Chapter 1

Introduction

Deoxyribonucleic acid (DNA) takes a determining task for the entire activity of any cell in living organisms as it is the key molecule for the realization of genetic information, storage, and duplication. Parts of DNA are called genes, which are the basic physical and functional units of heredity. Gene that holds an exclusive sequence of DNA nucleotides passes the information that is needed to build the functional molecules called Proteins. On the other hand, genes also take the responsibility to produce that type of molecules which can help the cell to assemble those proteins. The journey from DNA to protein is probably the most sophisticated process in the world. This process has two major steps, transcription and replication. Both these processes are involved in making copies of the DNA in a cell. Executing these two processes, the DNA has to unzip first, and the unzipping of duplex DNA is known as denaturation of the DNA. In vitro, the denaturation process has been studied experimentally. Theoretical studies have also been going side by side. In this thesis, we investigate the various aspects of the DNA denaturation through statistical mechanics and thermodynamic properties. Additionally, we study the dynamics of DNA through the molecular dynamics simulation method.

The structural behavior of DNA and the dynamics of denaturation of DNA are discussed in Appendix A (5.6). The DNA research is very wide in range from medical field to physics field and it is also one of the fastest growing topics. Therefore, it is difficult to touch upon all upgoing research which are realted to it in the literature review. So we would like to focus only on the issues which are related to this thesis work. In this chapter, we present a brief literature survey of the theoretical as well as of the experimental studies on the denaturation of DNA.

1.1 DNA *in-vitro*

DNA synthesis is an artificial creation of DNA opening. In nature, DNA opening is a process that is followed by DNA replication in all living cells. *In-vitro*, DNA opening is also studied. The two strands of DNA can be separated by increasing the temperature of the DNA-solution. DNA denaturation is also designated as the "DNA melting at thermal ensemble". Heat energy breaks the hydrogen bonds between the bases of DNA, and it uncoils the duplex-DNA. DNA denaturation can be studied in-vitro in many different ways also. The same study can be done by changing the pH of the solution as well as by pulling either of the strands by any mechanical device. The DNA denaturation in force ensemble is known as the "force-induced DNA melting". These are the way to understand the mechanism of replication and transcription. In recent years, the study of DNA is not limited to only physiologically relevant issues but also manipulating it for technology purposes [3–5] as DNA holds many exotic properties like storing molecular memory [6, 7], elasticity, and many more.

1.1.1 DNA Melting in Thermal Ensemble

Double strands DNA helix is dissociated to two single strands through the process of heating and this is called DNA melting or DNA denaturation. The heat energy weaken the hydrogen bonds that confine the two strands of DNA and after a certain temperature double starnds start to break. The temperature, at which half of the base pairs in the total chain are denaturated, is referred to as the melting temperature (T_m) [8–10]. The bond strength of G-C pairs and A-T pairs are in an approximate ratio of 1.25-1.5 as the GC pairs have three while AT pairs have two hydrogen bonds [11–21]. So it signifies that the amount of heat energy requires for denaturing a DNA chain depends on its nucleotide composition. When the temperature is increasing, after a certain value of temperature bubbles are created in the sequence and it is the portion of DNA from where the denaturation of the chain starts. The spot of the first bubble forming depends on the strength of the bit section of a DNA heterogeneous chain. After that bubbles get bigger, and therfore the base pairs are broken. The UV (ultraviolet) absorption spectroscopy is the most common method for studying DNA denaturation experimentally [8, 9, 22]. In UV spectroscopy technique, first the DNA is dissolved in an aqueous buffer and then it is heated. In the process of double strand breaking we get a significant increase in the UV absorbance. A schematic graph of the UV spectroscopy is

shown below (see Fig 1.1). The melting temperature(T_m) of a DNA chain depends

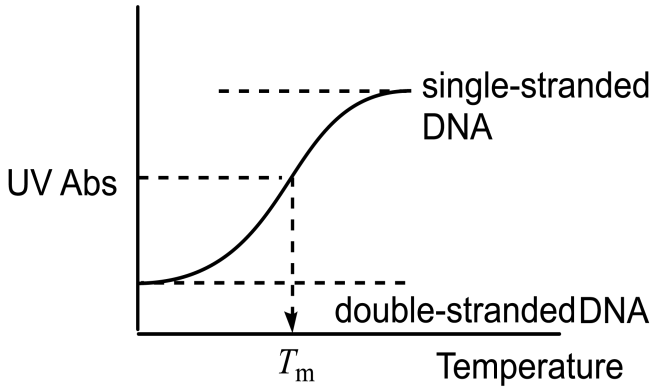


Figure 1.1: Melting Curve: Increase in fraction of open base pairs with temperature.

on various detail of the DNA molcule and the surroundings where it belongs. Some of the major factors are: DNA chain length, nucleotide sequence manner (A,T,G,C combination), the soluation properties (pH, salt concentration, specific cations and anions properties, buffers, hydrophobicity etc.), excluded volume effect[1]. The free space avilable for opening of DNA molecules is also a major parameter in DNA melting. In this section, we discuss the thermal melting of DNA in the presence of salt (low and high salt concentration) and different geometrical confinement effect on DNA.

1.1.2 DNA Melting in Force Ensemble

In living organisms, the process of DNA unzipping is not driven by thermal energy; rather, it is driven by the mechanical force. Enzymes and some specific proteins apply this force to separate the dsDNA into two ssDNA [8, 23]. Force-induced DNA unzipping is more real compare to the thermal-induced melting. So in force-induced study, dsDNA is thermally stable in the absence of any mechanical force. The force-induced unzipping of DNA also shows a continuous phase transition like the thermal ensemble. In the force-induced DNA melting, there are two ways in which force can be applied depending on the interest of the study. If DNA unzipping is the study of interest, then force should be applied in the perpendicular direction, and when the force is applied in the parallel direction, it is called rupture. Please see Fig.1.2. The advancement of different experiment

[1]excluded volume refers to the area in a solvent that is occupied by another molecule.

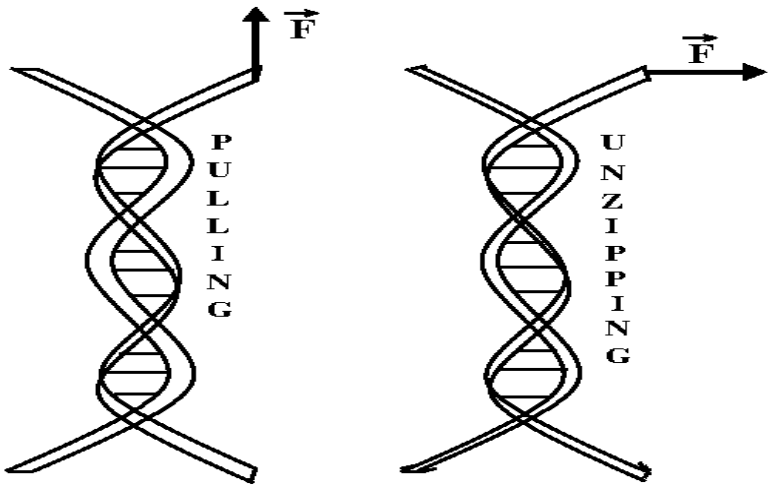


Figure 1.2: Force induced DNA melting: along the DNA axis & perpendicular to the DNA axis.

techniques like atomic force microscopy (AFM) [24] (range of forces between 10 and 10^4 pN), optical tweezers (OT) [25] (force range of 10^{-1} to 100 pN), magnetic tweezers (MT) [26] (force range between 10^{-2} to 100 pN), bio-membrane force-probe [27], and many others [23, 28] help to understand the functions of DNA. The SMFS (Single-Molecule Force Spectroscopy) experiments also play an important role in understanding not only DNA but other biomolecules also [23, 29]. Through Fig.1.3, the experiment set-ups are shown.

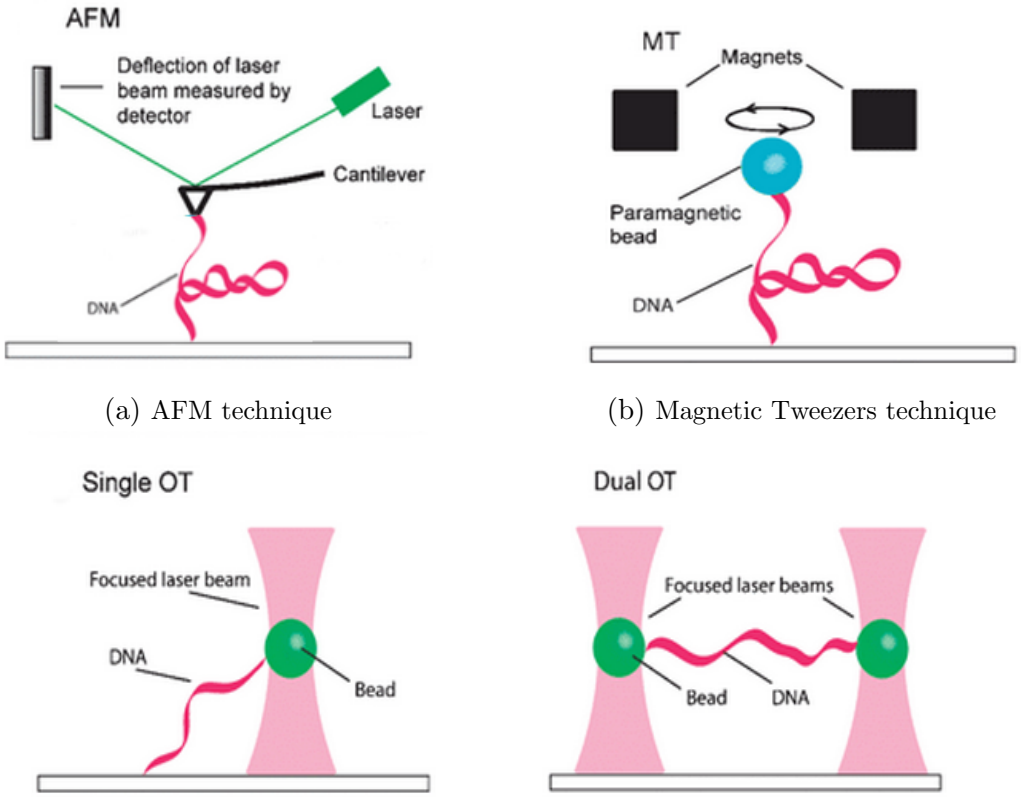
Literature review

The genetic information of life is interpreted through biology, chemistry and also physics as biology cannot breakout from the fundamental laws of physics and chemistry. Systematically mathematics and statistics manifest these behavior in equation format according to theory. In the next two sections we try to discuss these studies of DNA unzipping in thermal ensemble and in force ensemble.

1.2 Role of salt concentration

Experimental and Theoretical studies:

Nucleotides are charged polyanion. The DNA molecule has a negative charge backbone of phosphates. To reduce the coulombic repulsion between the two negative strands, the cations like sodium, potassium, and magnesium ions are essential. The properties of these ions, such as charge, size, concentration, etc. determine



(a) AFM technique

(b) Magnetic Tweezers technique

(c) Optical Tweezers (OT) technique

Figure 1.3: Mechanical stretching of DNA molecule through different experimental techniques. Figures are taken from [30]

the stability and the dynamics of DNA molecules. The experiment with DNA solution has been studied in the mid-1950s [31, 32], and the hypothesis of DNA melting has been used to study DNA stability. In 1983 polymerase chain reaction (PCR) was developed, and through this technique, one can make many copies of a specific DNA segment [33]. It has also been found that ions highly influence the melting temperature of DNA [34–36]. Since salt like NaCl, MgCl or KCl dissolves in the DNA solution and free its cations such as Na^+, Mg^{2+} or K^+ respectively hence polyelectrolytes DNA molecules is influenced by these cations and in the field of DNA research it creates an appealing area to study the DNA melting in the presence of salts. There are many studies that show that the cations screen the negative charge of DNA strands, and therefore, DNA gets more stability [34, 37, 38]. Owen *et al.* [39] studied how the melting temperature depends on the concentration of Na^+ ions and as well as the heterogeneity of the chain in 1969 experimentally. Their study also supports the screening phenomenon of negative

charge DNA strands. R Owczarzy *et al.* [40] in 2004 did a substantial study with 92 different DNA sequences with respect to a wide range of salt concentration and find a correlation between the scaling of the T_m of DNA duplex oligomers and ion concentrations. In another study, P. Yakovchuk *et al.* [41] showed that how the salt present in the solution affects the stacking interactions in DNA and hence to the overall stability of the molecule. Vuletić *et al.* [42] studied the changing DNA conformation in the presence of salt also. In most of these experiments, the stability of DNA has been studied under low or moderate (0.1 - 1.0 M) concentrations of salt. There are some experiments [2, 43] where the studies were executed at relatively high salt concentrations, and some strikingly different behavior in the DNA molecule was observed. These experiments found that in this range of salt concentration, the stability of DNA is changed. Interestingly, in a similar set of experiments, several research groups tested the condensation process and melting profile in DNA under different ethanol concentrations [44, 45]. In a recent review, Bose *et al.* [46] showed that abundant cations inside the cell as well as outside the cell play important roles in sustaining cancer cells and at the same time in the decay of immune cells. Cancer cells possess a significant electrical character [47] compared to normal somatic cells. Cancer cells become more electrostatic [48] during oncogenesis.

After the discovery of the DNA structure model, researchers put forward to relate the structure with basic theoretical physics. Statistical models are used to describe the DNA structure and dynamics since 1960s [49–54]. The experiment showed that the melting temperature of DNA immensely depends on the salt concentration, so researchers proposed some empirical relations to anticipate the experimental melting temperature. J. Santalucia Jr., proposed the nearest neighbor (NN) thermodynamics parameters to study the salt effect on the melting transition in DNA [55]. Salt dependent melting temperature data of Owen *et al.* was explained theoretically by Frank-Kamenetskii [35]. The electrostatic free energy calculations is also a tool to understand the DNA denaturation. N. Korolev *et al.* used that technique to describe the salt effect on the stability of the DNA [56]. Free energy difference indicates that the stability of DNA significantly depends on ions properties. Researchers calculated the base pair opening probabilities with the self consistent phonon approximation (SCPA) theory [57, 58] and showed that the base pair opening probabilities also varies with salt concentration [59]. On the other hand, the base pair opening probabilities were also calculated through

partition function implementing the unified nearest neighbor parameters. Krueger *et al.* [60] used this partition function method to calculated the base pair opening probabilities. Over the last few decades, the process of thermal DNA melting has been researching through Poland-Scheraga model (PS model) [51, 61, 62] extensively. Poland-Scheraga model was proposed in 1966. An alternating sequence of bound and denaturated states are taken in this model. Jost and Everaers [63] added a correction in this PS model to introduce the salt effect in this model and the experimental data were also compared with this corrected unified PS model. The melting behavior of DNA was studied by Peyrard-Bishop-Dauxois (PBD) model also [64, 65]. In this model, the hydrogen bonding between the two bases represented by Morse potential and anharmonic potential plays the role of stacking energy. N Theodorakopoulos [12] used this PBD model to study the melting behavior of a long DNA chain. Amar Singh *et al.* [38, 66] used that PBD model and gave a theoretical perspective of DNA melting in the ionic solution for homogeneous as well as heterogeneous DNA chain. Their theoretical results show a good agreement with experiment results. A molecular thermodynamic model for DNA melting in ionic and crowded solutions was developed by Y. Lui [67]. The model gives good predictable results of melting temperature in the ionic solution, and it shows a good match with the simulation and experimental results. It is noticed that the melting temperature (T_m) increases with increasing salt concentration. The rate of increase is sharp at low salt concentration, and its rate of sharpness becomes much less at high salt concentration.

1.3 DNA in confined geometry

Experimental and Theoretical studies:

In vivo, DNA molecule is confined in a limited space such as the cell chamber or a channel and is in highly dense solvent conditions [68–71]. This confinement restricts the conformation and movement of DNA molecules in the cell. The thermodynamic properties of DNA molecules highly depend on the confined space and on the solvent properties [72]. It is known that conformational properties of biopolymers under confinement have relevance in living systems like DNA packing in eukaryotic chromosomes, viral capsids, etc. [73]. In order to have a fair understanding of these properties, we should have a deep understanding of the denaturation and renaturation mechanisms of DNA in the confined state. These kinds of studies enable us to use DNA for effective design and to control the

self-assembly behavior of the molecule for various applications. The experimental findings show that the confinement strongly affects the entropy of DNA molecules [74]. In an interesting work, Derrington *et al* [75] show that how *Mycobacterium smegmatis* porin A (MspA) can be used to identify the sequence of DNA. The MspA is a short and narrow constriction that can be used to distinguish the individual nucleotides of single-stranded DNA. In an important work, Lau *et al.* [76] showed that a small DNA strand initially confined in a 4 nm diameter nanotube exhibit dynamics similar to the unconfined molecule, but the behavior is drastically changed when the diameter is reduced to 3 nm. From the point of view of biological purposes and DNA manipulation, carbon nanotubes are used as templates for DNA encapsulation, intracellular penetration via endocytosis and delivery of biological payloads, etc. [73, 77]. Recently, we have witnessed a few works, theoretical as well as experimental, which demonstrate the local properties of the biopolymers translocation in nanochannels [78–81]. Some experimental studies have been conducted on the translocation of biopolymers through conical geometry [82, 83]. It is a matter of interest that which geometry gives the highest possible resolution for reading information along the DNA contour [80, 81]. Franceschini et al. [84] show that during the time of translocation of DNA, how the interior channel possession charges help for the smooth sliding. Their results indicate that the net internal negative surface charge is important to allow the smooth sliding of the opposing negatively charged DNA when it passes through the connector or β-clamp temperature with varying the angle of the conical geometry using a statistical model. Another aspect of confinement can be noticed in the DNA encapsulation technique. The importance of this molecule lies in its wide applications. In gene therapy technique DNA is protected by a physical barrier. There are many different techniques that have been using to get better results like complexation with polycations [85], charged copolymers of different architecture, cationic lipids or liposomes [86]. As other options, DNA can be confined within gel [87, 88], polymeric nanocapsules (micelles) [89, 90], and microparticles. There are many elegant and versatile approaches for DNA encapsulation [91]. Carbon nanotubes also have been proved itself as the potential candidates for DNA encapsulation [92]. The thermodynamical spontaneity of DNA encapsulation of carbon nanotube under different conditions is still a significant area of research. The threshold diameter of this tube is also a vital issue to investigate since below the threshold encapsulation is inhibited [93]. Many sensitive parameters are involved in DNA encapsulation techniques like the medium and topology of the carrier,

thermodynamic parameters, etc.

1.4 DNA passing through different geometrical pores:

Experimental and Theoretical studies:

In recent years, nanopores are found as a promising tool for single-molecule analysis. Nanopores sensors are used widely to study the biopolymers like proteins and nucleic acids [79]. By applying a voltage across the nanopore, the DNA can be sequentially pass through these nanopores [94–98]. The lipid-embedded α hemolysin channels [99–106] are used as model nanopores to explore the secondary structures of nucleic acids by using the electrical force to unzip duplex regions [107, 108]. The major shortcoming that hinders prospective biotechnological nanopore applications are the fragility of these lipid bilayers and lateral diffusion of these channels in the membrane [109, 110]. One of the recent development in this field is the use of solid-state nanopores [111–115] for sensing purpose as they are better than biological nanopores [116, 117] in terms of high stability, easy controllability for size, adjustable surface properties and the potential for their integration into devices and arrays [118, 119]. These solid-state nanopores of diameters between the ssDNA and dsDNA cross-sections (somewhere between 1.5-2 nm) are used to determine the unzipping kinetics of DNA secondary structure by applying a shear force to induce DNA unzipping [120–122]. However, a direct measurement of single-molecule unzipping kinetics through solid-state pores has yet to be reported. Lots of interesting results have been reported that showed the unzipping of duplex DNA and hairpin DNA using biological pores [120, 123]. Recently, the unzipping kinetics of dsDNA through a sub 2 nm solid-state nanopores has been studied by Yao Lin *et al.* [124]. DNA unzipping through these nanopores has applications in the field of clinical and laboratory use of nanopore sensors, which could be easily realized by integrating the solid-state nanopores into all in one bioanalytical device. However, the current patterns corresponding to the whole unzipping process of DNA duplexes have not yet been reported for solid-state nanopores. On the other hand, the conventional *DNA mapping* scheme, which uses enzymatic based labeling, is quite expensive and exhaustive. A new concept of *denaturation mapping*, in comparison, is simple and affordable [125, 126]. Reisner *et al.* have predicted the DNA sequence through the *optical mapping* of the denaturating DNA that is confined in a rectangular nanochannel [127]. By combining

the experimental method with the computer simulation based on Poland-Scheraga model [128], which evaluates the sequence-dependent melting probabilities, this method offers a new horizon to predict the genome sequence.

1.5 Existing research gap

The easiest way to break the dsDNA structure is to heat the DNA solution. Hence the DNA denaturation is also called DNA melting, and it has been studied throughout a long time since the discovery of DNA structure. Researchers have been scrutinizing this problem through the experiments as well as through the theoretical and simulation methods. However, the complexity of the problem provides always more questions to continue the research. The objective of this thesis is to fill the theoretical research gap to an extent with the existing experimental investigations. Some of the research gaps which are addressed in this thesis are listed here.

- The stability of the DNA molecule and the effects of salt on the $B-A$ transition, and on the condensation of the DNA molecule have been studied in detail by several researchers. There are many experiments, and as well as the theory has been published by various research groups. In all these studies, the stability of DNA has been studied under low or moderate (0.1-1.0 M) concentrations of salt. Several experiments conducted by the groups revealed that short as well long DNA molecules become more stable as the content of cations in the solution increases. The theoretical work following these experiments attempted to explain this. In another set of experiments that were executed at relatively high salt concentration, some strikingly different behavior in the DNA molecule was observed. These experiments found that in this range of salt concentrations, the stability of DNA gets shattered. There are no or very few theoretical descriptions of the stability of DNA molecules at higher salt concentration.

- In most of the previous studies (experimental, theoretical, and numerical simulation), the DNA stability in the confined environment has been taken in research for many different aspects like in vivo or DNA encapsulation technique, etc., but theoretical comparison study of different geometry of the confinement has been ignored. Most of the theoretical studies have been worked out on cylindrical geometry exclusively, but there is very few theoretical results are available by considering DNA in a conical geometry.

- The recent years witnessed significant advances in nanopore technology and DNA sequencing. One of a very efficient method that is developed in recent years is the denaturation mapping of DNA molecule that is trapped inside a cylindrical geometry but when DNA translocates through different geometry how the stability of the chain depends on the parameters of this geometry is not studied so much as it demands specifically in theoretical aspect. There are still open questions, like ; (a) when DNA is passing through a nanopore how the entropy changes ongoingly depending on the chain length that have been passed through this channel, (b) which confined geometry can give more significant result for denaturation mapping of DNA, (c) how the stability differs for different geometry at the time of DNA translocation?

We expect future research on above-discussed issues that will elucidate the response of DNA molecule in the presence of high salt, in being confined by different geometry and translocates through different geometry.

1.6 Objectives of the present work

Objectives of the present work are the following.

1. Thermal and mechanical stability of dsDNA molecule at higher salt concentration.
2. Melting of DNA in different confined geometries.
3. Stability of DNA passing through different geometrical pores.

To achieve the above objectives, we use the Peyrard-Bishop Dauxois model (PBD) and molecular dynamics simulation. The PBD model is a statistical model of DNA, proposed in 1989. All simulations are conducted with the GROMACS package. "GROningen MAchine for Chemical Simulations (GROMACS) is a molecular dynamics package mainly designed for simulations of proteins, lipids, and nucleic acids". In Chapter 2, the PBD model and the MD method are discussed in details.