activation energy barrier, as the metastable bubble is formed by both single stranded elasticity and the stacking of base-pairs, which might further affects the TA dwell times.

As a future study, one might explore further the mapping between the zipping in bubble closure (of ladder model) and polymer translocation or adsorption, which is not clearly done in the current thesis. The exact mapping will lead to a better understanding of the zipping dynamics of the bubble closure in the ladder model. In the same ladder model, the metastable residence time is understood *via* mean first passage time approach (MFPT), where the full coupled problem (coupling of the angle between the arms, $\theta(t)$, and the end-to-end distance of the bubble, $R(t)$) is not solved. A similar study can also be extended to the helical model, for example mean first passage time in coordinates, end-to-end distance of the bubble and the mean twist angle in the bubble, $\Delta\phi(t)$. One could continue to study the full coupled problem analytically. In the ladder model, even though there is not explicit twisting potential, the non-zero twist with zero mean, arises in the model because of the two semi-flexible chains [112]. The effect of this non-zero twist can be studied in bubble closure dynamics.

The helical model is not limited to the exploration of the bubble closure dynamics, but can be used to study, for example mechanical denaturation of DNA. Denaturation is not only achieved *via* increasing the temperature (thermal denaturation), but also by mechanical force through single molecule manipulation experiments. Thereby denaturation of DNA can be achieved well below the melting temperature. The dynamics of mechanical denaturation of DNA [69, 113, 114, 115] or overstretching transition of DNA [116, 117] could be studied using our helical model. It would be interesting to determine whether such a simple mesoscopic model can be applied to both thermal and mechanical denaturation.

Our results ought to be tested at the experimental level. For example, Libchaber's experiments [13] are done for single size of DNA, $N = 30$ bps. These experiments can be

extended for variable sizes of DNA constructs. The activation energies can be extracted from the Arrhenius plot. As we have shown that the activation energy does not depend on the length of DNA, this point could be clarified by these experiments. The same experiments could also be carried out for homopolymer DNAs with tagged fluorescent particle at the ends of DNA constructs, to support the BDL mechanism of closure.

Appendix A

DNA Hybridization Dynamics

In order to compare it to the zipping dynamics in bubble closure, we also studied the hybridization (or renaturation) dynamics of the helical DNA. The initial configuration is achieved by switching off the Morse potential to an equilibrated DNA, leaving 10 bps at only one end of the DNA. At $t = 0$, we start applying the Morse potential. In Fig. A.1, snapshots of a trajectory are shown.

The evolution of the bubble size, normalized with initial bubble, $L(t)/L(0)$ for both

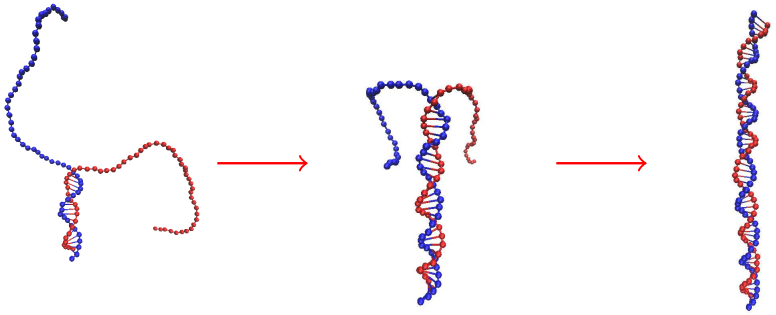

FIGURE A.1: Snapshots of a typical trajectory are shown for $N = 60$ bps and $\beta\kappa_\phi = 300$.

bubble closure and hybridization against time are shown in Fig. A.2. The bubble size in bubble closure case starts saturating after some time which denotes the metastable regime. We have chosen different lengths of DNA for comparison to make the initial bubble sizes equal, $L(0) = 50$ bps. The zipping dynamics in hybridization is different and slower than the zipping dynamics in bubble closure. In the case of the ladder model, described in Chapter 3, both zipping dynamics in hybridization and bubble dynamics were comparable (see Fig. 3.7), which is not the case here.

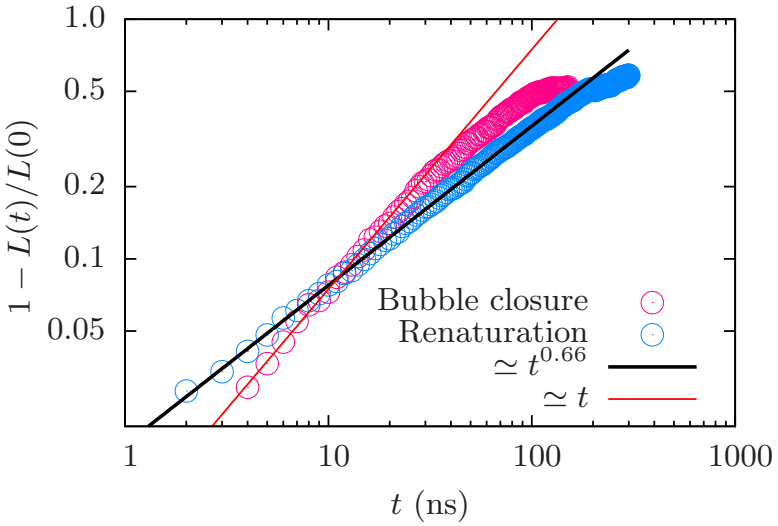

FIGURE A.2: The bubble size in DNA hybridization ($N = 60$ bps) and in bubble closure ($N = 70$ bps) are plotted against time for $\beta\kappa_\phi = 300$. The black line is power law fit to renaturation data with the scaling exponent of 0.66. The red line is fitted linear law at short times for the bubble dynamics. This figure should be compared with Fig. 3.7.

The bubble dynamics in renaturation follows a power law, $1 - L(t)/L(0) \simeq t^{0.66} \simeq t^{1/1.5}$. This exponent is comparable with the exponent that we found in ladder model. By contrast, the bubble dynamics in bubble closure does not follow a power law, whereas it is described by Eq. (4.9) in Chapter 4. The bubble dynamics follows a linear law at short times (given by red line in Fig. A.2) which is expected from Eq. (4.9) in Chapter 4. So the bubble dynamics in bubble closure and renaturation does not follow the same laws, which is not the case in ladder model. One reason could be that the elastic restoring torque, $\kappa_\phi \phi_{\rm el}$, and torque coming from the bending energy of two strands, $\mathcal{T}_{\rm bend}(t)$, in Eq. (4.3) does not enter in renaturation dynamics. Another reason could be that the separation between the single strands in renaturation is larger and the non-equilibrium response to stretching of strands when the ds part of DNA rotates, should also be considered in the dynamics.

Bibliography

[1] W. C. K. Poon and D. Andelman. *Soft Condensed Matter Physics in Molecular and Cell Biology (Scottish Graduate Series)*. Taylor & Francis, January 2006.

[2] J. D. Watson and F. H. C. Crick. *Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid*. Nature, **171**(1953) 737.

[3] F. H. C. Crick and J. D. Watson. *The complementary structure of deoxyribonucleic acid*. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, **223**(1954) 80.

[4] R. E. Franklin and R. G. Gosling. *Molecular configuration in sodium thymonucleate*. Nature, **171**(1953) 740.

[5] R. Wing, et al. *Crystal structure analysis of a complete turn of b-dna*. Nature, **287**(1980) 755.

[6] C. O. Pabo and R. T. Sauer. *Protein-dna recognition*. Annual Review of Biochemistry, **53**(1984) 293.

[7] Z. Kudritskaya and V. Danilov. *Quantum mechanical study of bases interactions in various associates in atomic dipole approximation*. Journal of Theoretical Biology, **59**(1976) 303.

[8] V. R. Cooper, et al. *Stacking interactions and the twist of dna*. Journal of the American Chemical Society, **130**(2008) 1304.

[9] C. Bouchiat, et al. *Estimating the persistence length of a worm-like chain molecule from force-extension measurements*. Biophysical Journal, **76**(1999) 409.

[10] C. Bustamante, et al. *Entropic elasticity of lambda-phage DNA*. Science, **265**(1994) 1599.

[11] B. Alberts, et al. *Molecular biology of the cell*. Garland, 4$^{\text{th}}$ edition, 2002.

[12] R. M. Wartell and A. S. Benight. *Thermal denaturation of DNA molecules: A comparison of theory with experiment.* Physics Reports, **126**(1985) 67.

[13] G. Altan-Bonnet, et al. *Bubble dynamics in Double-Stranded DNA.* Physical Review Letters, **90**(2003) 138101.

[14] S. Tyagi and F. R. Kramer. *Molecular Beacons: Probes that Fluoresce upon Hybridization.* Nature Biotechnology, **14**(1996) 303.

[15] O. Gotoh. *Prediction of melting profiles and local helix stability for sequenced dna.* Advances in Biophysics, **16**(1983) iii .

[16] B. S. Alexandrov, et al. *A nonlinear dynamic model of dna with a sequence-dependent stacking term.* Nucleic Acids Research, **37**(2009) 2405.

[17] D. Poland and H. A. Scheraga. *Occurrence of a phase transition in nucleic acid models.* The Journal of Chemical Physics, **45**(1966) 1464.

[18] D. Poland and H. A. Scheraga. *Phase transitions in one dimension and the HelixCoil transition in polyamino acids.* The Journal of Chemical Physics, **45**(1966) 1456.

[19] M. E. Fisher. *Effect of excluded volume on phase transitions in biopolymers.* The Journal of Chemical Physics, **45**(1966) 1469.

[20] M. E. Fisher. *Walks, walls, wetting, and melting.* Journal of Statistical Physics, **34**(1984) 667.

[21] Y. Kafri, et al. *Why is the DNA denaturation transition first order?* Physical Review Letters, **85**(2000) 4988.

[22] Y. Kafri, et al. *Melting and unzipping of DNA.* The European Physical Journal B - Condensed Matter, **27**(2002) 135.

[23] H. C. Fogedby and R. Metzler. *DNA bubble dynamics as a quantum coulomb problem.* Physical Review Letters, **98**(2007) 070601.

[24] T. Ambjrnsson, et al. *Sequence sensitivity of breathing dynamics in heteropolymer dna.* Physical Review Letters, **97**(2006) 128105.

[25] K. P. N. Murthy and G. M. Schtz. *Mean bubble formation time in DNA denaturation.* EPL (Europhysics Letters), **96**(2011) 68003.

[26] A. Bar, et al. *Loop dynamics in DNA denaturation.* Physical Review Letters, **98**(2007) 038103.

[27] J. Palmeri, et al. *Thermal denaturation of fluctuating DNA driven by bending entropy.* Physical Review Letters, **99**(2007) 088103.

[28] J. Palmeri, et al. *Thermal denaturation of fluctuating finite dna chains: The role of bending rigidity in bubble nucleation.* Physical Review E, **77**(2008) 011913.

[29] M. Manghi, et al. *Coupling between denaturation and chain conformations in DNA: stretching, bending, torsion and finite size effects.* Journal of Physics: Condensed Matter, **21**(2009) 034104.

[30] A. K. Dasanna, et al. *Strand diffusion-limited closure of denaturation bubbles in dna.* EPL (Europhysics Letters), **98**(2012) 38002.

[31] M. Peyrard and A. R. Bishop. *Statistical mechanics of a nonlinear model for DNA denaturation.* Physical Review Letters, **62**(1989) 2755.

[32] T. Dauxois, et al. *Entropy-driven dna denaturation.* Physical Review E, **47**(1993) R44.

[33] T. Dauxois, et al. *Dynamics and thermodynamics of a nonlinear model for dna denaturation.* Physical Review E, **47**(1993) 684.

[34] M. Peyrard. *Nonlinear dynamics and statistical physics of dna.* Nonlinearity, **17**(2004) R1.

[35] M. Peyrard, et al. *Nonlinear analysis of the dynamics of DNA breathing.* Journal of Biological Physics, **35**(2009) 73.

[36] T. Dauxois, et al. *Thermodynamic instabilities in one dimension: Correlations, scaling and solitons.* Journal of Statistical Physics, **107**(2002) 869.

[37] T. Dauxois, et al. *Localized breather-like solution in a discrete klein-gordon model and application to dna.* Physica D: Nonlinear Phenomena, **57**(1992) 267 .

[38] G. J. Martyna, et al. *Nos[e-acute]–hoover chains: The canonical ensemble via continuous dynamics.* The Journal of Chemical Physics, **97**(1992) 2635.

[39] A. Campa and A. Giansanti. *Experimental tests of the peyrard-bishop model applied to the melting of very short dna chains.* Physical Review E, **58**(1998) 3585.

[40] M. Barbi, et al. *Helicoidal model for DNA opening.* Physics Letters A, **253**(1999) 358.

[41] S. Cocco and R. Monasson. *Statistical mechanics of torque induced denaturation of DNA.* Physical Review Letters, **83**(1999) 5178.

[42] M. Barbi, et al. *A twist opening model for dna.* Journal of Biological Physics, **24**(1999) 97.

[43] S. Zdravković. *Helicoidal peyrard-bishop model of dna dynamics.* Journal of Nonlinear Mathematical Physics, **18**(2011) 463.

[44] A. Campa. *Bubble propagation in a helicoidal molecular chain.* Physical Review E, **63**(2001) 021901.

[45] J.-H. Jeon, et al. *A semiflexible chain model of local denaturation in double-stranded dna.* The Journal of Chemical Physics, **124**(2006) 164905.

[46] J. Kim, et al. *A breathing wormlike chain model on DNA denaturation and bubble: Effects of stacking interactions.* The Journal of Chemical Physics, **128**(2008) 055101.

[47] O.-c. Lee, et al. *How double-stranded dna breathing enhances its flexibility and instability on short length scales.* Physical Review E, **81**(2010) 021906.

[48] A. Hanke and R. Metzler. *Bubble dynamics in DNA.* Journal of Physics A: Mathematical and General, **36**(2003) L473.

[49] D. T. Gillespie. *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions.* Journal of Computational Physics, **22**(1976) 403 .

[50] D. T. Gillespie. *Exact stochastic simulation of coupled chemical reactions.* The Journal of Physical Chemistry, **81**(1977) 2340.

[51] S. K. Banik, et al. *Stochastic approach to dna breathing dynamics.* EPL (Europhysics Letters), **71**(2005) 852.

[52] T. Ambjörnsson, et al. *Breathing dynamics in heteropolymer dna.* Biophysical Journal, **92**(2007) 2674 .

[53] T. Ambjörnsson, et al. *Master equation approach to dna breathing in heteropolymer dna.* Physical Review E, **75**(2007) 021908.

[54] S. Wärmländer, et al. *Imino proton exchange in DNA catalyzed by ammonia and trimethylamine: evidence for a secondary Long-Lived open state of the base pair.* Biochemistry, **39**(1999) 607.

[55] M. Bishop, et al. *Molecular dynamics of polymeric systems.* The Journal of Chemical Physics, **70**(1979) 1299.

[56] J. Baschnagel, et al. *Monte carlo simulation of polymers: Coarsegrained models, in computational soft matter: From synthetic polymers to proteins edited by n. attig et al.* NIC Series, (2004) 83.

[57] P. M. Morse. *Diatomic molecules according to the wave mechanics. II. vibrational levels.* Physical Review, **34**(1929) 57.

[58] T. A. Knotts, et al. *A coarse grain model for DNA.* The Journal of Chemical Physics, **126**(2007) 084901.

[59] M. Manghi, et al. *Probing dna conformational changes with high temporal resolution by tethered particle motion.* Physical Biology, **7**(2010) 046003.

[60] F. Zhang and M. A. Collins. *Model simulations of DNA dynamics.* Physical Review E, **52**(1995) 4217.

[61] S. Nos. *A unified formulation of the constant temperature molecular dynamics methods.* The Journal of Chemical Physics, **81**(1984) 511.

[62] W. G. Hoover. *Canonical dynamics: Equilibrium phase-space distributions.* Physical Review A, **31**(1985) 1695.

[63] D. J. Evans and B. L. Holian. *The NoseHoover thermostat.* The Journal of Chemical Physics, **83**(1985) 4069.

[64] K. Drukker and G. C. Schatz. *A model for simulating dynamics of dna denaturation.* The Journal of Physical Chemistry B, **104**(2000) 6108.

[65] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids.* Oxford science publications. Oxford University Press, USA, June 1989.

[66] M. Sayar, et al. *Twist-writhe partitioning in a coarse-grained DNA minicircle model.* Physical Review E, **81**(2010) 041916.

[67] H. Limbach, et al. *Espressoan extensible simulation package for research on soft matter systems.* Computer Physics Communications, **174**(2006) 704 .

[68] E. Sambriski, et al. *A mesoscale model of DNA and its renaturation.* Biophysical Journal, **96**(2009) 1675.

[69] A. Florescu and M. Joyeux. *Thermal and mechanical denaturation properties of a DNA model with three sites per nucleotide.* The Journal of Chemical Physics, **135**(2011) 085105.

[70] T. E. Ouldridge, et al. *Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model.* The Journal of Chemical Physics, **134**(2011) 085101.

[71] T. E. Ouldridge, et al. *Dna nanotweezers studied with a coarse-grained model of dna.* Physical Review Letters, **104**(2010) 178101.

[72] H. R. Warner. *Kinetic theory and rheology of dilute suspensions of finitely extendible dumbbells.* Industrial & Engineering Chemistry Fundamentals, **11**(1972) 379.

[73] K. Kremer and G. S. Grest. *Dynamics of entangled linear polymer melts: a molecular-dynamics simulation.* The Journal of Chemical Physics, **92**(1990) 5057.

[74] S. Whitelam, et al. *The role of collective motion in examples of coarsening and self-assembly.* Soft Matter, **5**(2009) 1251.

[75] S. P. Mielke, et al. *Brownian dynamics simulations of sequence-dependent duplex denaturation in dynamically superhelical DNA.* The Journal of Chemical Physics, **123**(2005) 124911.

[76] S. P. Mielke, et al. *Brownian dynamics of double-stranded DNA in periodic systems with discrete salt.* Physical Review E, **77**(2008) 031924.

[77] M. C. Linak, et al. *Moving beyond WatsonCrick models of coarse grained DNA dynamics.* The Journal of Chemical Physics, **135**(2011) 205102.

[78] S. Niewieczerza and M. Cieplak. *Stretching and twisting of the DNA duplexes in coarse-grained dynamical models.* Journal of Physics: Condensed matter, **21**(2009) 474221.

[79] A. Morriss-Andrews, et al. *A systematically coarse-grained model for DNA and its predictions for persistence length, stacking, twist, and chirality.* The Journal of Chemical Physics, **132**(2010) 035105.

[80] S. Buyukdagli, et al. *Towards more realistic dynamical models for DNA secondary structure.* Chemical Physics Letters, **419**(2006) 434.

[81] A. K. Dasanna, et al. *Slow closure of denaturation bubbles in DNA: twist matters.* Physical Review E, **87**(2013) 052703.

[82] M. Doi and S. F. Edwards. *The Theory of Polymer Dynamics.* Oxford University Press, USA, November 1986.

[83] C. Bouchiat and M. Mézard. *Elasticity model of a supercoiled dna molecule.* Physical Review Letters, **80**(1998) 1556.

[84] N. L. Goddard, et al. *Sequence dependent rigidity of single stranded dna.* Physical Review Letters, **85**(2000) 2400.

[85] C. Ke, et al. *Direct measurements of base stacking interactions in dna by single-molecule atomic-force spectroscopy.* Physical Review Letters, **99**(2007) 018302.

[86] S. B. Smith, et al. *Overstretching b-dna: The elastic response of individual double-stranded and single-stranded dna molecules.* Science, **271**(1996) 795.

[87] B. Tinland, et al. *Persistence length of single-stranded DNA.* Macromolecules, **30**(1997) 5763.

[88] M. Manghi and R. R. Netz. *Variational theory for a single polyelectrolyte chain revisited.* The European Physical Journal E, **14**(2004) 67.

[89] S. Chandrasekhar. *Stochastic problems in physics and astronomy.* Reviews of Modern Physics, **15**(1943) 1.

[90] C. W. Gardiner. *Handbook of stochastic methods : for physics, chemistry and the natural sciences.* Springer series in synergetics, 13. Springer, November 2002.

[91] R. Kubo, et al. *Statistical Physics II: Nonequilibrium Statistical Mechanics (Springer Series in Solid-State Sciences, 31).* Springer, November 1995.

[92] D. Poland and H. A. Scheraga. *Theory of helix-coil transitions in biopolymers: statistical mechanical theory of order-disorder transitions in biological macromolecules.* Academic Press, 1970.

[93] D. Panja, et al. *Non-equilibrium dynamics of single polymer adsorption to solid surfaces.* Journal of Physics: Condensed Matter, **21**(2009) 242101.

[94] A. Ferrantini and E. Carlon. *Anomalous zipping dynamics and forced polymer translocation.* Journal of Statistical Mechanics: Theory and Experiment, **2011**(2011) P02020.

[95] K. Luo, et al. *Driven polymer translocation through nanopores: Slow-vs.-fast dynamics.* EPL (Europhysics Letters), **88**(2009) 68006.

[96] Y. Yin and X. S. Zhao. *Kinetics and dynamics of DNA hybridization.* Accounts of Chemical Research, **44**(2011) 1172.

[97] T. Sakaue. *Sucking genes into pores: Insight into driven translocation.* Physical Review E, **81**(2010) 041808.

[98] A. Y. Grosberg, et al. *How long does it take to pull an ideal polymer into a small hole?* Physical Review Letters, **96**(2006) 228105.

[99] Y. Kantor and M. Kardar. *Anomalous dynamics of forced translocation.* Physical Review E, **69**(2004) 021806.

[100] N. Destainville, et al. *Microscopic mechanism for experimentally observed anomalous elasticity of DNA in two dimensions.* Biophysical Journal, **96**(2009) 4464.

[101] M. Guéron, et al. *A single mode of DNA base-pair opening drives imino proton exchange.* Nature, **328**(1987) 89.

[102] M. Bandyopadhyay, et al. *Dna breathing dynamics: Analytic results for distribution functions of relevant brownian functionals.* Physical Review E, **83**(2011) 031905.

[103] B. S. Alexandrov, et al. *Toward a detailed description of the thermally induced dynamics of the core promoter.* PLoS Computational Biology, **5**(2009) e1000313.

[104] J. F. Marko. *Stretching must twist DNA.* Europhysics Letters (EPL), **38**(1997) 183.

[105] J. M. Huguet, et al. *Single-molecule derivation of salt dependent base-pair free energies in dna.* Proceedings of the National Academy of Sciences, **107**(2010) 15431.

[106] D. L. Ermak and J. A. McCammon. *Brownian dynamics with hydrodynamic interactions.* The Journal of Chemical Physics, **69**(1978) 1352.

[107] J. Rotne and S. Prager. *Variational treatment of hydrodynamic interaction in polymers.* The Journal of Chemical Physics, **50**(1969) 4831.

[108] H. Yamakawa. *Transport properties of polymer chains in dilute solution: Hydrodynamic interaction.* The Journal of Chemical Physics, **53**(1970) 436.

[109] M. Peyrard, et al. *Experimental and theoretical studies of sequence effects on the fluctuation and melting of short dna molecules.* Journal of Physics: Condensed Matter, **21**(2009) 034103.

[110] J. SantaLucia. *A unified view of polymer, dumbbell, and oligonucleotide dna nearest-neighbor thermodynamics.* Proceedings of the National Academy of Sciences, **95**(1998) 1460.

[111] S. Geggier and A. Vologodskii. *Sequence dependence of dna bending rigidity.* Proceedings of the National Academy of Sciences, **107**(2010) 15421.

[112] H. Zhou, et al. *Elastic property of single double-stranded dna molecules: Theoretical study and comparison with experiments.* Physical Review E, **62**(2000) 1045.

[113] S. Cocco, et al. *Unzipping dynamics of long dnas.* Physical Review E, **66**(2002) 051914.

[114] S. Cocco, et al. *Force and kinetic barriers to initiation of dna unzipping.* Physical Review E, **65**(2002) 041907.

[115] E. Orlandini, et al. *Mechanical denaturation of dna: existence of a low-temperature denaturation.* Journal of Physics A: Mathematical and General, **34**(2001) L751.

[116] F. Romano, et al. *Coarse-grained simulations of dna overstretching.* The Journal of Chemical Physics, **138**(2013) 085101.

[117] M. Manghi, et al. *Mesoscopic models for dna stretching under force: New results and comparison with experiments.* The European Physical Journal E, **35**(2012) 1.