ABSTRACT

Title of Dissertation:	EFFECT OF MISMATCHED BASE PAIRS ON DNA PLECTONEMES
	Parth Rakesh Desai, Doctor of Philosophy, 2022
Dissertation directed by:	Professor Siddhartha Das, Department of Mechanical Engineering, University of Maryland.

Base pair mismatches in DNA occur during replication and can result in mutations and certain types of cancer. The exact mechanism by which mismatch repair proteins recognize mismatches is still not well understood. Structures of mismatch recognition proteins bound to a mismatch indicate that the process involves introducing a sharp bend in the DNA and flipping out the mismatched base. Under external torsional stress, an elastic rod with a defect would buckle at the defect, provided the defect reduces the local bending stiffness. *In vivo*, if the same energetic scenario prevails, it could localize (or pin) the mismatch at the plectoneme end loop (plectoneme refers to a structure formed by the DNA when it buckles, and its helical axis wraps or writhes around itself in the presence of a critical torsional stress) and make the mismatched base pair more accessible to the mismatch repair protein. In genomic DNA, however, the entropic cost associated with plectoneme localization could make pinning unfavorable. Magnetic-tweezers-based studies of DNA supercoiling, performed at high salt concentrations, have shown that in DNA harboring a single mismatch, the plectoneme will always localize at the mismatch. Theoretical studies

have predicted that under physiological salt concentrations, plectoneme localization becomes probabilistic, i.e., the plectoneme does not always localize at the mismatch. Plectoneme localization under physiological salt conditions is dependent on the number of mismatches and tension applied to the DNA. However, both experimental and theoretical approaches are currently limited to positively supercoiled DNA. In the current dissertation, we aim to study plectoneme localization, in physiologically relevant conditions, using state-of-the-art molecular dynamics (MD) simulations and single molecule magnetics tweezers-based experiments.

In order to simulate plectoneme localization we first develop a framework using the widely available sequence and salt dependent OxDNA2 model. We verify that the OxDNA2 model can quantitively reproduce a reduction in bending rigidity due to the presence of the mismatch(es), similar to all-atom MD simulations. We then verify that the current framework can reproduce the experimentally observed plectoneme pinning (at the location of the mismatches). Next, we simulate plectoneme pinning under physiologically relevant conditions. We find that the plectoneme pinning (at the location of the mismatches) becomes probabilistic and this probability of plectoneme pinning increases with an increase in the number of mismatches. We also simulate a longer 1010 base pair long DNA to study the influence of entropic effects on plectoneme pinning.

Next, we extend the simulation framework to simulate a negatively supercoiled, *i.e.*, under-wound, DNA molecule. *In vivo*, DNA is maintained in a negatively supercoiled state. Negative supercoiling can result in local melting at the mismatched base pairs: this local melting would further reduce the local bending rigidity at the

mismatched base pairs and could enhance plectoneme pinning. We find that negative supercoiling significantly enhances plectoneme pinning in comparison with equivalent levels of positive supercoiling. We also find that the mismatched base pairs are locally melted and the plectoneme end loop is bent significantly more as compared to the positive supercoiling case. Additionally, we simulate the 1010 base pair long DNA under two different negative super-helical densities, i.e., two different degrees of unwinding. We find that the super helical density does not affect the plectoneme pinning probabilities. We also conduct simulations of DNA under different stretching forces (0.3 pN, 0.4 pN and 0.6 pN). Negatively supercoiled DNA under relatively high stretching force (~0.6 pN) absorbs tortional stress by locally melting instead of supercoiling. Simulations of DNA under different forces allow us to study the effect of mismatches on the competition between supercoiling and local melting in a negatively supercoiled DNA. We find that higher stretching forces, up to a maximum set by the onset of melting, increase plectoneme pinning at the location of mismatch.

Finally, we propose and develop a single molecule assay to validate the simulations results presented in the previous chapters. Previous single-molecule magnetic tweezers measurements of mismatch DNA buckling and pinning were limited to the high force (~2 pN) – high salt (>0.5 M NaCl) regime. We propose to overcome this limitation by attaching a small gold nano-bead via a di-thiol group close to the mismatched base pairs, which permits direct observation of transient DNA buckling at the mismatch. We generate a DNA substrate that can be used to directly observe plectoneme pinning at the mismatch. We perform single-molecule magnetic tweezers

measurements to verify that the presence of the di-thiol group does not result in anomalous pinning in an intact DNA molecule.

EFFECT OF MISMATCHED BASE PAIRS ON DNA PLECTONEMES

by

Parth Rakesh Desai

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:
Professor Siddhartha Das, Chair
Dr. Keir C. Neuman
Professor Peter Chung
Professor Paul Paukstelis
Professor Don DeVoe
Professor Jason Kahn, Dean's Representative

© Copyright by
Parth Rakesh Desai
2022

Acknowledgements

Many people have contributed directly and indirectly to the conclusion of this thesis and I would like to extend my sincerest thanks to all of them.

Firstly, I would like to thank my advisors, Dr. Siddhartha Das and Dr. Keir C. Neuman, for their invaluable support and guidance throughout my dissertation research. I would also like to thank the member of my dissertation committee; Dr. Kahn, Dr. Chung, Dr. DeVoe and Dr. Paukstelis; for their suggestions.

I would also like to thank my friends and family for their invaluable help during this journey. I would also like to thank my colleagues from the Das lab and Neuman lab for all their help.

Finally, I would like to thank the people of the Mechanical engineering graduate office, NHLBI office, international student services at NIH and UMD, for their help with all the paperwork.

Table of Contents

List of Tables

Chapter 2

Chapter 3

List of Figures

List of Abbreviations

Deoxyribonucleic acid (DNA)

Ribonucleic acid (RNA)

Adenine (A)

Cytosine (C)

Guanine (G)

Thymine (T)

Linking number (*Lk*)

Twist (*Tw*)

Writhe (*Wr*)

Mismatch recognition proteins (MMR)

Molecular Dynamics (MD)

Magnetic tweezers (MT)

Bovine serum albumin (BSA)

1× Phosphate buffered saline (PBS)

Chapter 1: Background and Motivation

1.1. DNA

Deoxyribonucleic acid (DNA) was first discovered by Friedrich Miescher in 1869[1]. Avery et al.[2] later observed that DNA was responsible for transfer of traits when dead bacteria are mixed with a live population, confirming DNA to be the carrier of genetic information. In 1953, based on the x-ray diffraction data of Rosalind Franklin[3,4], James Watson and Francis Crick proposed that DNA consists of two intertwined polynucleotide chains which form a double helix[5]. The two polynucleotide chains consist of units that are complementary and linked by hydrogen bonding.

The genetic information stored in DNA is usually propagated to ribonucleic acid (RNA), through transcription, and then propagated from RNA to proteins and enzymes. In vivo, DNA is a relatively long and slender molecule. In cells, DNA needs to be stored in a fashion which allows it to be accessible to various proteins. In eukaryotic cells, DNA is first wrapped around histones and then further packaged with proteins into a complex called chromatin. These chromatin complexes are then condensed into chromosomes, which are stored in the cells' nuclei.

Interaction of DNA with proteins affects the structure of DNA, which in turn regulates the protein functions[6]. The effects of protein on DNA structure have received a lot of attention[7,8]; however, how the presence of a mismatch affects the DNA structure

and protein interactions is not well understood. In the current thesis we study how the presence of defects (or mismatches) affects the DNA structure.

1.1.1. DNA structure

DNA consists of two strands, which intertwine to form a right-handed helix. Each strand is made of polymer consisting of repeating units called nucleotides. A nucleotide consists of three groups: a phosphate group, a sugar group, and a base. The phosphate group and sugar group acts as the DNA backbone and connects the nucleotides to each other. The base group is an aromatic group, which consists of one of four possible bases. These four bases are Adenine (A), Cytosine (C), Guanine (G) and Thymine (T). Each base has a specific complementary base (AT and GC).

1.1.2. DNA topology

DNA topology refers to the study of how two strands of DNA are intertwined. If DNA is topologically constrained, *i.e.*, its ends are not allowed to rotate relative to each other, a topological invariant called *linking number* (*Lk*) can be defined[9,10]

$$Lk = Tw + Wr.$$

Here *Lk* refers to the global measure of wrapping of the two DNA strands around each other. This wrapping can be partitioned between twist, *Tw*, the degree of intertwining of the two single strands about a common axis, and writhe, *Wr,* the degree of self-wrapping of the DNA helix axis. Torsional stress applied to a piece of linear DNA under tension, through a change in *Lk,* results in a change in *Tw* until a critical $Tw_{critical}$ is reached, after which the DNA buckles into a structure in which the helical axis of DNA wraps around itself forming a writhed structure referred to as a

plectoneme[11,12]. Tortional stress, if applied in the direction of the helical axis of DNA, will result in positive supercoiling, i.e., over-winding of DNA ($\Delta Lk > 0$) (Fig 1.1). The tortional stress when applied in direction opposite to the helical axis of the DNA results in negative supercoiling, i.e., under-winding of DNA ($\Delta Lk < 0$) (Fig 1.1). Negative supercoiling of DNA leads to destabilization of individual base pairs.

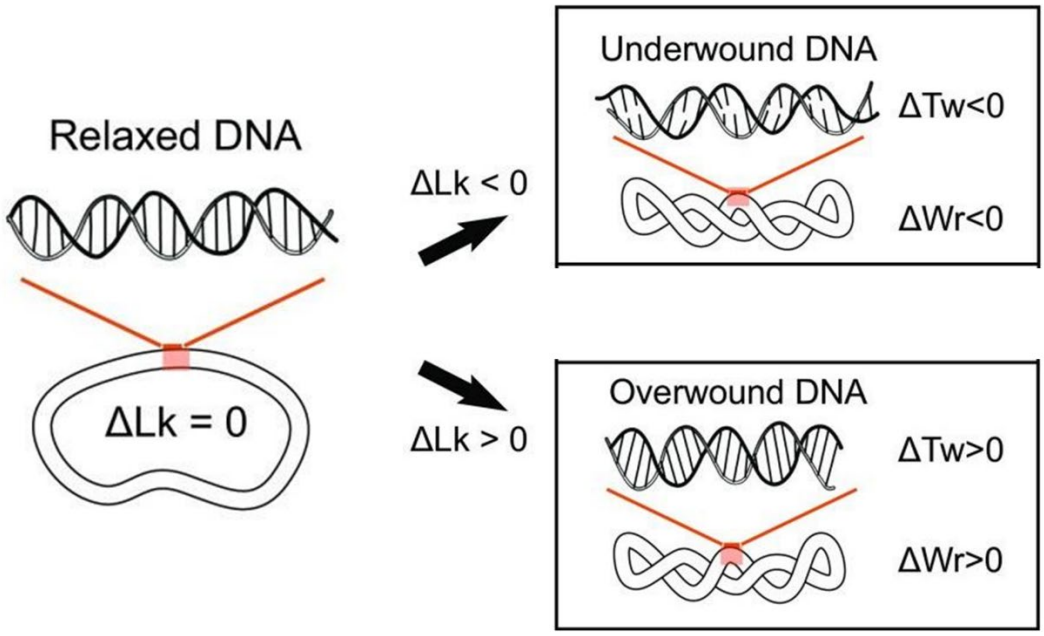

Figure 1.1 Overview of DNA Topology. Relaxed DNA has inherent Tw_0. A change in linking number (ΔLk) is accommodated by a combination of a change in twist (ΔTw) and a change in writhe (ΔWr). Figure adapted from Seol *et al.*[13]

1.1.3. Mismatched base pairs can affect DNA topology

Defects in duplex DNA can arise from base-pair mismatches[14,15], damaged[16] or modified bases[17], locally melted DNA bubbles[18], or a protein introduced kink in the DNA[19]. In the current thesis we focus on a specific type of defect, namely a mismatched base pair, and study its effect on supercoiled DNA. Mismatches in DNA arise during the replication process. DNA replication is performed by DNA polymerase. Generally,

a DNA polymerase will produce one error for every 10^7 nucleotides synthesized. Cells also contain additional proteins that will further reduce this error rate by three orders of magnitude. These proteins that correct mismatches are referred to as mismatch repair proteins (MMR). Usually in cells, a mismatch is detected by the MutS protein. MutS then recruits another repair protein (MutL) in order to start the repair process[20]. It is not well understood how MutS recognizes a mismatch , but crystal structures of MutS bound to DNA have shown that the bound DNA adopts a sharply bent structure with the mismatched base flipped out[21–23](Fig. 1.2 (A)).

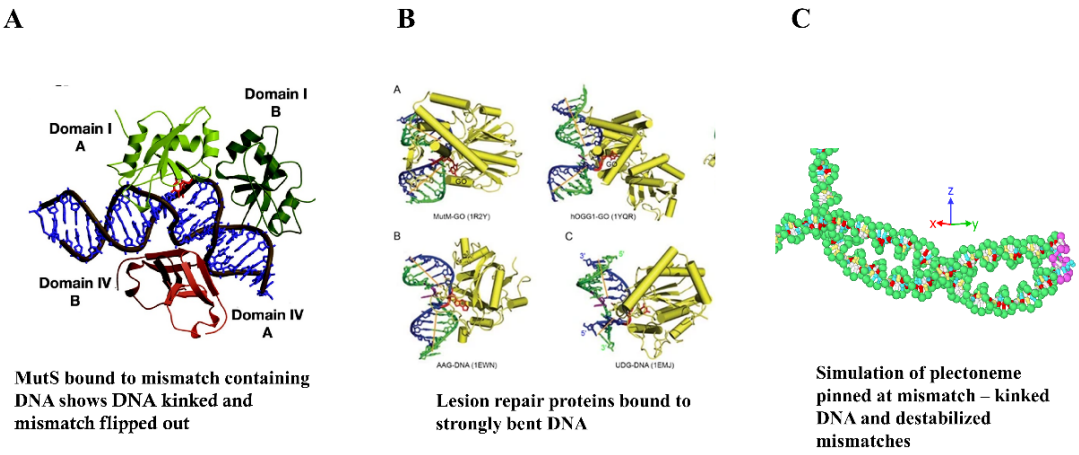

Figure 1.2 Repair proteins in complex with DNA. **(A)** Crystal structure of Taq MutS bound to DNA containing G:T mismatch. The DNA is kinked, and the T base is flipped out. Adapted from Obmolova et al.[20,22] **(B)** DNA glycosylases bound to DNA containing defects. The defected bases are flipped out. Adapted from Yang et al.[24] **(C)** Simulation snapshot from the current work showing supercoiling can cause kinking and base destabilization at the mismatch (the mismatch is shown by purple spheres).

An elastic rod with a defect that locally decreases the bending rigidity of the rod, when subjected to torsional stress, would buckle at the defect[25]. However, in a genomic DNA that contains millions of base pairs, it is possible that thermal fluctuations and entropic effects could mask the influence of the defect. Alternatively, analogous to the buckling at defect observed in the elastic rod, depending on the type

and the size of the defect, the enthalpic gain associated with buckling at the DNA defect could result in a preferential buckling and plectoneme formation at the defect[14,26]. Furthermore, the type and the size of the defect could significantly alter the structure of the end loop containing the defect; for example, localizing the defect at the sharply kinked tip of a plectoneme may promote the flipping out of the base at the mismatch[27] . Indeed, simulations indicate that the distal end of a plectoneme in a supercoiled DNA can adopt a sharply kinked conformation in which the base is flipped out in intact DNA (Fig. 1.2 (C)). Suggestively, some classes of mismatch recognition proteins (MutS, glycosylase, etc. See Fig 1.2 (A, B)) specifically recognize mismatched or damaged DNA bases through a process in which the DNA is sharply bent and the base is flipped out at the defect site[21–24] (Fig. 1.2 (B)). Recently, a single-molecule FRET study has shown that in a linear DNA substrate, MutS discriminates between intact and mismatched bps by a process that involves both bending and unbending of DNA at the mismatch.

The importance of a sharply bent DNA in mismatch recognition can also be inferred from a recent simulation study. Sharma et al.[28] quantified the effect of a mismatch on the local rigidity of a DNA molecule via all atom simulations. They found that for slightly bent DNA, the presence of a mismatch negligibly affects the bending rigidity (Fig. 2.1). However, when the DNA is strongly bent the decrease in rigidity due to the mismatch becomes significantly more pronounced (by several k_BT or more). It follows, therefore, that discriminating a mismatch from intact duplex DNA based on energetic differences associated with DNA bending requires a sharp bend. Given the similarity in the structure of the distal end of a plectoneme observed in simulations and

the crystal structure of mismatch recognition proteins bound to mismatched DNA (compare Fig. 1.2 (A) and Fig. 2.3 (c)), it follows that recognition of DNA defects *in vivo* could be enhanced if they were localized at the end of plectonemes due to facilitated buckling. Whereas this is an attractive model to achieve improved rate and efficiency of *in vivo* detection of mismatches, the probability as well as the dynamics of plectoneme formation at mismatches have not been established. *In vitro* studies of mismatch repair generally focus on studying the repair process downstream of mismatch recognition by MutS[20,29–31]. One of the steps in mismatch repair involves stand discrimination by MutS-MutL complex. The MutS-MutL complex searches for a methylated d(GATC) site to determine the correct DNA strand to be excised during mismatch repair. Few studies have shown that supercoiling can aid in the mismatch repair by bringing the methylated site close to mismatch[30]. In a recent coarse-grained simulation study, Brackely et al.[32] found that protein target search is significantly shorter if the target sequence is located on a loop of a DNA. Previously, Van den Broek et al.[33] have shown that supercoiling accelerates restriction enzymes' (EcoRV) target site search. To the best of our knowledge, no *in vitro* or *in vivo* experiments have studied the effect of supercoiling on the initial mismatch recognition.

In addition to mismatch recognition proteins, many other DNA binding proteins introduce a bend in DNA[34], or bind two distal binding sites simultaneously to form a DNA loop, a common motif for transcriptional repression or activation[35]. Defects could affect the kinetics and thermodynamics of protein induced bending or looping, which could in turn affect transcription and DNA processing more generally. Finally,

plectonemes are believed to be able to diffuse along DNA *in vivo*[36]and it is unclear how the presence of mismatches could impact this process.

Here, we study the effect of a mismatched base pair defect on the supercoiling of DNA. We provide biological context to our results where appropriate, but the focus of our study is understanding the effect of mismatches on supercoiled DNA.

1.2. Molecular Dynamics Simulation

Molecular dynamics (MD) simulation can be a very useful tool to study the physical mechanism underlying atomistic interactions. Generally, MD simulation involves dividing a system into its constituent atoms and applying Newton's laws of motion to the constituent atoms to study the system's properties.

1.2.1. Simulation studies of DNA

Simulation studies of DNA are usually performed on three length and time scales. Quantum mechanics simulations are usually performed at a scale that considers interactions of each electron in the DNA. These simulations can typically simulate a few base-pairs for a timescale in the order of few pico-seconds. Quantum mechanics simulations can be used to study interactions between individual nucleotides[37], binding of ions to DNA[38], etc.

To simulate a DNA with a few hundred base pairs, all atom simulations can be performed. All atom simulations implicitly account for the electron-electron interactions. This allows the simulations to be performed for a few micro-seconds. Well known all atom force fields like the CHARMM[39,40] family of force fields and the AMBER[41–43] family of force filed have been extensively tested and compared with

experiments[44–46]. All-atom simulations have been used to study the effects of mismatch on DNA bending[28], effect of mismatch on base flipping[47], interaction between DNA and DNA binding proteins[48], etc.

The simulation of a longer DNA can be performed using coarse-grained models. Coarse-graining involves eliminating some degrees of freedom of the DNA molecule, which in turn flattens the free-energy curve and accelerates the dynamics. Different degrees of coarse-graining can be used to study DNA processes at different length and time scales. A popular approach to coarse-graining involves using 2-6 molecules to simulate the motions of each base pair. Using this approach a number of coarse-grained models have been developed to study various expects of DNA mechanics[49–55]. Generally, coarse-grained models can be parameterized using either the "bottom-up" or the "top-down" approach. The "bottom-up" approach uses forces obtained from all atom or quantum mechanics simulations to parametrize the coarse-grained force field. The models develop though the "bottoms-up" approach are suitable to study motions of individual base-pairs because they are parameterized to reproduce interactions of individual bps[49,56,57]. The "top-down" approach is more ad-hoc and the coarse-grained models parameterized using this approach aim to reproduce many of the experimentally relevant properties[51,53,58]. The popular OxDNA[58] model is parameterized using this "top-down" approach. Since we use a sequence dependent version of OxDNA model (OxDNA2) in the current thesis, we discuss the limitations of the OxDNA model in more detail in section 1.2.2.

In order to simulate DNA molecules longer than 10 kilo-base pairs, a continuum-level coarse-grained model is needed. In this coarse-graining approach,

motions of multiple bps are combined into one molecule. In the continuum-level simulations, DNA is modelled as a semi-flexible polymer using the worm-like chain model. The model can incorporate various experimentally measured DNA properties like the bending modulus, persistence length, twist modulus, etc. The worm-like chain allows for simulations of thousands of base pair long DNA for a timescales of the order of seconds. Worm-like chain model has been used to study plectoneme relaxation[59], plectoneme dynamics[60], etc.

1.2.2. OxDNA2 model

The OxDNA2 is a coarse-grained model of DNA[50,61]. The model approximates each nucleotide to be a rigid ellipsoid with three interaction sites. The three interaction sites correspond to the phosphate group, sugar group and base group. The model uses the Debye-Hückel approximations (based on the assumption of weak charge density) to account for the electrostatic interactions at different salt concentrations. The OxDNA2 model incorporates DNA sequence dependence through sequence dependent stacking interactions and sequence dependent hydrogen bonding interactions.

The aim of current thesis is to understand plectoneme pinning at mismatches in a supercoiled DNA. Coarse-grained models, like the OxDNA2 model are needed in order observe supercoiling in the timescale and length scales relevant to experiments. Previously, the sequence independent version of OxDNA (OxDNA1) model has quantitatively reproduced the response of DNA to tension and torsional stress[62]. This was the primary consideration in selection of the OxDNA2 model for simulations in the current thesis. The OxDNA2 model has also been used to study various structural features of DNA[60,63]. Recently, Kriegel et al.[63], used magnetic tweezers, OxDNA2

simulations and all atoms simulation to study the effect of temperature on DNA supercoiling. Kriegel et al.[63] showed that OxDNA2 can semi-quantitively reproduce single molecule magnetic tweezers and all atom observations of temperature dependence of helical twist in DNA. Although, OxDNA2 could reproduce the reduction in twist at higher temperatures, the mechanism of this reduction was completely different than all atom simulations.

Kriegel et al.'s study highlights the advantages and limitayions of using a coarse-grained model, like the OxDNA2 model, to study DNA mechanics under extreme stress. The advantage of coarse-graining is that it provides the ability to study DNA supercoiling in experimentally relevant time scales and length scales. On the other hand, Kriegel et al. showed that, even when OxDNA2 model can reproduce experimental observations, the specific mechanism might not be accurate.

Our aim is to use the OxDNA2 model to study the effect of mismatches on supercoiled DNA. In chapter 2, we discuss some of the limitations that should be considered when using coarse-grained simulations to study phenomena that affect individual nucleotide. Here, we would like to emphasize that, even though we find that OxDNA2 model can reproduce experimentally observed pinning at the mismatch; the exact mechanism of plectoneme pinning may not be accurately represented by the OxDNA2 model. Specifically, the extent of DNA bending at plectoneme end loop may not be accurately captured by the OxDNA2 model. To the best of our knowledge, the effect of multiple mismatches on bending in the plectoneme end loop has not been quantified by any existing experimental or all atom simulation studies.

Here, we would like to highlight that the single molecule experiments developed in chapter 5, can directly test the mechanism of plectoneme pining at the mismatch. Vologodskii et al.[64] have previously shown that strong bends in DNA can result in a reduction in the DNA extension. The single molecule experiments can use this reduction in DNA extension to precisely measure the extent of bending in the plectoneme end loop. The single molecule study will allow us to quantitively compare the mechanism of plectoneme pinning in experiments with the mechanism of pinning observed in simulation.

1.3. Single-molecule magnetic tweezers based measurements of DNA topology

In the past two decades, various single-molecule techniques have been developed to study DNA topology[65]. Single-molecule techniques can be used to manipulate and study DNA conformational change in real time[66–68], these conformational changes of DNA can be used to understand mechanisms of DNA binding proteins[69]. Although, various single molecule techniques have been developed to manipulate and study DNA topology[65], here we focus on traditional single-molecule magnetic tweezers.

Typically, in traditional magnetic tweezers (MT), the molecule of interest is attached to the glass surface of a flow cell at one end and the other end is tethered to a magnetic bead. Magnets placed above the flow cell can be used to apply force and toques on the bead. Since the molecule of interest is tethered at both ends, the force and torque applied to the magnetic bead will stretch and twist the molecule of interest. Response of the molecule to force and torques can be studied by tracking the motion of tethered magnetic bead (Fig. 4.6 (A)).

Conventional magnetic tweezers[70] use permanent magnet to produce a horizontal magnetic field, under which the rotation angle of the magnetic bead is strongly controlled[71]. Precise forces can be applied to the tethered magnetic bead by varying the position of the permanent magnets with respect to the tethered beads. Generally, magnetic tweezers are excellent candidates for constant force experiments, however they can only manipulate the motion of the magnetic bead in one direction (the direction of the magnetic force). Traditional magnetic tweezers can apply precise forces in the range of 10^{-3}-10^{2} pN [67]. The temporal resolution of the magnetic tweezers are limited to a range of 10^{-2} seconds due to Brownian motion of the tethered magnetic bead[72]. The limited temporal resolution makes traditional magnetic tweezers poor candidates to study dynamic transitions. This limitation can be partially overcome by using gold rotor bead tracking (AuRBT)[73].

In a AuRBT assay (Similar to Fig. 4.2), in addition to the magnetic bead, a second smaller gold bead is attached to the molecule of interest. In a Brownian motion limited regime, the temporal resolution of traditional magnetic tweezers is limited by the size of the magnetic bead. The magnetic bead is typically 1 μm in diameter, using 80-140 nm gold particles as additional probes can improve the temporal resolution of the magnetic tweezers by several orders of magnitude[73]. A critical limitation of the current AuRBT is the need of relatively high stretching forces (~ 5 pN) to obtain a good signal to noise ratio for gold bead tracking. In studies requiring supercoiled DNA, the high stretching force limits the suitability of AuRBT to only studies of positively supercoiled DNA. Negative supercoiling may play a crucial role in various DNA-protein interactions. In order to overcome the need of high stretching force, we propose

an alternative attachment method (using a di-thiol group) in chapter 4. The goal of this thesis is to propose and develop an experimental approach in order to study plectoneme pinning in a negatively supercoiled DNA under physiological salt conditions. We plan to use the modified AuRBT technique to study plectoneme pinning.

Chapter 2: Coarse-Grained Modelling of DNA Plectoneme Pinning in the Presence of Base-Pair Mismatches*

2.1 Introduction

The double helical structure of duplex DNA underpins the process of semi-conservative replication while stabilizing the individual strands of DNA and protecting the bases against damage. The intertwining of the two DNA strands also imposes topological constraints that complicate processes involving duplex unwinding[74], but also affords opportunities for long range control of global DNA conformation[35,75] and accessibility to processes involving singe-stranded DNA. The global degree of intertwining of the two DNA strands in the duplex, or equivalently over- or under-winding of the helix is referred to supercoiling[11,12]. *In vivo*, the degree of supercoiling affects cellular processes including gene expression[76], enzyme binding[77]and genome organization[78]. Supercoiling is maintained in a state of dynamic homeostasis through the action of a class of enzymes termed topoisomerases, which control physiological processes of DNA metabolism such as transcription and replication that alter DNA topology[79].

Contents of this chapter are published in *Nucleic Acids Research*, Volume 48, Issue 19, 4 November 2020, Pages 10713–10725.

If dsDNA is topologically constrained, *i.e.* its ends are not allowed to rotate relative to each other, a topological invariant called *linking number* (*Lk*) can be defined:[9,10]

$$Lk = Tw + Wr$$

Lk refers to the global measure of wrapping of the two DNA strands around each other. This wrapping can be partitioned between twist, *Tw*, the degree of intertwining of the two single stands about a common axis, and writhe, *Wr,* the degree of self-wrapping of the dsDNA helix axis.

Torsional stress applied to a piece of linear dsDNA, through a change in *Lk,* results in a change in *Tw* until a critical $Tw_{critical}$ is reached, after which the dsDNA buckles into a structure in which the helical axis of dsDNA wraps around itself forming a writhed structure referred to as a plectoneme[11,12].

Supercoiling of DNA is affected by both external environmental factors including the pH [80,81], tension applied on the DNA[82], concentration of multivalent ions[83]; and intrinsic DNA properties such as DNA sequence[84,85], mismatches[14], or other structural inhomogeneities . The response of duplex DNA to tension and torsional stress has been extensively studied using single-molecule approaches, principally magnetic and optical tweezers[67]. In conjunction with these experimental measurements, theoretical models have been developed to describe the mechanics of torsional DNA buckling and plectoneme formation[86,87]. Computer simulations have also been invaluable for understanding the dynamics and molecular-level details of plectoneme formation[59,60,62,88–92]. Despite these advances, many aspects of DNA supercoiling,

particularly the effects of DNA defects on DNA supercoiling, remain poorly understood.

Here, we study the effect of a mismatched base pair defect on the supercoiling of DNA. Recent single-molecule magnetic tweezers-based experiments by Dittmore *et al.*[14] found that for positively supercoiled (overwound) DNA in the limit of high salt concentration (0.5 M to 1 M NaCl) and high force (>2 pN), a plectoneme forms and remains localized at the location of a single mismatched base pair. This provides a possible mechanism for detection of mismatches by mismatch repair enzymes. However, due to the requirement of a strongly bent end loop to infer plectoneme pinning, the single-molecule experiments could not measure pinning under physiological conditions (0.1 M NaCl salt concentration and forces on the order of 1 pN or less). Brahmachari et al.[26] developed a statistical-mechanical model describing the localization of plectonemes under physiologically relevant conditions (0.1 M NaCl and 1 pN). The results of these calculations indicate that the plectoneme localization at the mismatch becomes probabilistic (i.e., the plectoneme is not always pinned at the mismatch) under physiological conditions due largely to entropic effects. This model used a simple analytical approach in which the mismatch was assumed to result in a local reduction in the bending energy. This combination of theoretical and experimental studies gave intuitive explanations for various phenomenological signatures; however, they do not provide a structural understanding of the microscopic origin of plectoneme pinning by base-pair mismatches. In addition, the effects of defects on negatively supercoiled DNA, a prevalent form of DNA topology *in vivo*, has remained largely

unexplored. Finally, the dynamics of plectoneme pinning is yet another aspect that remains unexplored from a theoretical, computational, or experimental standpoint.

In this chapter, we develop a computational framework within the OxDNA2 model[50,61,93] to study plectoneme localization at mismatches. We verify that results obtained in this framework are consistent with all atom simulations of the effects of defects on DNA bending rigidity (Fig. 2.1 (b)) and can reproduce experimentally and theoretically observed plectoneme localization at high salt and high forces. We use the computational framework to study the effect of force and salt concentration on the localization of plectonemes at mismatches. This framework can also be used to study the effect of defects on the dynamics of the supercoiling process and the effect of mismatches on negatively supercoiled DNA.

The use of computational modelling provides molecular-level details of the plectoneme formation process, which are not accessible with existing experimental or theoretical approaches. We establish a computational approach to model the effects of mismatches on DNA supercoiling that offers a complementary approach to statistical mechanical and experimental approaches: the combination of these approaches will permit a comprehensive understanding of DNA mechanics in the presence of defects.

2.2 Materials and Methods

We perform molecular dynamics (MD) simulations of a DNA molecule using the OxDNA2 model[50,61]. All the simulations presented in the current paper are performed using the sequence dependent variation of the OxDNA2 model[61] implemented in the LAMMPS[94] simulation software by Henrich et al[93]. The OxDNA

model has previously been shown to reproduce the behavior of DNA under tension and torsional stress[62] and has also been used to study various structural features of DNA[60,63,95–97]. In the following sections the simulation parameters are given in the reduced Lennard-Jones (LJ) units, we will provide equivalent parameters in SI units where necessary. The conversion of LJ units to SI units is not always trivial, specifically because coarse graining eliminates many degrees of freedom, the potential energy surface is flattened, and hence the dynamics are accelerated. Here, the LJ units are converted to SI units using the parameters provided on the OxDNA2 model website.

2.2.1 Modelling of mismatched base-pairs

The OxDNA2 model treats each nucleotide as a rigid body with only three interaction sites (base, phosphate, and sugar groups)[50,61]and as such eliminates many degrees of freedom of individual base pairs compared to all atom simulations, but this coarse-graining approximation is necessary to reach length scales and time scales relevant to the study of supercoiling dynamics of DNA. The OxDNA2 model uses sequence dependent stacking and hydrogen bonding interactions; other interactions between nucleotides are independent of nucleotide type[61]. A base pair mismatch in OxDNA2 is simulated by removing the hydrogen bond between the bases. Harrison et al.[97] have previously used the sequence independent version of OxDNA (OxDNA1) to simulate mismatches in a similar manner. Previous studies have shown that many types of mismatched base-pairs have hydrogen bonds[98]; therefore, approximating a mismatched base pair without hydrogen bonds is perhaps a crude approximation. However, as we demonstrate below, such an approximation reproduces the decrease in

bending stiffness at the mismatch obtained from all atom simulations. Eliminating these degrees of freedom may alter the dynamics of individual base-pairs. But here we are interested in studying the effect of these mismatches on the bending rigidity of the local structure of duplex and this bending rigidity is not altered due to this approximation (as shown by the favorable comparison with the all atom simulation result, see Fig. 2.1 (b) below). Ditmore et al.[14] and Brahmachari et al.[26] have shown that the reduction in local bending rigidity due to the presence of mismatched base-pairs promotes plectoneme pinning.

Our goal is to develop an approach within the oxDNA2 model to accurately model the buckling of DNA containing mismatched base pairs. To this end we first measure the free energy required to bend a 15 bp long DNA with different mismatches parameterized in oxDNA2 and compare the results to all atom simulations performed by Sharma et al.[28] We use the same sequence of DNA as used in Sharma et al[28]. To calculate the free energy of bending we simulate a 15 bp long DNA with the mismatch located at the center (8th base pair). We employ umbrella sampling to efficiently sample the free energy landscape. The bend angle (θ) is the angle defined between the center of mass of 3 blocks of 5 bp's (block 1: 2-5 bp, block 2: 6-10 bp, block 3: 11-14 bp), the mismatched base pair is located at the center of the DNA (base 8/23) (Fig. 2.1 (a)). The ends of the DNA are unconstrained. We then apply a harmonic constraint, of type

$$U(\theta) = \frac{k}{2}(\theta - \theta_{ref})^2$$

using the COLVARS[99] module in LAMMPS. Here $k = 0.064\left(\epsilon_{LJ}/\mathrm{rad}^2\right)$. We carry out both forward and backward sampling of the free energy landscape by first gradually

decreasing θ_{ref} from 180° to 90° in steps of 5°, and then increasing θ_{ref} from 90° to 180° in steps of 5°. We simulate each step for $10^4\tau_{LJ}$ (30.3 ns) of which the first $10^2\tau_{LJ}$(0.3 ns) are equilibration steps and not considered in subsequent data analysis. The MD simulations are carried out in the canonical constant number of particles, constant volume, and constant temperature (NVT) ensemble using the Langevin thermostat with a damping factor of 0.5 $\tau_{LJ}$and 10 τ_{LJ} for the transitional and rotational degrees of freedom, respectively. The equations of motion are integrated using the velocity-Verlet algorithm with a timestep of $0.001\tau_{LJ}$ (3.03 fs). We can then extract the free energy associated with bending using the weighted histogram analysis method (WHAM)[100] as implemented by Grossfield (http://membrane.urmc.rochester.edu/?page_id=126).

Sharma et al.[28] analyzed the results of their all atom simulations using the definition of angle provided by the COLVARS module and the definition of angle used in the 3DNA software. Sharma et al.[28] also noted that their results are slightly affected by the definition of bend angle.

We use the bending free energy obtained by Sharma et al.[28] using the 3DNA software to compare the oxDNA2 model with all atom simulation. The 3DNA software fits a linear helical axis to the two fragments of a bent DNA molecule. The bend angle is defined as the angle between these two helical axes. This definition of the bend angle closely resembles the definition of bend angle for a coarse-grained model like oxDNA2.

We find that for a G:T mismatch the reduction in bending rigidity observed using the OxDNA2 model (Fig. 2.1 (b)) is similar to the reduction observed by Sharma

et al.[28] using an all atom model (within an error of ~ 4 k_BT) . We also provide the bending free energy of an intact G:C base pair using the OxDNA2 model for comparison (Fig. 2.1 (b)). We find that the bending free energy of G:T mismatch in the OxDNA2 model is higher than the free energy obtained from all atom simulations. In the context of the simulation results presented here, this implies a higher bending rigidity of OxDNA2 model: accordingly, the probabilities of plectoneme pinning will be underestimated in the simulation work here.

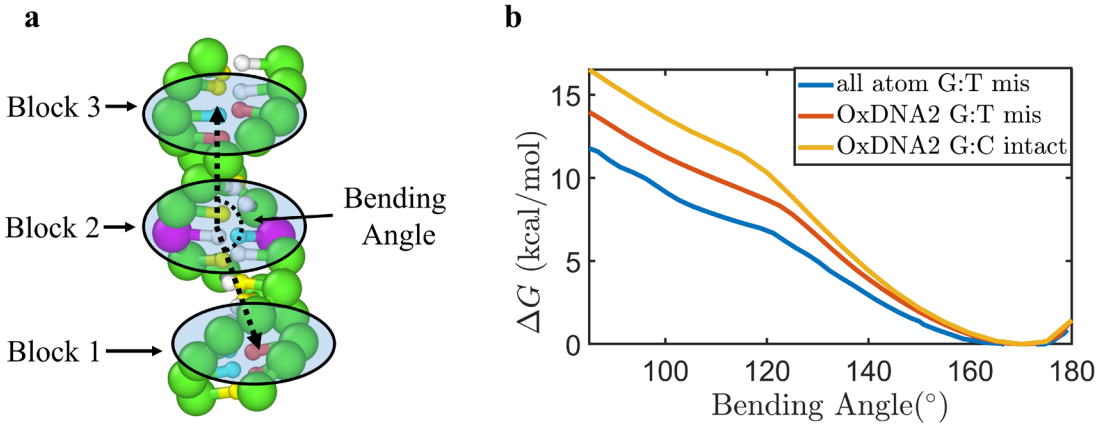

Figure 2.1 Bending free energy of G:T mismatch. **(a)** Illustration of the bending angle constraint applied for umbrella sampling. Bending angle is calculated as the angle between vectors connecting center of mass of Block 2 to Block 1 and the vector connecting center of mass of Block 2 to Block 3. Green sphere represent intact bps, purple sphere represents G:T type mismatched bps. **(b)** Comparison of free energy versus bend angle for a G:T mismatch obtained using an all atom model[28] (blue) and the OxDNA2 model (red). Free energy vs bend angle of a intact G:C bp (yellow curve) obtained using the OxDNA2 model are provided for reference.

2.2.2 Intrinsic bend of DNA

Kim et al.[101] have shown that DNA sequence can influence plectoneme pinning. They found that intrinsic bend of a DNA sequence can be a significant factor in

determining plectoneme pinning. We study the intrinsic bend in the DNA sequence used, to ensure we do not have a significantly bent section in the simulated DNA.

Previous studies[101] used dinucleotide steps to calculate the intrinsic bend in a DNA segment. Here instead of using the dinucleotide step method, we explicitly calculate the bend angle at various bp steps. The dinucleotide step method does not account for force applied to DNA and how that might affect the intrinsic bend at various base-pairs. To overcome these limitations, we explicitly calculated the intrinsic bend at various base pairs.

To calculate the intrinsic bend, we conduct simulations of stretched and untwisted DNA to calculate the twist stored in a DNA molecule with mismatched base pairs. We perform simulations for a 610 bp and 1010 bp DNA with either no mismatched bases, or with 2, 4, or 6 consecutive mismatched bases in the center of the DNA. We perform a single simulation for each mismatch size under different force and salt conditions. The procedure used to perform these simulations is similar to the procedure used for DNA under torsional stress. Simulation for the 610 bp DNA under the high force-high salt condition is carried out for $2.3 \times 10^7 \tau_{LJ}$ ($69.69\ \mu s$). Simulation for the 610 bp DNA under the low force-low salt condition is carried out for $1.8 \times 10^7 \tau_{LJ}$ ($54.54\ \mu s$), and simulation for 1010bp DNA under low force-low salt condition are carried out for $1.6 \times 10^7 \tau_{LJ}$ ($48.48\ \mu s$).

To calculate the bend angle at a bp, we use the same method used to calculate the end loop angle as described in the Methods section. Briefly, the bend angle at a base-pair is defined between the center of mass of 3 blocks of 5 bp. If the bp is at

position i, the 3 blocks would be block 1: i-12 to i-8 bp, block 2: i-2 to i+2 bp, block 3: i+8 to i+12 bp.

Fig. 2.2 provides the bend angles at each base pair along the DNA. We find for intact DNA (0 mismatches) the bend angle is almost constant along the DNA. For a DNA with mismatches, we find that the bend angle is significantly lower (i.e., the DNA is highly bent) at the location of mismatch. We also find that the bend angle at the mismatch decreases with more mismatches.

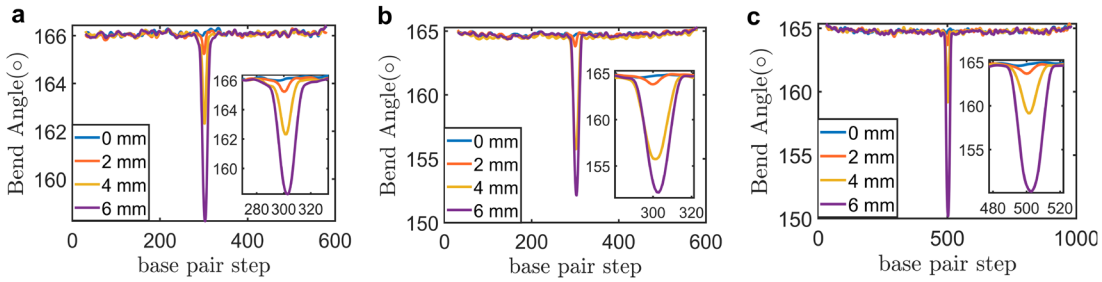

Figure 2.2. Intrinsic bend. (a-c) Average angle at each bp along the DNA backbone for different number of mismatches introduced at the center of the DNA for (a) 610 bp DNA with F=2 pN and 1 M monovalent salt, (b) 610 bp DNA with F=0.3 pN and 0.2 M monovalent salt, and (c) 1010 bp DNA with F=0.3 pN and 0.2 M monovalent salt. Insets provide a magnified view of the angle at the location of mismatches.

2.2.3 Effect of number of mismatched base pairs on DNA bending

To calculate the effect of number of mismatches on the free energy of bending, we simulate a 30 bp long DNA with the mismatch located at the center. We employ umbrella sampling to efficiently sample the free energy landscape. The method used to calculate the effect of number of mismatches on DNA bending is similar to the method used in section 2.2.1. The collective variable (θ) is the angle defined between the center of mass of 3 blocks of 5 base pairs (block 1: 4-8 bp, block 2: 13-17 bp, block 3: 23-27

bp). The ends of the DNA are unconstrained. We then apply a harmonic constraint, of type

$$U(\theta) = \frac{k}{2}\left(\theta - \theta_{ref}\right)^2,$$

using the COLVARS[99] module in LAMMPS. Here $k = 0.064\left(\epsilon_{LJ}/rad^2\right)$. We carry out both forward and backwards sampling of the free energy landscape by first gradually decreasing θ_{ref} from 180° to 30° in steps of 3°, and then increasing θ_{ref} from 30° to 180° in steps of 3°. We simulate each step for $3 * 10^4 \tau_{LJ}$ (90.9 ns) of which the first $750\tau_{LJ}$ (2.25 ns) are equilibration steps and not considered in subsequent data analysis. The MD simulations are carried out in the NVT ensemble using the Langevin thermostat with damping factors of 1 $\tau_{LJ}$and 10 τ_{LJ} for the transitional and rotational degrees of freedom, respectively. The equations of motion are integrated using the velocity-verlet algorithm with a timestep of $0.003\tau_{LJ}$(9.09 fs). We then extract the free energy associated with bending using the weighted histogram analysis method (WHAM)[100] as implemented by Grossfield (http://membrane.urmc.rochester.edu/?page_id=126).

2.2.4 Simulation Details

For the supercoiling simulations, we consider a 610 and 1010 bp long DNA with 50% G-C content and 0, 2, 4, or 6 consecutive mismatches at the center of the DNA. We constrain the 5 base pairs at the top end of the DNA in the x-y plane and constrain 5 base pairs of the other end in the x-z plane. This allows the DNA ends to

freely move in the y and z directions in response to tension and torsional stress while ensuring that the boundary base pairs do not rotate so that the super-helical density remains constant during the simulation (Fig. 2.3). To ensure that the nucleotides don't pass around the DNA ends, we apply a purely repulsive harmonic potential of type $U(z) = k(z - z_{ref})^2$ that acts in the x-y plane and repels all but the boundary base pairs if they move beyond the boundary base pair in the z direction (Fig. 2.3). Here $k = 1000(\epsilon_{LJ}/rad^2)$ and z_{ref} is $5\sigma_{LJ}$away from the last base pair at each end.

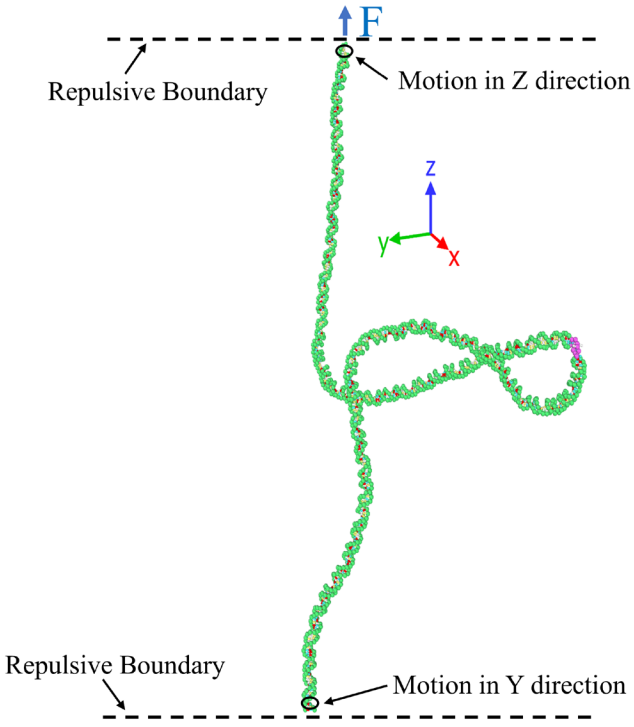

Figure 2.3 Simulation boundary condition. Once the tortional stress is applied, one end of the DNA can freely move in Y direction and the other end of DNA can freely move in Z direction. A fixed force F is applied to the top end. A repulsive boundary condition applied to the ends of DNA ensures that no bps pass the DNA ends.

We define the equilibration time τ_{equi}using the autocorrelation function of the DNA end-to-end distance (R_{end}). The auto-correlation function of ${R_{end}}^{60}$, is calculated as

$$C_x(t) = \frac{\langle (R_{end}(t) - \langle R_{end} \rangle)(R_{end}(0) - \langle R_{end} \rangle) \rangle}{\left(\langle {R_{end}}^2 \rangle - \langle R_{end} \rangle^2 \right)}$$

The equilibration time is then given by

$$\tau_{equi} = \int_0^{t_{cut}} C_x(t)\, dt.$$

Here the upper limit of the integral t_{cut} is selected as the first time when $C_x(t) = 0$. It is necessary to use this limit to eliminate the effect of the noisy tail of $C_x(t)$. To collect data, we discard the first $2 * \tau_{equi}$timesteps. In the current work, all the simulations have been carried out for at least $12 * \tau_{equi}$.

We start each simulation from an equilibrated condition. We apply torsional stress by rotating the top end of the DNA around the z axis for fixed number of turns (N_{turns}). To select the value of N_{turns}, we first conducted preliminary simulations and then selected the smallest value of N_{turns} for which we could obtain a stable plectoneme within a reasonable simulation time. A larger value of N_{turns} would produce a larger plectoneme that would stabilize faster and hence require less simulation time. On the other hand, a larger plectoneme will restrict the motion of the plectoneme along the DNA backbone and hence it would require a longer simulation to accurately capture the entropic effects of plectoneme diffusion along the DNA backbone.

To apply tension, we apply a force along the z-axis to the top base-pair. For all simulations we use the Langevin C[102] thermostat with a timestep of $0.01\tau_{LJ}$, and unless otherwise specified, damping factors of 5 τ_{LJ} and 10 τ_{LJ} for the transitional and rotational degrees of freedom respectively. The Langevin thermostat models the interaction of particles with an implicit solvent. The viscosity of implicit solvent can be controlled by the varying the damping factors.

For simulations at a salt concentration of 1 M, we performed 5 simulations (using different random numbers as the seed for the Langevin thermostat) of a 610 bp long DNA for each mismatch value. Simulations were performed for $2.6 \times 10^7\tau_{LJ}$ $(78.78\,\mu s)$ for the case of 0 mismatches; $2.7 \times 10^7\tau_{LJ}$ $(84.08\,\mu s)$ for 2 mismatches, and $2 \times 10^7\tau_{LJ}$ $(60.6\,\mu s)$ for 4 and 6 mismatches. For 4 and 6 mismatches, we used a damping factor of 15 τ_{LJ} for the rotational degrees of freedom. For simulations at a salt concentration of 0.2 M, we performed simulations of 610 bp and 1010 bp DNA molecules. For the 610 bp DNA, we performed 20 distinct simulations for each mismatch size for approximately $3.5 \times 10^7\tau_{LJ}$ $(105\,\mu s)$ using a timestep of 0.01 τ_{LJ} and a damping factor of 15 τ_{LJ} for the rotational degrees of freedom.

For the 1010 bp DNA we performed 10 simulations for each mismatch. The simulations were run with a timestep of $0.01\tau_{LJ}$ for approximately $3.7 \times 10^7\tau_{LJ}(112.11\mu s)$. Combined these simulations required a total of 5.2×10^6 CPU hours. The simulations were performed using NIH high performance computing system, Biowulf.

2.2.5 Plectoneme detection algorithm

To detect a plectoneme we used a modified version of the algorithm provided by Matek et al.[62] The plectoneme detection algorithm is as follows:

- Find the center of mass of each base pair (bp) in the DNA.
- Start at the bottom bp and loop over all bp.
- Find the distance between the current bp (i) and all other base pairs (j) beyond a cut-off of N_c bp along the DNA backbone, i.e., distance between bp i and bp j, for all $j > N_c$.
- If any distance is less than $d_{cuttoff}$; the bp is identified as the beginning of plectoneme.
- To detect the end of plectoneme we utilize the concept of contact points. Once the beginning of the plectoneme is detected, we skip the next N_c bp's along the backbone and identify the bp closest to the bp identified as the start of the plectoneme. This bp is defined as the plectoneme end.
- The plectoneme center is defined as the mean of the bp indices of the beginning and the end of the plectoneme.

We typically choose $N_c = 80$, and $d_{cuttoff} = \frac{N_c \times 0.34}{2} nm$. The algorithm allows us to detect multiple plectonemes, but the size of each plectoneme detected is larger than N_c. We very rarely observe multiple plectonemes for the conditions simulated here. If multiple plectonemes are observed, the center of each plectoneme is considered to calculate the plectoneme pinning probability. We also verified that changing N_c does not significantly alter the results.

2.2.6 Calculation of end loop angle

The end loop angle is defined between the center of mass of 3 blocks of 5 bp. If the center of the plectoneme is considered as base *i,* the 3 blocks would be block 1: *i*-12 *to i*-8 bp, block 2: *i*-2 *to i*+2 bp, block 3: *i*+8 *to i*+12 bp.

We also note that our plectoneme detection algorithm does not always provide the exact center of the plectoneme. For the case of high force and high salt concentration, where the plectoneme center is pinned at the location of the mismatch, the plectoneme detection algorithm calculates the center of the plectoneme with an error of around ± 10 bp. Keeping this in mind, to calculate the angle of the end loop we vary the center of the plectoneme between *i* ± 15 bp and select the lowest angle as the angle of the end loop.

2.2.7 Calculation of twist

To calculate the twist between two consecutive base pairs, we use the method defined by Skoruppa et al. [103]; specifically the definition of *Triad II*, as defined by Skoruppa et al., is used for all calculations.

Briefly, the twist for each bp step can be calculated by defining two vectors, one vector connecting the centers of mass of two nucleotide in a bp ($\hat{b}_{bp}$) and the second vector connecting two base pairs along the DNA contour($\hat{n}_{bp}$). Twist can then be defined as the angle between the two ($\hat{b}_{bp}$) for each base pair in the base pair step along the ($\hat{n}_{bp}$). A more detailed description can be found in Skoruppa et al[103].

2.2.8 Calculation of writhe

To calculate the writhe, we carried out simulations of DNA under tension but no torsional stress. We analyzed these simulations to obtain $Lk_0 = Tw_0$, which is the twist inherent in the DNA with different number of mismatches.

Writhe in DNA is then defined as:

$$Wr = N_{turns} - (Tw - Tw_0)$$

where, Tw is the average twist in DNA under tension and torsional stress.

2.2.9 Calculation of diffusivity

To obtain the diffusivity of the plectoneme we first calculate the mean squared displacement (MSD) of the plectoneme center P_{cent}.

$$MSD(\Delta t) = \frac{\sum_{i=1}^{t_{total}-\Delta t}((P_{cent}(i+\Delta t) - P_{cent}(i)) * 0.34 * 10^{-3})^2}{t_{total} - \Delta t}$$

We calculate MSD for $\Delta t \in \{0,20\}$frames, where Δt is the time lag, we select trajectories that have at least 400 continuous plectoneme centers. We then average all the MSD curves for a given condition

The diffusion constant can be obtained from the MSD curve using the relation $MSD = 2Dt$, we perform linear fitting for the first 8 points on the MSD vs time curve to obtain the diffusion constant. The MSD curves are provided in Fig. 2.9 (e-g).

2.2.10 Visualization

Snapshots of the simulation configurations is carried out using the OVITO[104] simulation software. LAMMPS[94] provides the center of mass and orientation of each nucleotide in the form of a quaternion. We post-process this information using MATLAB to find the location of base interaction sites and the backbone site for all nucleotides. The process used is similar to the process employed in the visualization code provided by Oliver Henrich with the OxDNA module in LAMMPS.

2.3 Results and Discussion

2.3.1 Coarse-grained estimates of the effects of DNA defects on bending free energy

Coarse-grained simulations of DNA, in particular OxDNA2, afford the requisite trade-offs between molecular-level details and computational speed and efficiency to permit millisecond-scale simulations of kilobase length DNA[60]. OxDNA has been successfully applied to study DNA supercoiling and has been benchmarked against single-molecule measurements of DNA supercoiling[62]. For these reasons we decided to develop a mismatch model of DNA in the OxDNA2 framework. We first validated the effects of defined mismatches on DNA bending elasticity for OxDNA2. Fig. 2.4 (a,b) shows the variation of the total Gibbs free energy (ΔG) of the DNA molecule as a function of the bend angle imposed at the center of the DNA molecule containing different numbers of consecutive mismatches at the center of the DNA for high (Fig. 2.4 (a)) and low (Fig. 2.4 (b)) salt concentrations. A bending angle of θ_b=170° corresponding to a nearly straight DNA, is the lowest free energy configuration of DNA with and without mismatches. Previous studies[105–107] have also

found the lowest free energy configuration to be at an angle slightly less than 180° (i.e. a completely straight segment of DNA). Bending at the DNA center reduces this angle. The presence of mismatches reduces the bending energy of the DNA at the location of the mismatch[28]. Accordingly, ΔG increases from the minimum at θ_b=170° as the included angle decreases for all DNA molecules, but for a given bend angle the energy progressively decreases with increasing mismatch size, from a maximum for no mismatch. This data, establishing the effect of mismatch size on bending energy, is in qualitative agreement with results from previous studies[14,26,28] and suggests that the mismatch model is feasible for investigating DNA plectoneme formation in the presence of mismatches through coarse-grained approaches. It is difficult to quantitatively compare the effect of multiple mismatches on bending rigidity observed in OxDNA2 model to existing studies. To the best of our knowledge, the effect of multiple mismatches on bending free energy has not been quantified in experiments or all atom simulations.

We next performed simulations of positively supercoiled DNA (610 bp and 1010 bp) containing 0, 2, 4, or 6 G:T mismatches at the center of the DNA molecule (Fig. 2.4 (c) and 2.4 (d)). To compare the simulation results with experimental and theoretical results, simulations were run with two different conditions of monovalent salt and force: high monovalent salt (1 M) and high force (2 pN); and low monovalent salt (0.2 M) and low force (0.3 pN) (Fig. 2.4 (c) and 2.4 (d)). We refer to the conditions of 1 M monovalent salt and 2 pN of force as the high force - high salt condition, and conditions of 0.2 M monovalent salt and 0.3 pN as the low force - low salt condition.

We have summarized the different simulation condition in Table 2.1. The most important feature distinguishing the plectonemes formed in these two cases is the angle of the end-loop (defined in the insets of Fig. 2.4 (c, d) and in the methods section). The end-loop angle is smaller under the high force and salt condition, indicating a larger extent of deformation or kinking. For example, in the snapshots shown in Fig. 2.4 (c, d), this angle is 46.8^o and 143.3^o for the high and low force and salt conditions, respectively.

Table 2.1 Details of simulation conditions.

DNA Length (bp)	**Force Applied (pN)**	**Implicit Monovalent Salt concentration (M)**	**Turns applied (ΔLk)**	**Referred to as**
610	2	1	3.46	High force - High salt
610	0.3	0.2	2.31	Low force - Low salt
1010	0.3	0.2	2.87	Low force - Low salt

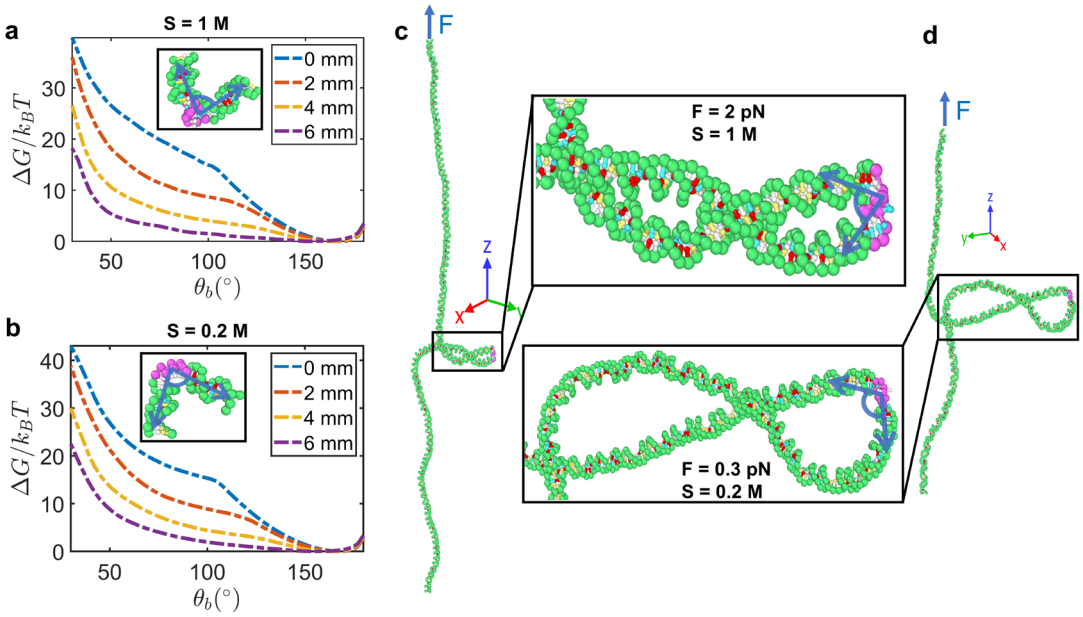

Figure 2.4 DNA mismatches lower bending energy and nucleate plectonemes in supercoiled DNA. (a, b) Free energy of a 30-bp DNA molecule as a function of the bending angle θ_b (in degrees) imposed at the center of the DNA (i.e., the 15th bp) (see also the methods section). The results are shown for (a) high salt concentration (1 M monovalent salt) and (b) low salt concentration (0.2 M monovalent salt). For both cases, there is no external stretching force applied to the DNA and the results are shown for different numbers of consecutive G:T mismatches, introduced at the center of the DNA. The insets illustrate the DNA configuration (with 6 consecutive G:T mismatches introduced at the center of the DNA) corresponding to a bending angle of approximately 95°, indicated by the blue arrows. Here the green spheres represent the sugar group of the intact base pairs whereas pink spheres represent the sugar group of the mismatched bases. Red, yellow, white, and blue spheres represent the A, C, G and T base groups, respectively. **(c, d)** Simulation snapshots of DNA plectonemes formed by positively supercoiling (over-winding) a 610-bp DNA molecule with 6 mismatched bases under the conditions of (c) High salt (1 M monovalent salt) and high force (F = 2 pN) and (d) Low salt (0.2 M monovalent salt) and low force (F = 0.3 pN). The DNA is subjected to a constant force one end, while the other end is fixed. The torque generated in twisted DNA (3.46 and 2.31 turns in (c) and (d), respectively) results in buckling of the DNA to form a plectoneme. Expanded views of the plectonemes along with the definition of the bend-angle at the mismatch location (blue arrows) are shown in the insets of (c) and (d). The plectoneme end loop angle is 46.8° and 143.3° for the high force - high salt and low force - low salt and salt conditions, respectively. A detailed description of the end loop bend angle can be found in the methods section.

Kinking the end-loop brings negatively charged segments of the DNA closer to each other and contributes an additional electrostatic free energy cost to plectoneme

formation. Under high salt conditions, the electrostatic repulsion is screened over a much smaller distance leading to a reduced Debye length. This, in turn, reduces the electrostatic repulsion between the buckled segments and hence reduces the energetic cost of buckling. Furthermore, the external force imposes a second energetic cost associated with buckling related to the work done against the force in decreasing the extension of the DNA by an amount equal to the loop size. As a result, smaller loops with increased kinking are favored under higher applied forces. Previous theoretical and experimental studies obtained similar force and salt concentration dependent increases in the kinking of the end loop[108]. Our results demonstrate that kinking is enhanced in the presence of mismatches, but that the extent of kinking is governed by the applied force and salt concentration.

2.3.2 Mismatches enhance plectoneme pinning

A key question related to the supercoiling of DNA containing mismatches is the degree to which the mismatch localizes or pins the plectoneme by stabilizing a sharp bend at the plectoneme tip. Practically, this is related to both the degree of bending or kinking at the plectoneme tip and the probability of the mismatch being located at the tip of the plectoneme. Consistent with previous studies[108], we observe a decrease in the end loop angle between the high force - high salt and low force - low salt conditions.

For both the simulation conditions, we observe a decrease in the end loop angle (when the plectoneme center is pinned at the mismatches) with increasing number of mismatches (Fig. 2.5 (d-f)). This is consistent with the decrease in the local bending

energy in proportion to size of the mismatch (Fig. 2.4 (a, b)). Previous experimental studies [101] have shown that the intrinsic bend of a DNA segment can affect plectoneme pinning. We characterized the intrinsic bend of the DNA sequence used here to ensure that the specific DNA sequence does not affect the plectoneme pinning probability (see section 2.2.2).

The probability of the mismatch being localized at the tip of the plectoneme depends on the DNA length, applied force, monovalent salt concentration, and mismatch size. Under the high force - high salt condition (Fig. 2.5 (a)), the center of the plectoneme always coincides with the location of the mismatch for 2, 4, and 6 mismatches. This observation is in agreement with the single molecule magnetic tweezers experiments conducted by Dittmore et al.[14] and results of previous simulation study[109]. Furthermore, the DNA is increasingly sharply bent at the mismatch as the number of mismatched bases increases (Fig. 2.5 (d)). However, for the low salt - low force condition there is a non-unity (<1) probability that the mismatch is located at the center of plectoneme, though this probability rapidly increases with the number of mismatches (Fig. 2.5 (b, c)). This result, which has been predicted theoretically[26], is described in detail below. The finite probability of buckling at the mismatch for the low salt and low force conditions leads to a bifurcation of the end loop angle: the end loop angle is constant, independent of the mismatch size, when the mismatch is not located at the tip of the plectoneme but decreases significantly with increasing mismatch size when the mismatch coincides with the tip of the plectoneme (Fig. 2.5 (e, f)). In Fig. 2.5 (d-f), we show the effect mismatches on plectoneme end loop angle. We use standard deviation (SD) to show the variations in plectoneme end loop angle. In

MD simulations the plectoneme end loop is not fixed. We observe that the plectoneme end loop angle is highly variable (similar to Fig. 3.6). If the plectoneme is pinned at the mismatch, the plectoneme end loop angle will vary in accordance with the bending rigidity at the mismatch. In the low force – low salt condition (Fig. 2.5 (e, f)), we observe that the variation in plectoneme end loop (as captured by the SD) is lower when the plectoneme is located away from the mismatch. When the plectoneme is pinned at the mismatch, the SD in plectoneme end loop angle increases with higher number of mismatches. Compared to intact DNA, the mismatches are more flexible and hence more prone to thermal motion. MD simulations can capture these variations in end loop angle.

Here we note that the standard deviation (SD) is not a good measure of error in the simulations. The variability of the plectoneme end loop is a feature of the simulation and this variability in plectoneme end loop structure has also been observed in previous experiments of supercoiled DNA[110]. Here, the standard error of mean (SEM) is a more appropriate measure of the error in calculation of the plectoneme end loop. SEM provides a measure to ensure that MD simulations have sufficiently captured the variation in plectoneme end loop angle. In Fig. 2.5 the SEM is smaller than the symbol size.

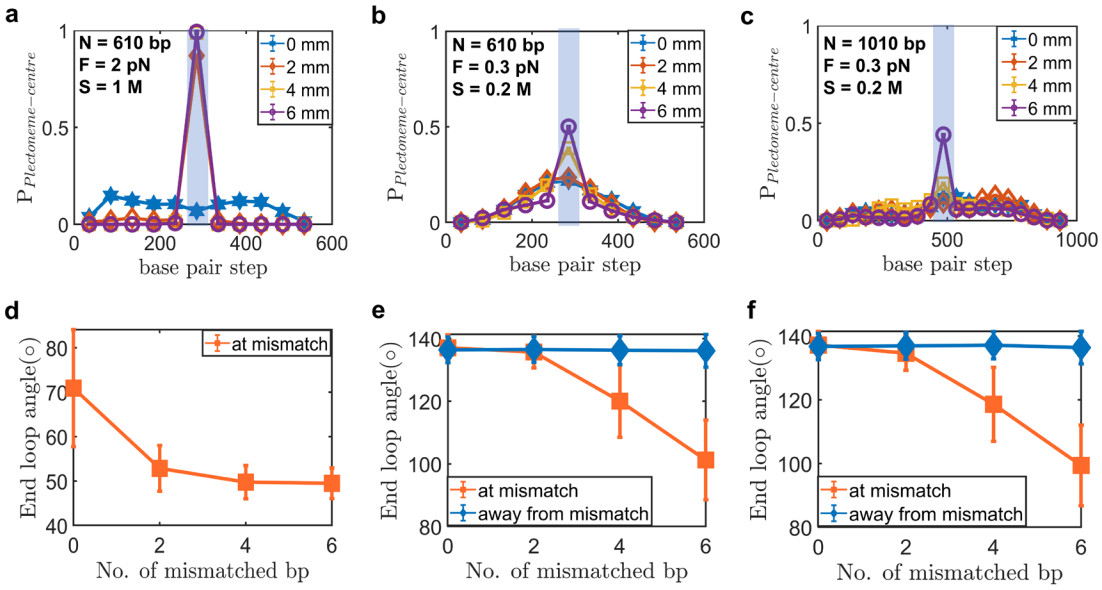

Figure 2.5 Plectoneme end loop kink angle and position. (a-c) Probability distribution of the plectoneme center position. Probability of finding the plectoneme center ($P_{Plectoneme\text{-}center}$) at a given DNA base pair for (a) 610 bp DNA with F=2 pN and 1 M monovalent salt, (b) 610 bp DNA with F=0.3 pN and 0.2 M monovalent salt, and (c) 1010 bp DNA with F=0.3 pN and 0.2 M monovalent salt. The location of the DNA mismatch is denoted by the blue shaded region. Error bars (standard error of the mean, SEM) are smaller than the symbol size.
(d-f) Average end-loop angle in degrees (see the details of the averaging procedure in the Methods section) of the DNA plectoneme [this angle is defined by blue arrows in the insets of Figs. 2.2 c and Figs. 2.2 d] as a function of mismatch size (in base pairs) for (d) 610 bp DNA with F=2 pN and 1 M monovalent salt, (e) 610 bp DNA with F=0.3 pN and 0.2 M monovalent salt, and (f) 1010 bp DNA with F=0.3 pN and 0.2M monovalent salt. The orange line represents the average end loop angle when the plectoneme center is pinned at the mismatch, whereas the blue line represents the average end loop angle when the plectoneme center is not pinned at the mismatch. Error bars represent standard deviation (SD).

The difference in the plectoneme localization between the two conditions (high salt-high force and low salt-low force) can be explained through a simple free energy argument.

For the high force - high salt case, bending or kinking at the mismatch is favored (Fig. 2.4 (a)). Fig. 2.4 (a) indicates that sharper bending requires higher energy. Higher forces will lead to a stronger bending of the end loop since this minimizes the extent of the plectoneme loop, which in turn minimizes the change in extension of the DNA and

the corresponding work against the external force. In this scenario it will be energetically favorable if the bending occurs at the location of the mismatch where the bending rigidity is decreased. However, there is an entropic cost of buckling at the mismatch (equivalent to localizing the plectoneme at the mismatch), which limits the possible configurations of the plectoneme as compared to the case where the plectoneme can form at any location on the DNA molecule. Mathematically, we can write this as:

$$\Delta G \approx \Delta G_{bending} = \Delta H_{bending} - T\Delta S_{bending} \quad (1)$$

where ΔG, ΔH, and ΔS are the changes in the Gibbs free energy, the enthalpy, and the entropy, respectively.

For the case of a high salt - high force, the plectoneme is pinned at the location of the mismatches leading to a large loss of entropy. Therefore, we can split the energy contributions as follow:

$$(\Delta G)_{high-salt} \approx \Delta G_{bending,high-salt} = \Delta H_{bending,high-salt} - T\Delta S_{bending,high-salt}$$
$$= (Large\ negative) + (Large\ positive) \quad (2)$$

The equilibrium is driven by the competition between a large favorable bending enthalpy (due to the localization of the bending deformation at segments with lower bending rigidity) and a large unfavorable bending entropy (due to the pinning of the plectoneme center at the mismatch). In the current work, we find that for a force of 2 pN applied to a 610 bp DNA with 2, 4 and 6 consecutive mismatched base pairs, the enthalpic gain due to bending at the mismatch overcomes the entropic loss due to pinning; hence the plectoneme is always pinned at the mismatch (Fig. 2.5 (a)).

Next, we consider the free energy for the low salt - low force condition. Here too, we can write the free energy change as:

$$(\Delta G)_{low-salt} \approx \Delta G_{bending,low-salt} = \Delta H_{bending,low-salt} - T\Delta S_{bending,low-salt}$$

$$= (Small\ negative) + (Small\ positive) \quad (3)$$

Here both the enthalpy and the entropic components are smaller in magnitude. The equilibrium is driven by the competition between a weakly favorable bending enthalpy (due to a higher bending angle of the end loop, the enthalpic difference for bending at the defect for different numbers of mismatches is lower as compared to the high force case) and a weakly unfavorable pinning entropy (due to a larger plectoneme for the low force-low salt case compared to the high force-high salt case, the entropic cost of pinning is lower for a given length of DNA). In the case of low force - low salt condition for a 610 bp DNA, we find an increase in probability of plectoneme pinning with increasing number of mismatches. To probe the effect of the entropic contribution to the pinning probability, we studied a 1010 bp DNA with 0, 2, 4 and 6 bp mismatches under the same conditions. Pinning in a longer DNA result in a larger entropic loss due to the presence of a larger number of configurations arising from plectoneme diffusion. Consistent with this reasoning, we find that for the 1010 bp DNA, the enthalpic gain due to bending at the mismatch is only able overcome the entropic loss for the case of 6 mismatches (Fig. 2.5 f). We note that the arguments related to the relative entropic and enthalpic contributions to the Gibbs free energy of bending are intended to provide a qualitative explanation of the pinning process. Quantitative calculations of the entropy and enthalpy have been performed in theoretical treatments of the buckling process of intact and defect harboring DNA (Brahmachari et al.[26]). The process of

quantifying the entropic contributions to the free energy of bending via simulations is not immediately obvious.

2.3.3 Simulation results of plectoneme pinning qualitatively agree with statistical mechanical model

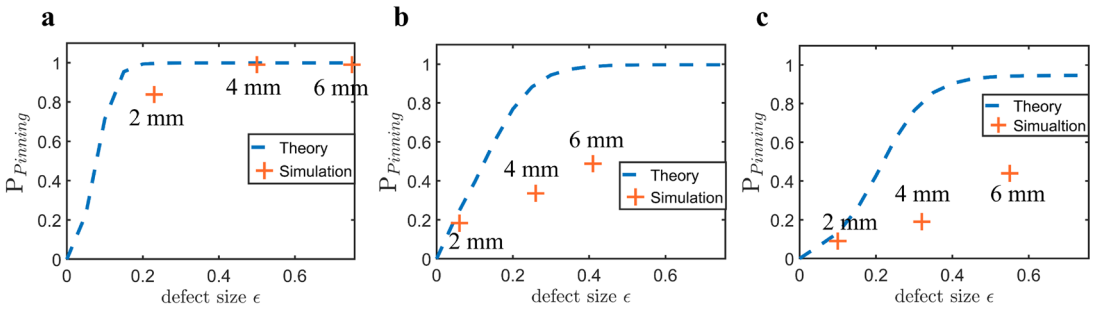

Figure 2.6 Theoretical predictions of probability of plectoneme pining at the defect site. Probability of the plectoneme being pinned at the mismatch (blue dashed line) are plotted as a function of the defect size, using theoretical model of Brahmachari et al.[26] The orange crosses correspond to the probability of buckling at the mismatch obtained from MD simulations for defect sizes associated with 2, 4, and 6 mismatches (see Table 2.2) . (A) For the high force-high salt case with a 610 bp DNA, both theory and simulations indicate that 2, 4, or 6 mismatches stably pin the plectoneme. (B) For the low force-low salt case with a 610 bp long DNA, the theory predicts appreciable pinning only for 4 and 6 mismatches, whereas the 2 bp mismatch is unable to pin the plectoneme. The overall trend of increasing pinning probability with increasing mismatch size is recovered from simulations (Fig. 2.5 (a-c)), but there is an increasing discrepancy between theory and simulations for 4 and 6 mismatches. (C) For the low force-low salt case with a 1010 bp long DNA, the theory predicts appreciable pinning only for the 6 bp mismatch. Similar to the results with the 610 bp DNA under the low force and low salt conditions, the MD simulations agree with the theory for a 2 bp mismatch, but the probability of pinning the plectoneme at the mismatch increase more slowly with increasing mismatch size than predicted by theory.

To verify the results of the MD simulations, we compared them with results obtained from a previously published theoretical model [26] that is built upon the statistical mechanical behavior of dsDNA as a semiflexible polymer. The model explicitly incorporates various free-energy components, such as the contribution from

DNA bending associated with the end loop and the plectoneme, the Debye-Hückel electrostatic energy associated with the plectoneme wrapping, and the stretching energy under an external force. The presence of a defect or base-pair mismatch is theoretically modeled as a mismatch-size dependent reduction in the energy component corresponding to DNA bending in the end loop. This energy reduction is incorporated through a defect size parameter, ε, which varies from 0, for no mismatches, to a maximum of 1 as the extent of the mismatch increases. The energy and size of the plectoneme end loop are rescaled by (1- ε). With this simplifying assumption, the energetic differences associated with a mismatch are lumped together in a single factor, which is likely an oversimplification of the underlying energetic considerations associated with mismatches.

MD simulations allow detailed microscopic understanding of the bending energy reduction associated with mismatches. Comparing the structure of the end loop and the energy associated with a kinked end loop, we obtain values corresponding to the defect-size (ε) parameter for different mismatch sizes. ε is defined as

$$\varepsilon = 1 - \frac{E_{\gamma\dagger}}{E_{\gamma}} \qquad (4)$$

Here, $E_{\gamma\dagger}$ is the energy of a kinked end loop and E_{γ} is the measure of energy of a defect free end loop. In order to compare the theoretical predictions with MD simulations, we calculate ε using the results of the MD simulations. We use Fig. 2.5 (d-f) to obtain the angle of the plectoneme end loop when the plectoneme is pinned at the mismatch and when it is not pinned at the mismatch. Fig. 2.4 (a) and Fig. 2.4 (b) can be used to obtain the energy required to bend a segment of DNA to a given angle. Using equation 4, we can obtain an approximate estimate of ε as a function of the number of mismatches in

the simulated DNA molecule. We provide the values of the structural parameter ε obtained from the simulations in table 2.2.

Table 2.2 Structural parameters obtained from simulation

	N = 610 bp, F = 2 pN, S = 1 M & $\Delta Lk = 3.46$	**N = 610 bp, F = 0.3 pN, S = 0.2 M & $\Delta Lk = 2.31$**	**N = 1010 bp, F = 0.3 pN, S = 0.2 M & $\Delta Lk = 2.87$**
Number of mismatches	**ε obtained from simulations**	**ε obtained from simulations**	**ε obtained from simulations**
0	0	0	0
2	0.23	0.06	0.1
4	0.5	0.26	0.32
6	0.75	0.41	0.55

Interestingly, the defect size parameter depends on the force and salt conditions. This suggests that the thermodynamic size of the defect, which is the parameter that controls the probability of buckling at the defect site, is not only dependent on the size of the mismatch but also depends on the force and the salt conditions. Using these values for the defect size, we find that the theoretical model predictions are in qualitative agreement with the simulation results. More specifically, for the high force and high salt conditions the simulations and the theoretical calculations quantitatively agree, whereas for the low force and low salt conditions (probabilistic regime) the theoretical probability of plectoneme pinning is over-estimated compared to the simulation results. The theoretical approach of Brahmachari

et al.[26] assumes the mismatches do not cause local torsional softening; however, we observe a slight torsional softening in the simulations. Although this torsional softening causes a small change in the critical linking number, it is unlikely to significantly affect the theoretical predictions. A second simplifying assumption in the theoretical approach is considering the end loop fixed in size for a given force and defect size. In simulations, however, we observe variations in the end loop size. These fluctuations in the end loop may additionally destabilize the pinned plectoneme at lower forces.

An important parameter in the theoretical model for quantitative comparison with simulation results is the discretization length used to calculate discrete sliding states of the plectoneme. This discretization affects counting of states that are accessible via sliding of the plectoneme. A small discretization length leads to a large number of states resulting in a higher entropic cost to pinning of the plectoneme, and vice versa. This discretization length becomes significant in determining the stability of the pinned plectoneme state, especially for short DNA chains, the simulation case at hand. This characteristic length is expected to be much shorter than the persistence length of DNA, about 50 nm (the coarse-graining length scale for semiflexible polymer behavior), longer than the base pair length scale (about 1 nm), and comparable to other characteristic lengths in the system, such as the force-induced correlation length (about 10-20 nm); however, the exact numerical choice is not clear. We find that using a discretization length of 5-10 nm results in the best agreement with the simulations. The data displayed in Fig. 2.6 were obtained using 5 nm as the characteristic length distinguishing the sliding states of the plectoneme.

2.3.4 Mismatches absorb twist but do not affect writhe

In addition to providing the location and bending angle of the tip of the plectoneme, the MD simulation approaches provide detailed information concerning every base pair in the DNA molecule. An important consideration that has not been addressed in prior experimental or theoretical studies is the effect of the mismatches on the torsional compliance of the DNA molecule. To address this, we calculated the average twist over each base-pair for both the intact and mismatch bases in each simulation (Fig. 2.7 (a-c)). The presence of mismatched bases is expected to cause a decrease in the local torsional stiffness. This effect can be seen clearly for the high force - high salt case (Fig. 2.7 (a)): the twist per bp step at the mismatch is higher than the twist per bp step at the intact DNA, and it increases with the size of the mismatch region, whereas the twist per bp step for the intact DNA remains constant. For the low force - low salt case, there does not appear to be a similar reduction in torsional stiffness (Fig. 2.7 (b,c)). Curiously, compared to the intact DNA, the twist per bp step at the mismatch decreases for the case of 2 mismatches for both 610 bp (Fig. 2.7 (b)) and 1010 bp (Fig. 2.7 (c)), but the twist per bp step increases for the 4 and 6-mismatch cases. We note that the twist at the 2 bp mismatch will be affected more by the flanking base pairs as compared to the 4 and 6-mismatch cases, which might be responsible for a lower twist at the mismatch.

To better understand the twist accumulation at the mismatch, we performed simulations of torsionally free and stretched DNA with the same sequence as the DNA under torsional stress. In torsionally relaxed DNA, we observed the same decrease in the twist at the mismatch step for the 2-mismatch case and subsequent increase in twist

for 4 and 6 mismatches. This would suggest that for the 2-mismatch case, the twist at the mismatch is lower even when no torsional stress is applied to the DNA. When torsional stress is applied, we observe a higher twist accumulation at the mismatch as compared to that at the intact base pairs due to torsional softening. The effect of torsional softening can be clearly seen for the high salt high force case (Fig. 2.7 (a)): here the twist at the mismatch is larger than that at the intact bases in the duplex. The twist at the mismatch for the low salt low force case is similar to the case with no torsional stress, likely due to the lower applied torsional stress. Previous, all atom molecular dynamics simulations have quantified the effect of a mismatch on the twist per bp step distribution in DNA[98,111]. These studies found that for the G:T mismatch used in the current study, the twist at the mismatch is similar to the twist at the intact DNA. To the best of our knowledge, there has been no study relating the number of consecutive mismatches to the twist per bp step distribution in the DNA.

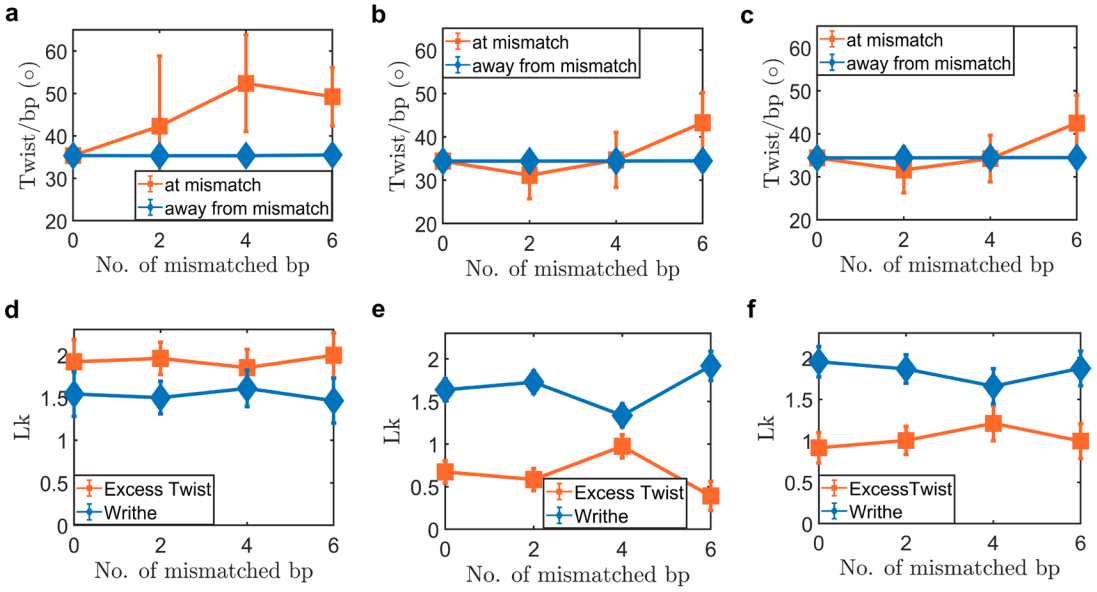

Figure 2.7 Twist and writhe in a positively supercoiled DNA containing mismatches. (a-c) Average twist per bp step in degrees (The details of the averaging procedure are provided in

the Methods section) of the DNA as a function of the number of consecutive mismatched base pairs introduced at the DNA center for (a) 610 bp DNA with F=2 pN and 1 M monovalent salt, (b) 610 bp DNA with F=0.3 pN and 0.2 M monovalent salt, and (c) 1010 bp DNA with F=0.3 pN and 0.2 M monovalent salt. Orange line represents the average twist per bp step at the mismatch, Blue line represents the twist per bp step averaged over all the intact base pairs. Error bars show standard deviation of measurement. **(d-f)** Distribution of excess twist and writhe as a function of the number of mismatches for (d) 610 bp DNA with F=2 pN and 1 M monovalent salt, (e) 610 bp DNA with F=0.3 pN and 0.2 M monovalent salt, and (f) 1010 bp DNA with F=0.3 pN and 0.2 M monovalent salt. The details of the averaging procedure can be found in the Methods section. Error bars show standard deviation of measurement.

We next probed possible global effects of the mismatches on the distribution of twist and writhe in the supercoiled DNA molecules. For all conditions and mismatch sizes, we find the twist and writhe stored in the DNA are not significantly altered by the presence of mismatches (Fig. 2.7 (d-f)). In the current work, the mismatches comprise a small fraction of the DNA (a maximum of 6 mismatches in a 610 bp DNA), hence the torsional softening at the mismatches does not significantly affect the twist, and hence the writhe, stored in the DNA. However, for the low force - low salt conditions (Fig. 2.7 (e, f)), we see a slight increase in the twist stored in the DNA for the 4-mismatch case. This increase is consistent for both the 610 bp and 1010 bp case and may be statistically significant. A future study of how the twist for the 4-mismatch case varies with increasing torsional stress will be helpful in determining the significance of this observation.

2.3.5 Mismatches allow DNA to accommodate the same writhe with a smaller plectoneme

Whereas the degree of bending or kinking at the tip of the plectoneme is an important determinant of the localization of the plectoneme that depends critically on

the size of the mismatch in addition to the salt and force conditions, it is not readily experimentally measurable. In contrast, the extent of the DNA molecule that is in the plectoneme can be accurately measured experimentally since it corresponds to the decrease in DNA extension associated with plectoneme formation. Fig. 2.8 provides the average size of the DNA plectoneme (length of DNA in the plectoneme) as a function of the number of consecutive bp mismatches introduced at the DNA center for different combinations of the applied force (F), salt concentration (S), and DNA length (N). The plectonemes are smaller under the high force - high salt condition as compared to the low force - low salt condition. This is also evident from the snapshots provided in Fig. 2.4 (c, d). The writhe in the DNA is not affected by the presence of mismatches (Fig. 2.7 d-f), yet we see a decrease in the plectoneme size with increasing number of mismatches. The presence of the mismatch causes a decrease in the end loop angle (Fig. 2.5 (d-f)) and a tighter plectoneme end loop reduces the plectoneme size while not affecting the writhe. Although we note that the decrease in plectoneme size is not universal and seems to disappear for the longer plectoneme.

We find that MD simulations can capture the variation in plectoneme size. In Fig. 2.8 we provide the standard deviation (SD) in measurements of the plectoneme size to show the variability of plectoneme size observed in the simulation. The SD of the plectoneme sizes is generally large. Here, we note that the large variations in the plectoneme size are not necessarily errors in the simulations. For example, under high forces (Fig. 2.8 (a)) we expect the plectoneme to be largely stable because any increase in the plectoneme size will result in a reduction in the DNA extension. A large amount of energy will be needed to pull against the high force and reduce DNA extension. We

see a SD of ~10 bps in the high force high salt case (see Fig. 2.8 (a)). Under lower force and salt concentrations (Fig. 2.8 (b)), we see larger variations in plectoneme sizes. Under lower forces, it will be easier for the plectoneme to elongate as it will have to perform less work against a lower force compared to the higher stretching force. Considering SD as a measure of error would imply that the errors in plectoneme sizes are lower in the high force- high salt condition compared to the low force – low salt conditions. We would generally expect a coarse-grained model like the OxDNA2 model to be more error prone under extreme conditions like the high force – high salt condition. Here, we provide the SD in plectoneme size to show the variation in plectoneme sizes and not as a measure of error.

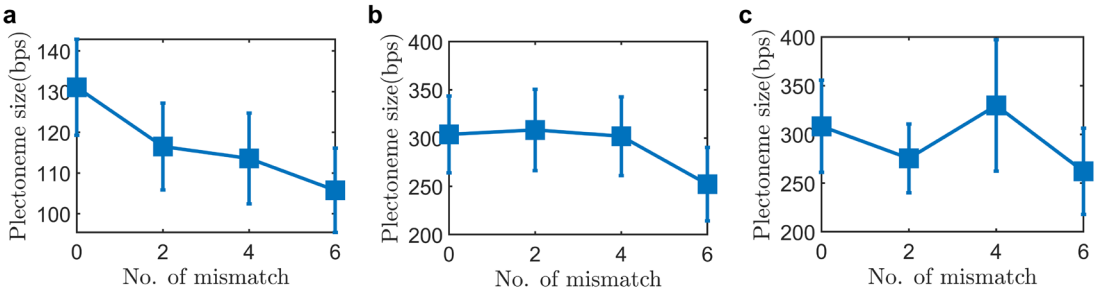

Figure 2.8 Influence of mismatches on the length of DNA in plectoneme. Average size (in bp) of the length of DNA in the plectoneme (see Methods section for details of the averaging procedure) as a function of the number of consecutive mismatched base pairs introduced at the DNA center for (a) 610 bp DNA with F=2 pN and 1 M monovalent salt, (b) 610 bp DNA with F=0.3 pN and 0.2 M monovalent salt, and (c) 1010 bp DNA with F=0.3 pN and 0.2 M monovalent salt. Note the differences in the y-axis scales. Error bars show standard deviation of measurement.

2.3.6 Mismatches reduce plectoneme diffusivity

While the statistics of plectoneme formation and conformation permit comparison with experiments, the simulation results provide a wealth of additional

information concerning plectoneme dynamics (Fig. 2.9 (a-d)). Tracking the position of the plectoneme along the DNA molecule over time provides a dynamic view of the motion and the pinning of the plectoneme as a function of the size of the mismatch region under different force and salt concentration conditions.

For intact DNA, the plectoneme is significantly less mobile under the high force - high salt condition than the low force – low salt condition (Fig. 2.9 (a) and 2.9 (c)). For DNA with four mismatches the plectoneme is strongly pinned at the mismatch location under the high force - high salt conditions (Fig. 2.9 (b)). This is reflected in the unity probability of the plectoneme being centered at the mismatch location (Fig. 2.5 (a)). Conversely, for DNA containing four mismatches under low force - low salt conditions, the plectoneme remains highly mobile, indicating a lack of pinning of the plectoneme at the mismatch. This is reflected in the low, significantly less than unity, probability of the plectoneme being centered at the mismatch location under low salt and low force conditions (Fig. 2.5 (b, c)).

To compare the mobility of the plectoneme under different conditions, we calculated the diffusion constant of the plectoneme center (Fig. 2.9 (h-j)). For plectonemes in both the high force-high salt and the low force-low salt regime, there is a decrease in the plectoneme mobility with increasing number of mismatches (Fig. 2.9 (h-j)). The diffusivity values presented in the current work present the short-time motion of the plectoneme. In the high force high salt conditions despite the fact that the plectoneme is pinned at the mismatch we observe slight changes in the plectoneme center (Fig. 2.9 (b)) and these slight fluctuations lead to the non-zero diffusivity values. We also observe higher diffusivity of plectoneme for the 1010 bp DNA (Fig. 2.9 (g))

compared to 610 bp DNA (Fig. 2.9 (f)). The lower diffusivity of the plectoneme in 610 bp DNA is likely due to the plectoneme size being a higher fraction of the DNA length. For the case of 610 bp DNA, the plectoneme size was half of DNA length in bps (~ 300 bp, See Fig. 2.8 (b)). In the case of 1010 bp DNA, the plectoneme is roughly a third of DNA length in bps (~ 300 bp, See Fig. 2.8 (c)).

The diffusion constant for intact DNA is similar to the value obtained by Matek et al.[62] However, the diffusivity obtained in the current work is 3 orders of magnitude larger than the experimentally observed diffusion constant. In the experiments performed by van Loenhout et al.[112], the diffusion constant is calculated for plectonemes larger than 4 kb. In the current study, the average plectoneme size is considerably smaller (Fig. 2.8), a smaller plectoneme will be more mobile. Another factor that may affect the diffusivity is the use of a coarse-grained model in the current work. Coarse-graining decreases the degrees of freedom in the system and hence accelerates the dynamics. Similarly, the use of implicit solvent decreases the effective viscosity, which further accelerates the conformational dynamics. As a result, the plectoneme diffusivity data provided here represents the qualitative trend of the effect of mismatches on plectoneme diffusivity, but the diffusion constants are likely over-estimated. It is possible to decrease the diffusivity of molecules (viscosity of an implicit solvent can be varied by changing the random forces applied by Langevin thermostat), but this would result in inefficient sampling. The simulation time is limited to the order of micro-seconds, and a higher diffusivity coefficient allows the simulation to sample phenomenon observed in experiments.

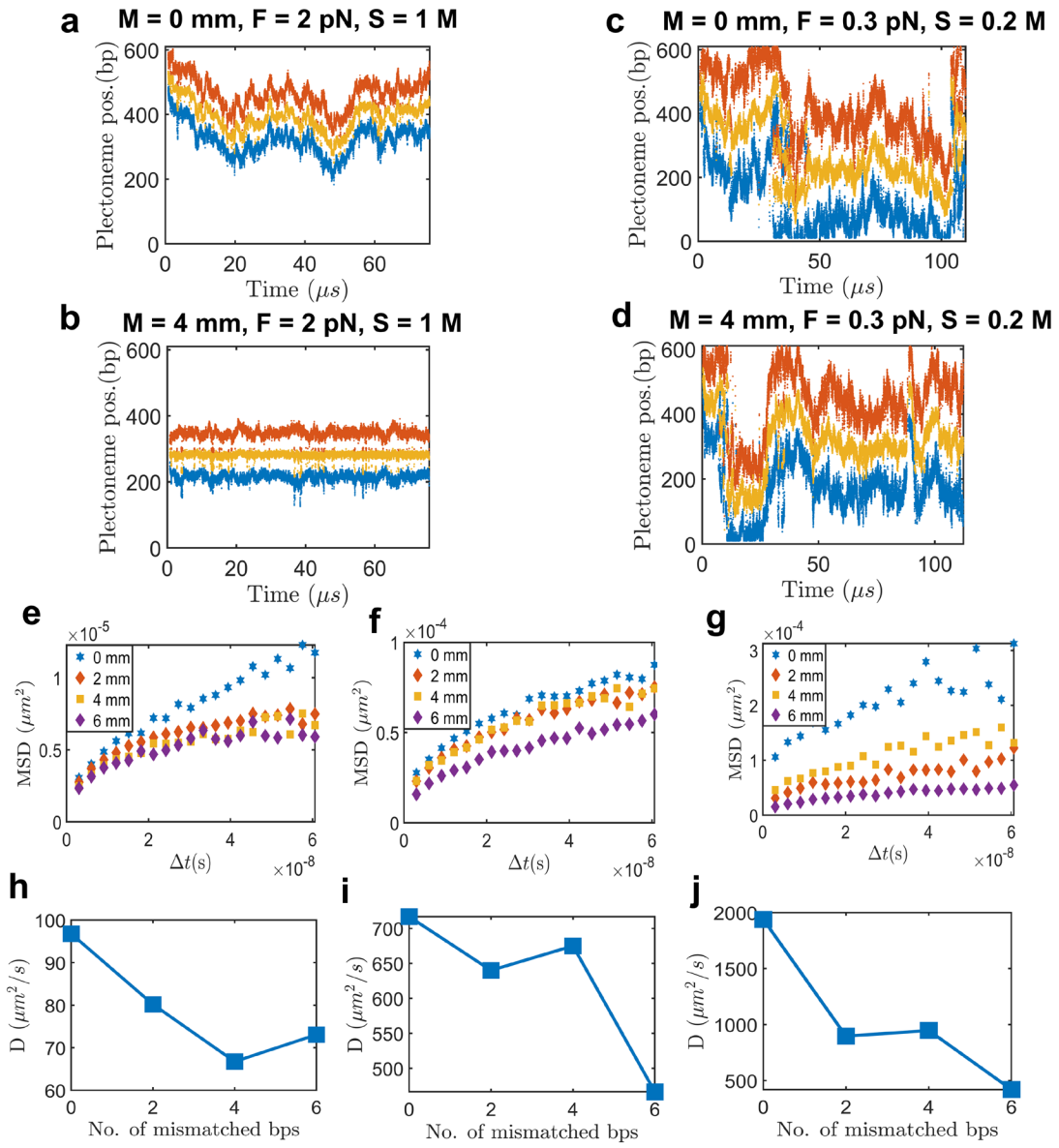

Figure 2.9 Mismatches slow plectoneme dynamics. **(a-d)** Time-dependent location of the beginning (blue line), center (yellow line), and end (red line) of the plectoneme for (a) high salt and high force with intact DNA (M=0), (b) high salt and high force with 4 consecutive mismatches (M=4) at the DNA center, (c) low salt and low force with intact DNA (M=0), and (d) low salt and low force with 4 consecutive mismatches (M=4) at the DNA center.**(e-g)** Mean squared displacement of the plectoneme center as a function of the number of consecutive mismatched base pairs introduced at the DNA center for (e) 610 bp DNA with F=2 pN and 1 M monovalent salt, (f) 610 bp DNA with F=0.3 pN and 0.2 M monovalent salt, and (g) 1010 bp DNA with F=0.3 pN and 0.2M monovalent salt. Note the change in the *y*-axis scaling. **(h-j)** Diffusivity of plectonemes as a function of the number of mismatches for (h) 610 bp DNA with F=2 pN and 1 M monovalent salt, (i) 610 bp DNA with F=0.3 pN and 0.2

M monovalent salt, and (j) 1010 bp DNA with F=0.3 pN and 0.2M monovalent salt. Note the change in the *y*-axis scaling.

2.4 Conclusions

To the best of our knowledge these results represent the first simulation-based evidence of the manner in which the interplay of force, salt concentration, and DNA size affects the localization of a supercoiled plectoneme in a DNA containing mismatches. We find that in the physiological regime of monovalent salt concentration and force, entropy plays an important role in preventing plectoneme pinning. As a result, under physiological conditions, small mismatch defects position plectonemic domains probabilistically; larger mismatches lead to gradually more deterministic positioning. These effects should be taken into account while considering the molecular enzyme kinetics of base-mismatch-repair proteins. More generally, we provide a coarse-grained simulation approach to investigate the effects of mismatches on DNA supercoiling and plectoneme formation that can provide details of the conformations, structures, and dynamics of the system that are not accessible from analytical theories or current experimental approaches.

Chapter 3: Coarse-Grained Simulations of Plectoneme Pinning in a Negatively Supercoiled DNA containing Mismatched Base-Pairs

3.1 Introduction

Base-pair mismatches in DNA can arise during the replication process. *In vivo,* these mismatched base-pair need to be rapidly repaired to maintain genomic integrity. The mismatched base-pairs are repaired by mismatch repair proteins (MMR)[113,114]. Generally, the MutS protein (a canonical example of a MMR protein) will first bind to the mismatched base-pair. Once bound, the MutS protein will then signal other MMR proteins to start the mismatch repair process[115]. The process by which MutS recognizes DNA is not well understood. The crystal structure of MutS bound to the mismatched DNA has shown that the DNA adopts a sharply bent configuration with the mismatched base flipped out[21–23]. In fact, a sharply bent DNA appears to be a common motif in many crystal structures of DNA bound to other classes of repair proteins[24].

Recently, Dittmore et al.[14] performed single molecule magnetic tweezers measurements to study supercoiling of a DNA molecule with different number of mismatches. Dittmore et al.[14] found that under high force (~2 pN and above) and high salt (~0.5 M and above) conditions, the plectoneme would localize at the mismatched base pair. This provides a possible mechanism by which MMR proteins could recognize the mismatch. The presence of mismatch locally reduces the bending rigidity of DNA and this reduction makes the plectoneme pinning at the mismatch favorable. However, plectoneme pinning results in a loss of entropy due to the loss of mobility of

the plectoneme. The probability of plectoneme pinning is governed by the trade-off between the enthalpic gain associated with bending at the mismatch and an entropic loss due to pinning. These experiments however were unable to probe the plectoneme pinning process under physiological condition (~ 0.5 pN force and ~ 0.2 M salt concentration). Brahmachari *et al.* [26] later published a theoretical study which predicted a reduction in the probability of plectoneme pinning under physiological condition. In chapter 2 of the current thesis, we presented molecular dynamics (MD) simulation results which qualitatively agreed with the theoretical prediction of Brahmachari *et al.* All of the studies discussed above were limited to studying positively supercoiled, i.e., over-wound, DNA molecules. The simulation framework discussed in chapter 2 can be used to study the plectoneme pinning in a *negatively supercoiled*, i.e., an under-wound, DNA. Negative supercoiling may cause local melting at the location of the mismatch. This local melting can reduce the local rigidity of the DNA molecule even further and enhance the plectoneme pinning. Local melting can also allow the DNA to absorb the torsional stress without forming a plectoneme by absorbing the change in linking number entirely through untwisting (melting) the DNA strands. Generally, supercoiling is favored at low forces and melting is favored at higher forces. Simulating DNA under different forces reveals the effect of mismatches on the competition between buckling and local melting in a negatively supercoiled DNA.

In the current chapter we use the framework, developed and verified in chapter 2, to study the plectoneme pinning process in a negatively supercoiled DNA under physiological salt conditions. We find that negative supercoiling significantly enhances plectoneme pinning in comparison to the case with equivalent levels of positive

supercoiling. We also find that the mismatched base pairs are locally melted for a significant portion of the simulation and the plectoneme end loop is bent significantly more as compared to the positive supercoiling case. Additionally, we simulate the 1010 base pair long DNA under two different negative super-helical densities, i.e., two different degrees of unwinding. We find that the super helical density does not affect the plectoneme pinning probabilities.

We also conduct simulations of DNA under different stretching forces (0.3 pN, 0.4 pN, and 0.6 pN). Negatively supercoiled DNA under relatively high stretching force (~0.6 pN) absorb torsional stress by locally melting instead of forming a plectoneme. Simulations of DNA under different forces allow us to study the effect of mismatches on the competition between plectoneme formation and local melting in a negatively supercoiled DNA. We find that higher stretching forces increase the probability of plectoneme pinning at the location of the mismatch.

3.2 Materials and Methods

To study the effect of the mismatch on a negatively supercoiled DNA, we simulate a 610 bp and a 1010 bp long DNA with 0, 2, 4 and 6 consecutive mismatches in the center of the DNA. The sequence of DNA is identical to the DNA in the previous chapter. The simulations are carried out under physiological conditions of 0.2 M salt concentration. The force and supercoiling conditions applied are supplied in the table below. Unless otherwise stated, the method used to analyze plectoneme properties and structure are similar to the methods described in the previous chapter.

Table 3.1 Simulation details

DNA length (bp)	Force applied (pN)	Implicit monovalent salt concentration (M)	Turns applied (ΔLk)
610	0.3	0.2	- 2.31
	0.4		
	0.6		
1010	0.3	0.2	- 2.87
	0.4		
	0.6		
1010	0.3	0.2	- 3.82
	0.4		
	0.6		

3.3 Results and Discussion

3.3.1 Negative supercoiling enhances plectoneme pinning at mismatches

Fig. 3.1 provides representative snapshots of a positively supercoiled and negatively supercoiled 610 bp long DNAs containing mismatches. The DNA in both snapshots is subjected to a stretching force of 0.3 pN and an implicit salt concentration of 0.2 M is used to simulate the physiologically relevant conditions. The simulation

conditions are identical for the two cases (i.e., the cases of positively and negatively supercoiled DNAs) with the only difference being in the direction of the torsional stress, which for the case of the negatively supercoiled DNA underwinds the DNA. Fig. 3.1 shows a positively supercoiled DNA containing mismatches in the center. In the snapshot, the plectoneme is pinned at the location of the mismatch. Fig. 3.1(A) shows a bent end loop and the inset shows the configuration of the mismatches.

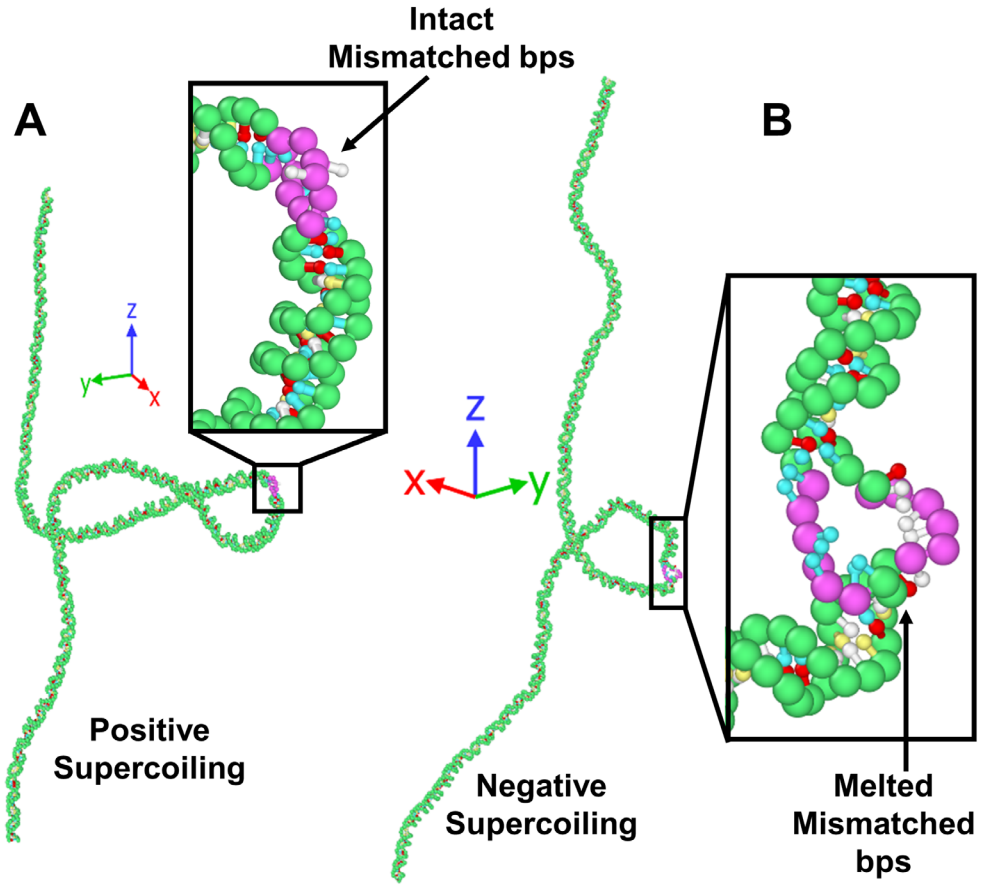

Figure 3.1 MD Simulation Snapshots. Simulation snapshots of **(A)** positively supercoiled (over-wound) and **(B)** negatively supercoiled (under-wound) 610 bp long DNA with 6 mismatched bps. The DNA is subjected to a constant force of 0.3 pN and an instantaneous torque at one end, while the other end is fixed. An implicit monovalent salt concentration of 0.2 M is used to perform the simulations. Green particles represent the sugar group of the matched bps and pink particles represent the sugar group of the mismatched bps. Red, yellow, white, and blue particles represent the A, C, G and T base groups, respectively.

We have previously shown that the presence of mismatches can facilitate plectoneme localization at the mismatch for a positively supercoiled DNA. Negative supercoiling of a DNA containing mismatches can result in local base pair melting at the mismatched bps [see the inset of Fig. 3.1(B)] present at the center of the DNA. This local melting in a negatively supercoiled DNA will result in an even greater reduction in the local bending rigidity, when compared to the positively supercoiled DNA, and therefore, will result in a further enhanced pinning. This reduction can be observed by comparing the average plectoneme end loop angle for a positively (Fig. 2.5 (e, f)) and negatively supercoiled DNA (Fig. 3.2 (C, D)), under similar force and salt conditions. In Fig. 3.2 (A, B) we compare the probability of plectoneme pinning at the mismatch for a positively and negatively supercoiled DNA under identical force and salt conditions. In Fig 3.2 (A, B), the 0 mismatch case represents the probability of plectoneme center being in a random location along the DNA. We find that negative supercoiling significantly enhances plectoneme localization at the mismatch. For the case of 1010 bp long DNA, we simulate DNA under two different specific supercoiling densities (which determine the magnitude of the applied torsional stress). We find that plectoneme localization is not significantly affected by the torsional stress applied (Fig. 3.2 (B)).

We have previously shown that plectoneme localization at the mismatch is facilitated by the reduction in local bending rigidity at the location of the mismatch. In Fig. 3.2 (C-E) we show the effect of mismatches on the plectoneme end loop angle. We find that, compared to positive supercoiling [Fig. 2.3 (D-F)], negative supercoiling (Fig. 3.2 (C-E)) causes a significantly higher bent plectoneme end loop (i.e., the

plectoneme end is significantly more kinked). The end loop angle observed here may depend on the specific parameters of the OxDNA2 model. To the best of our knowledge, the effect of multiple mismatches on plectoneme end loop angle has not been quantified in experiment or simulations and hence, it is difficult to make a quantitative comparison between the simulation data presented here and existing studies.

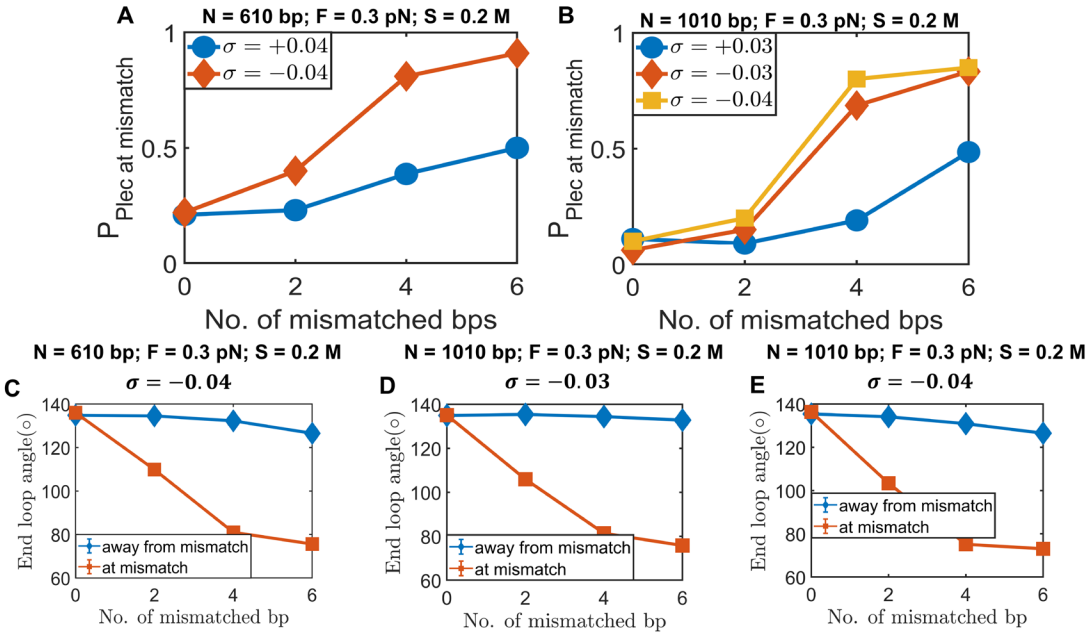

Figure 3.2 Plectoneme position and end loop angle. (A, B) The probability of plectoneme pinning at the mismatch as a function of the number of mismatches for negatively and positively supercoiled DNA. Here σ is the specific super helical density, positive supercoiling is defined by $+\sigma$ and negative supercoiling is defined by $-\sigma$. (A) 610 bp DNA with F=0.3 pN, 0.2 M monovalent salt, (B) 1010 bp DNA with F=0.3 pN, 0.2 M monovalent salt. Error bars (standard error of the mean, SEM) are smaller than the symbol size. **(C-E)** Average end-loop angle, in degrees, of the DNA plectoneme [this angle is defined by blue arrows in the insets of Figs. 2.2 c and Figs. 2.2 d] as a function of mismatch size (in base pairs) for (C) 610 bp DNA with F=0.3 pN, 0.2 M monovalent salt and σ = -0.04, (D) 1010 bp DNA with F=0.3 pN, 0.2 M monovalent salt and σ = -0.03 and (E) 1010 bp DNA with F=0.3 pN, 0.2 M monovalent salt and σ = -0.04. The orange line represents the average end loop angle when the plectoneme center is at the mismatch, whereas the blue line represents the average end loop angle when the plectoneme center is not pinned at the mismatch. Error bars (SEM) are smaller than the symbol size.

In Fig. 3.3, we compare the probability of local base pair melting at the location of the mismatch. In order to study the effect of mismatches on local base pair melting, we calculate the probability of a base pair (that is close to the location of the mismatch) being flipped out. Here we consider a bp to be locally melted if one or both nucleotides are flipped out. In Fig. 3.3, the center of a sequence of mismatch is considered base pair step 0, we calculate the base pair melting of the 10 base pairs adjacent to the mismatch. We find that the presence of mismatch causes an increase in the probability of local melting for both positively and negatively supercoiled DNA. Negative supercoiling [Fig. 3.3 (B, D)] enhances the local base pair melting at the mismatches. We also find that negative supercoiling causes melting of base pairs around the mismatches resulting in a bigger locally melted region when compared to positively supercoiled DNA.

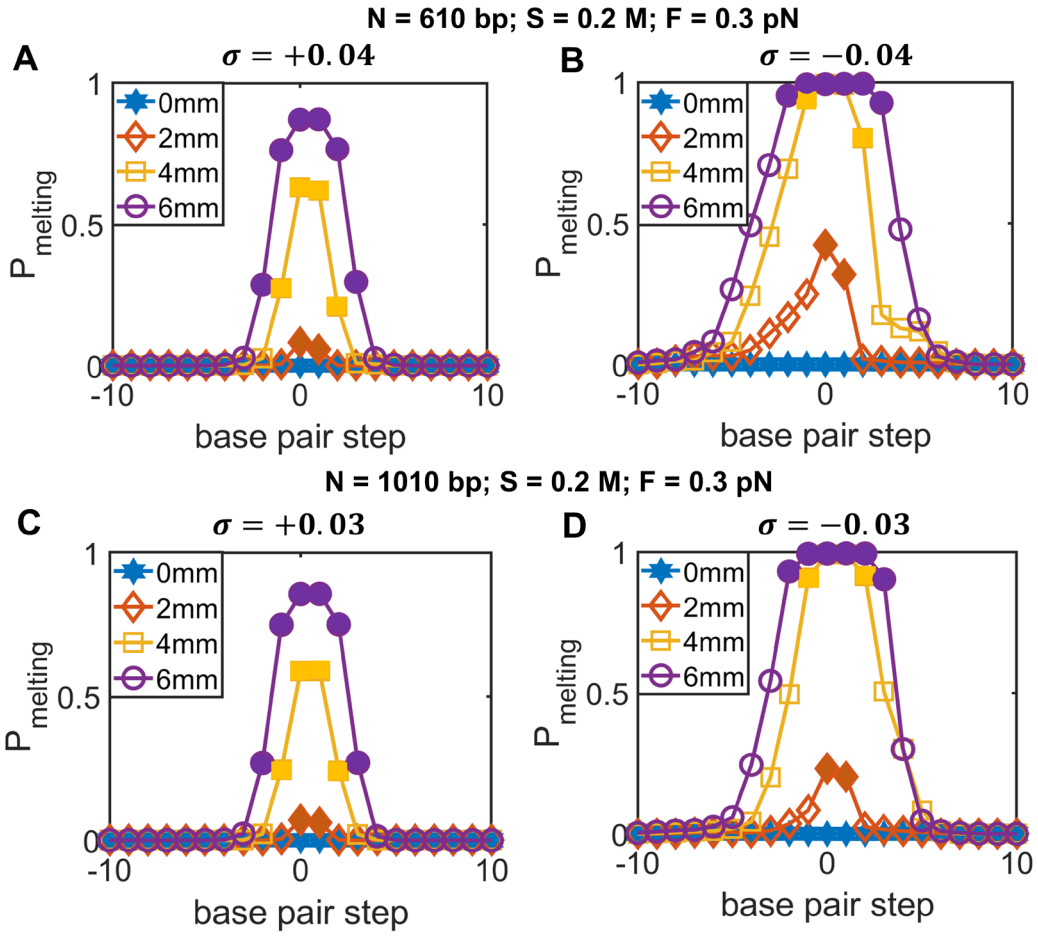

Figure 3.3 Probability of local melting. (A, B) The probability of base flipping as a function of bp location for 610 bp DNA with F=0.3 pN, 0.2 M salt and (**A**) σ = + 0.04, **(B)** σ = - 0.04. **(C, D)** The probability of base flipping as a function of bp location for 1010 bp DNA with F=0.3 pN, 0.2 M salt and (**C**) σ = + 0.03, **(D)** σ = - 0.03. The location of the mismatched bps is shown by filled symbols. Intact bps are shown as empty symbols. The center of consecutive mismatches is step 0 and the bp next to the center is shown as ±1 and so on. Error bars (SEM) are smaller than the symbol size.

3.3.2. Effect of stretching forces on plectoneme localization at the mismatch in negatively supercoiled DNA

In the previous section we have shown that negative supercoiling causes local base pair melting at the mismatch and hence, causes higher plectoneme localization at

the mismatches. The local base pair melting at the mismatches can be affected by the tension applied to the negatively supercoiled DNA. Single molecule measurements have shown that, even in an intact DNA molecule, negative supercoiling can start causing some local base pair melting at a tension of ~0.4 pN. Increasing this tension will result in additional melting and at around 1 pN of tension, the DNA will no longer form a plectoneme. At stretching forces higher than ~1 pN, the DNA absorbs the applied negative torsional stress by melting base pairs and no plectoneme formation is observed. We use the simulation framework developed in the previous chapter to study the effect of mismatches on the competition between plectoneme formation and local base pair melting in a negatively supercoiled DNA. We perform simulations of a negatively supercoiled DNA under three different tensions of 0.3 pN, 0.4 pN and 0.6 pN. We find that for all the stretching forces simulated here, the probability of plectoneme localization at the mismatch increases with the number of mismatches present in the DNA. We also find that the effect of stretching force on plectoneme localization is significant only for the case of 2 consecutive mismatches present at the center of DNA molecule (Figs. 3.4 (A-C)). For the case of 4 and 6 mismatches, we find that the mismatches are always locally melted even for the lowest stretching force (Fig. 3.5). Hence, we do not observe an appreciable increase in the probability of plectoneme localization with increased stretching force, for a DNA containing 4 and 6 mismatches.

In Figs. 3.4 (D-F) we show the end loop angle when the plectoneme is localized at the mismatches [similar to the orange line in Fig 3.2 (C-E)]. We observe that for all the simulation conditions, the plectoneme end loop angle decreases with an increase in the number of consecutive mismatches. We also find that the tension applied doesn't

have a significant effect on the plectoneme end loop angle for the case of 4 and 6 consecutive mismatches. For the cases of 4 and 6 consecutive mismatches, the mismatches are locally melted (Fig. 3.3) for all the applied tensions and hence the applied tension has negligible effect on the plectoneme end loop kinking. For the case of 2 consecutive mismatches, we find that the plectoneme end loop kinking increases (i.e., the end-loop angle decreases) with higher tension.

In Fig 3.5, we show the effect of force on the probability of base flipping (i.e., local melting). We find that for all simulated conditions, a DNA with 4 and 6 consecutive mismatches is always locally melted (i.e., the bases are flipped out) at the location of the mismatch. However, higher tensions applied to the DNA increase the size of the locally melted region by increasing the probability of base flipping in base pairs close to the mismatches. For the case of 2 consecutive mismatches, we find that the increased tension leads to an increase in the probability of base flipping of the mismatched base pairs. We also find that an increase in the applied tension leads to a bigger locally melted region around the mismatched base pairs. In Fig. 3.5, we observe an asymmetry in the probability of local melting around the mismatches. This asymmetry is likely due to the presence of A:T type bps, which are generally torsionally weaker than G:C type bps.

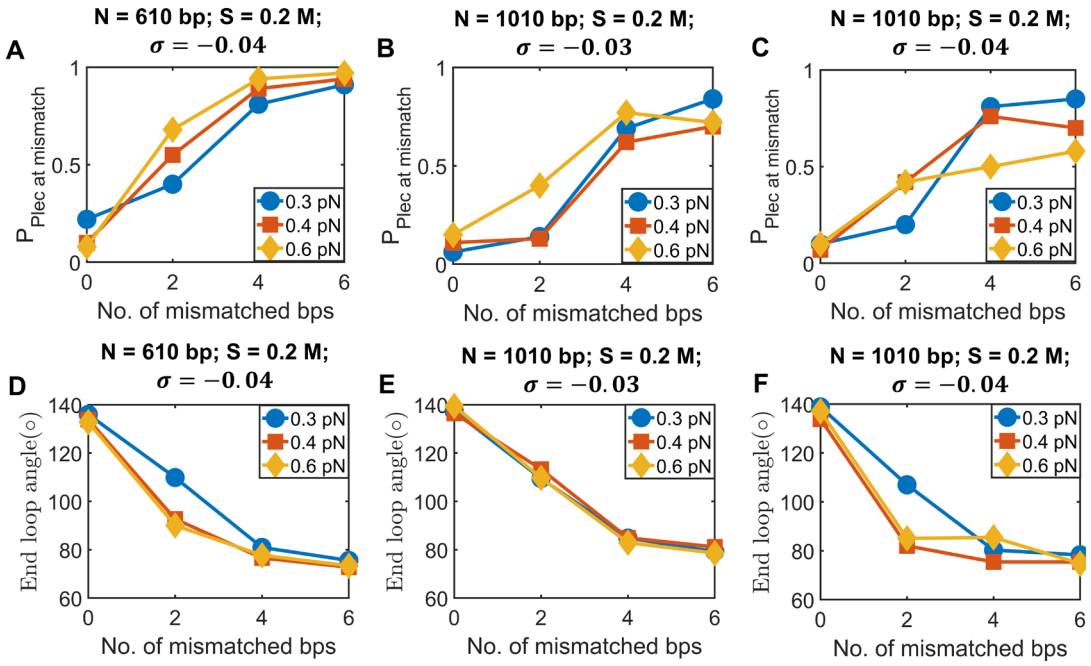

Figure 3.4 Effect of tension on plectoneme position and end loop angle. (A-C) The probability of plectoneme pinning at the mismatch as a function of the number of mismatches for negatively supercoiled DNA under different tension (A) 610 bp DNA, 0.2 M monovalent salt and $\sigma = -0.04$, (B) 1010 bp DNA, 0.2 M monovalent salt and $\sigma = -0.03$ (C) 1010 bp DNA, 0.2 M monovalent salt and $\sigma = -0.04$. **(D-F)** Average end-loop angle, in degrees, of the DNA plectoneme when the plectoneme is pinned at the mismatch as a function of mismatch size (in base pairs) for different tensions (D) 610 bp DNA, 0.2 M monovalent salt and $\sigma = -0.04$, (E) 1010 bp DNA, 0.2 M monovalent salt and $\sigma = -0.03$ (F) 1010 bp DNA, 0.2 M monovalent salt and $\sigma = -0.04$. Error bars (SEM) are smaller than the symbol size.

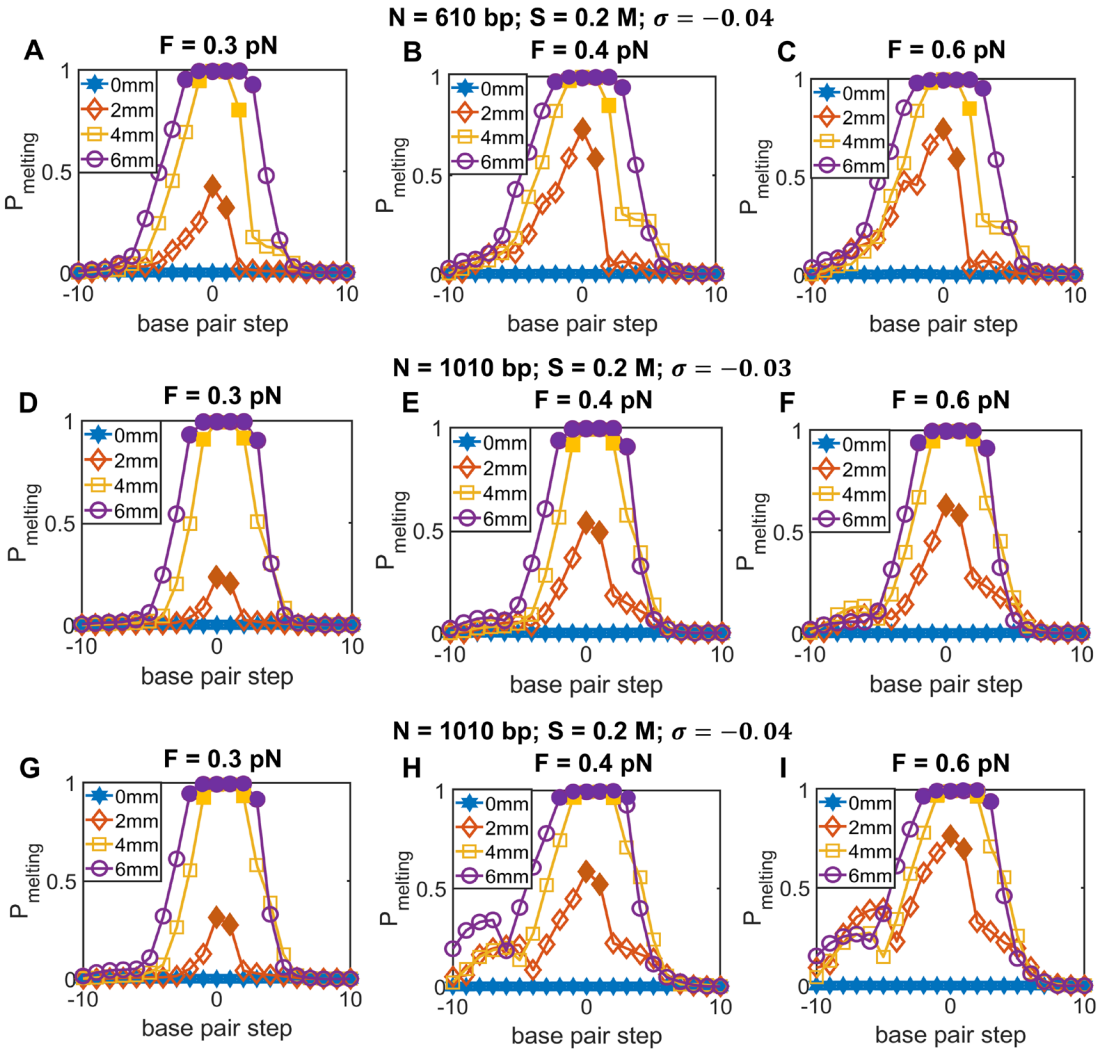

Figure 3.5 Effect of tension on base flipping. The probability of base flipping as a function of bp location for **(A-C)** 610 bp DNA in 0.2 M salt, $\sigma = -0.04$ and (A) F = 0.3 pN, (B) F = 0.4 pN, (D) F = 0.6 pN. **(D-F)** The probability of base flipping as a function of bp location for 1010 bp DNA in 0.2 M salt and $\sigma = -0.03$, and (D) F = 0.3 pN, (E) F = 0.4 pN, (F) F = 0.6 pN. **(G-I)** The probability of base flipping as a function of bp location for 1010 bp DNA in 0.2 M salt and $\sigma = -0.04$, and and (G) F = 0.3 pN, (H) F = 0.4 pN, (I) F = 0.6 pN. The location of the mismatched bps is shown by filled symbols. Intact bps are shown as empty symbols. The center of consecutive mismatches is step 0 and the bp next to the center is shown as ±1 and so on. Error bars (SEM) are smaller than the symbol size.

3.3.3. Molecular dynamics simulations capture fluctuations in plectoneme end loop caused by base flipping

MD simulations allow us to study the plectoneme localization and the effects of localization in granular detail. One example of the significance of garnering such details is as follows: some DNA binding proteins efficiently bind to a DNA that is bent at a specific angle and hence, the end loop angle of the plectoneme (precisely quantifiable from simulations) could affect the binding of mismatch repair proteins to the mismatch. Some proteins can actively bend DNA to a desired bend angle[116,117]. Crystal structure of mismatch repair protein (MutS) bound to a DNA containing G:T mismatched base pair[22], shows the DNA is bent at an angle of ~120°. Previous all atom simulations have shown that ~15 k_BT of free energy is required to bend a DNA containing G:T mismatch (Fig. 2.1) to an angle of ~120°. In addition to the free energy needed to bend the DNA, crystal structures of MutS bound to mismatch containing DNA show that mismatched base is flipped out. In absence of MutS, previous all atom studies have shown that ~ 10 k_BT free energy is required to flip out a mismatched G:T bp. When MutS is bound to the mismatch containing DNA, ~5 k_BT of free energy is required to flip out the mismatched nucleotide[118]. Here, it is important to note that the protein MutS locates mismatches in DNA without the consumption of ATP; MutS consumes ATP once it has located the mismatch[119] and starts the downstream mismatch correction process. Actively bending DNA and flipping out the mismatched base would require significant work by MutS, without consuming ATP. Negative supercoiling could potentially aid mismatch recognition by bending the DNA at mismatch and flipping out the mismatched base.

Fig. 3.6 shows the probability distribution of the plectoneme end loop angle for a DNA containing 2,4 and 6 mismatches at the center of the DNA. We find that for a DNA containing 4 and 6 mismatches, the end loop angle distribution has a single peak and a broad distribution. Such a single peak at similar values correspond to the fact that the bases are entirely flipped out for the cases of 4 and 6 mismatches. Interestingly, for the case of DNA containing 2 mismatches, we find that generally, the end loop angle distribution has two peaks. For the case of a DNA containing 2 mismatches, the mismatched base pairs have only a finite probability of being flipped out (Fig. 3.5). This is commensurate with these two peaks in the end loop angle distribution: one peak corresponds to the case of completely flipped out bases (hence the peak occurs at similar angle as the single peaks for the cases of 4 and 6 mismatches), while the second peak corresponds to the case of the intact (non-flipped out) bases.

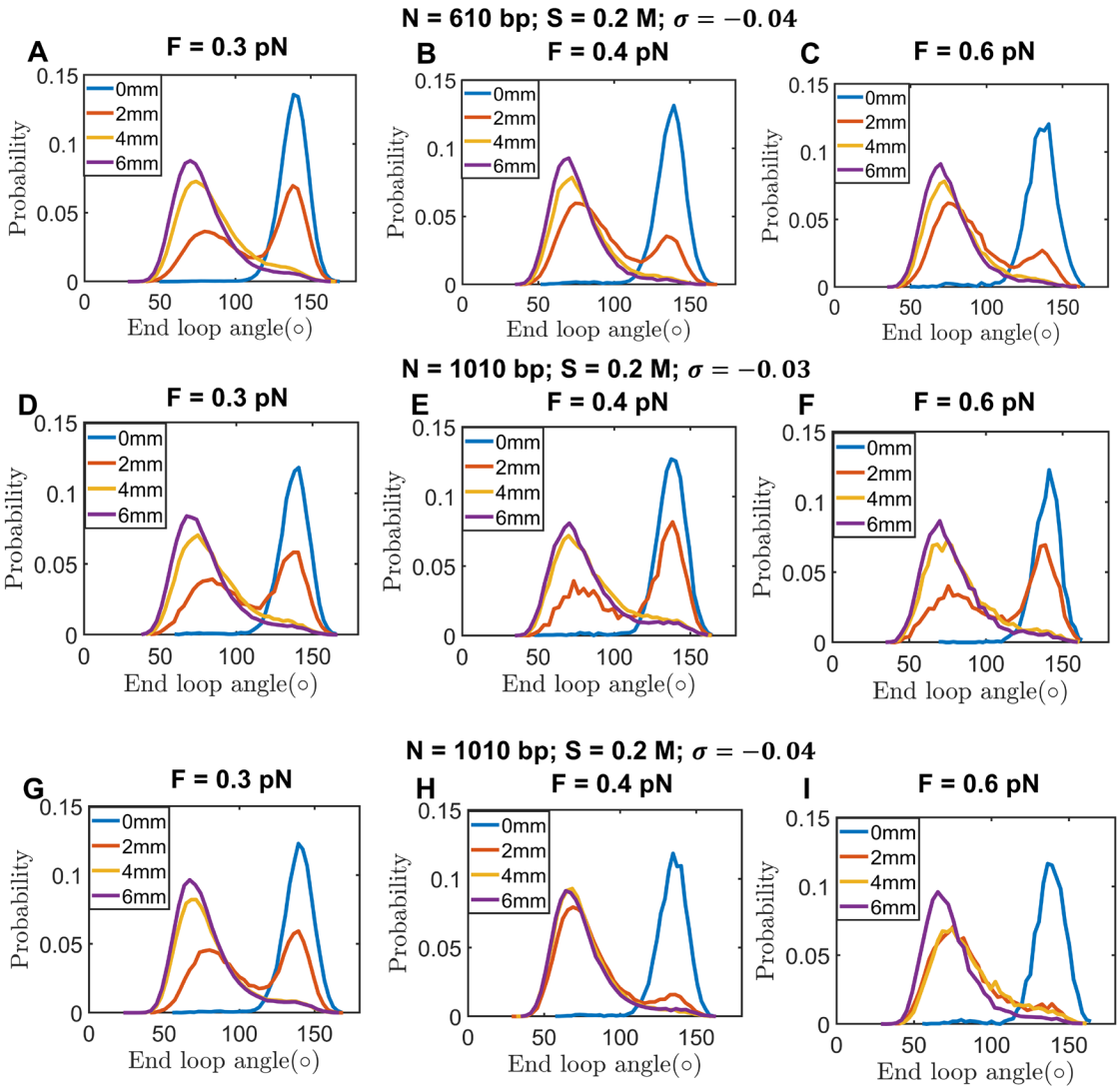

Figure 3.6 Distribution of plectoneme end loop angle. Probability distribution of the plectoneme end loop angle when the plectoneme is pinned to the mismatch for **(A-C)** 610 bp DNA in 0.2 M salt, $\sigma = -0.04$ and (A) F = 0.3 pN, (B) F = 0.4 pN, (C) F = 0.6 pN. **(D-F)** 1010 bp DNA in 0.2 M salt and $\sigma = -0.03$, and (D) F = 0.3 pN, (E) F = 0.4 pN, (F) F = 0.6 pN. **(G-I)** 1010 bp DNA in 0.2 M salt and $\sigma = -0.04$, and (G) F = 0.3 pN, (H) F = 0.4 pN, (I) F = 0.6 pN.

In Fig. 3.7, we plot the probability of the mismatched bp being flipped out as a function of the end loop angle. We find that for a DNA containing 4 and 6 mismatches, the mismatched base pairs are always flipped out for all end loop angle. This results in

a unimodal distribution (Fig. 3.6) of the end loop angle for a DNA containing 4 and 6 mismatches. For the case of a DNA containing 2 mismatched bps, we find that the probability of base flipping is not uniform. When the plectoneme end loop is strongly kinked (small end loop angle), the mismatched bps are flipped out. This corresponds to the first peak for the case of the 2 mismatched bp case. The presence of the base flipping is invariably associated with a smaller end-loop angle, stemming from the fact that the presence of the base flipping reduces the local bending rigidity of the DNA and hence make it easier to bend the DNA leading to a smaller end loop angle. By the same argument, a reduced presence of the base flipping should be associated with a larger end-loop angle. In other words, as evident from Fig. 3.7, for the case of 2 mismatched bps, a progressive increase in the end-loop angle leads to a progressive decrease in the value of $P_{base\ open}$. Most remarkably, this monotonic variation of $P_{base\ open}$ with the end loop angle ceases at around an end loop angle of $150\text{-}160^0$. $P_{base\ open}$ plateaus around that angle and then decreases sharply to zero as the end loop is increased further.

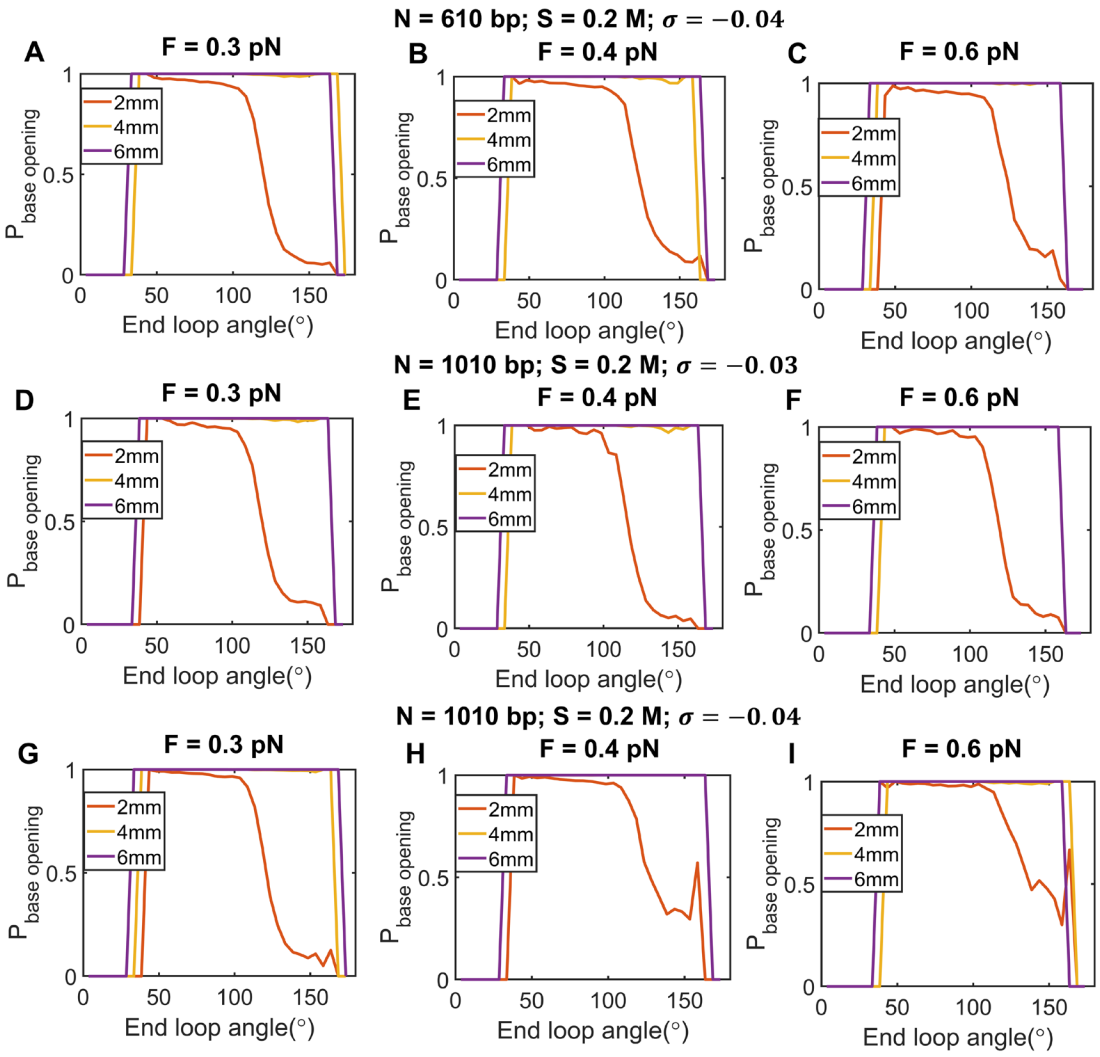

Figure 3.7 Effect of base flipping on plectoneme end loop. The probability of the mismatched base pair being flipped-out as a function of end loop angle for **(A-C)** 610 bp DNA in 0.2 M salt, $\sigma = -0.04$ and (A) F = 0.3 pN, (B) F = 0.4 pN, (C) F = 0.6 pN. **(D-F)** 1010 bp DNA in 0.2 M salt and $\sigma = -0.03$, and (D) F = 0.3 pN, (E) F = 0.4 pN, (F) F = 0.6 pN. **(G-I)** 1010 bp DNA in 0.2 M salt and $\sigma = -0.04$, and (G) F = 0.3 pN, (H) F = 0.4 pN, (I) F = 0.6 pN.

3.4 Conclusion

Plectoneme pinning could play a vital role in mismatch recognition. Single molecule experiments, MD simulations, and theoretical studies have recently shown that the presence of mismatches can cause plectoneme localization at the mismatch. These studies, however, were limited to a positively supercoiled DNA. Negative supercoiling can cause local melting at the location of the mismatch: this would result in a significant reduction in local bending rigidity at the location of the mismatch. This reduced bending rigidity can enhance the plectoneme localization at the mismatch. In the current chapter we have used the framework developed in previous chapters to study the effect of mismatches on plectoneme localization in a negatively supercoiled DNA. We find that compared to a positive supercoiling, negative supercoiling enhances the plectoneme pinning at the mismatch. We also find that plectoneme localization is not affected by the torsional stress applied to the DNA. We then study DNA under different applied tensions and find that higher applied tensions can cause an enhancement in plectoneme pinning for the case of 2 consecutive mismatches. We also find a direct correlation between the extent of end loop bending and base flipping in a mismatched base pair.

To the best of our knowledge, the simulation study presented here shows the first direct evidence of plectoneme pinning in a negatively supercoiled DNA containing mismatches.

Chapter 4: Observing Plectoneme Pinning in a DNA containing Mismatches Using Single Molecule Rotor Bead Magnetic Tweezer Assay

4.1 Introduction

In the previous chapters of this thesis, we have studied the effect of mismatched base pairs on DNA plectoneme using molecular dynamics simulations. The use of molecular dynamics was partially motivated by the limitation of traditional magnetic tweezer experiments. In this chapter we propose a rotor bead magnetic tweezer assay to overcome the limitations of traditional magnetic tweezer assays and verify the results provided in chapters 2 and 3.

In traditional magnetic tweezer assays, a magnetic bead is attached to one end of the DNA and the other end of DNA is attached to a glass slide[67]. Magnets can then be used apply tension and torque to the DNA through the tethered magnetic bead. Rotating the magnetic bead results in torsional stress being applied to the DNA. The application of torsional stress will result in a change in the twist of the DNA. Once a critical twist is attained, the DNA will buckle into a writhed structure, called a plectoneme. Additional turns applied to the bead will result in the elongation of the plectoneme. The formation of plectoneme results in a decrease of the DNA extension and the elongation of the plectoneme results in additional decrease of the DNA extension[11,12]. Magnetic tweezers have been used to study the plectoneme formation process[81]. Magnetic tweezer studies have also been used to study interactions of DNA with proteins[120–122].

A significant limitation of traditional single molecule magnetic tweezer assays is that the DNA extension provides limited information, so it is not possible to observe the location of the plectoneme. To overcome this limitation Dittmore *et al.* [14] introduced mismatched base pair approximately 8% of the DNA length from one end of the DNA. When supercoiled, the plectoneme that is pinned at the mismatch can absorb 8% of the DNA length from below the mismatch and the same amount from above the mismatch. Therefore, when the plectoneme is pinned at the mismatch and the plectonemes length has become 16% of the DNA length, the DNA can no longer elongate due to the plectoneme impinging on the surface. Additional turns applied to the DNA, under such circumstances, will result in a second buckling transition. A second buckling transition will only be observed if the plectoneme remains pinned at the mismatch. Therefore, the occurrence of this second buckling transition can be used to infer plectoneme pinning at the mismatch. Dittmore *et al.* [14] only observed pinning in the high force (> 2 pN) and high salt (> 0.5 M) regime. In the low force- low salt regime, it is possible that multiple plectonemes are already present [123,124] and even when a there is a single plectoneme it may not always be pinned at the mismatch [26]. This would lead to plectoneme pinning not being observed in the low force -low salt regime. Furthermore, since experiments performed by Dittmore *et al.* [14] required high forces to observe plectoneme pinning, they were not suitable to study plectoneme pinning in a negatively supercoiled DNA. Under high forces (> 1 pN) a negatively supercoiled DNA locally melts instead of forming a plectoneme.

In this thesis, we outline a strategy for extending magnetic tweezer assays to be able to measure the location of plectonemes. This strategy consists of attaching a non-

magnetic gold nanoparticle near the mismatched base pair via a di-thiol group. By simultaneously tracking the magnetic bead and gold nanoparticle, it is possible to determine whether the plectoneme is located at the mismatch or elsewhere along the DNA (Fig. 4.2). In order to track both the magnetic bead and gold nanoparticle, the magnetic tweezer instrument is combined with a total internal reflection microscope. A schematic of the instrument that will be used to simultaneously track the magnetic bead and the gold nano-bead is shown in Fig. 4.1. In Fig. 4.2, we provide an illustration to help the reader visualize the experiment.

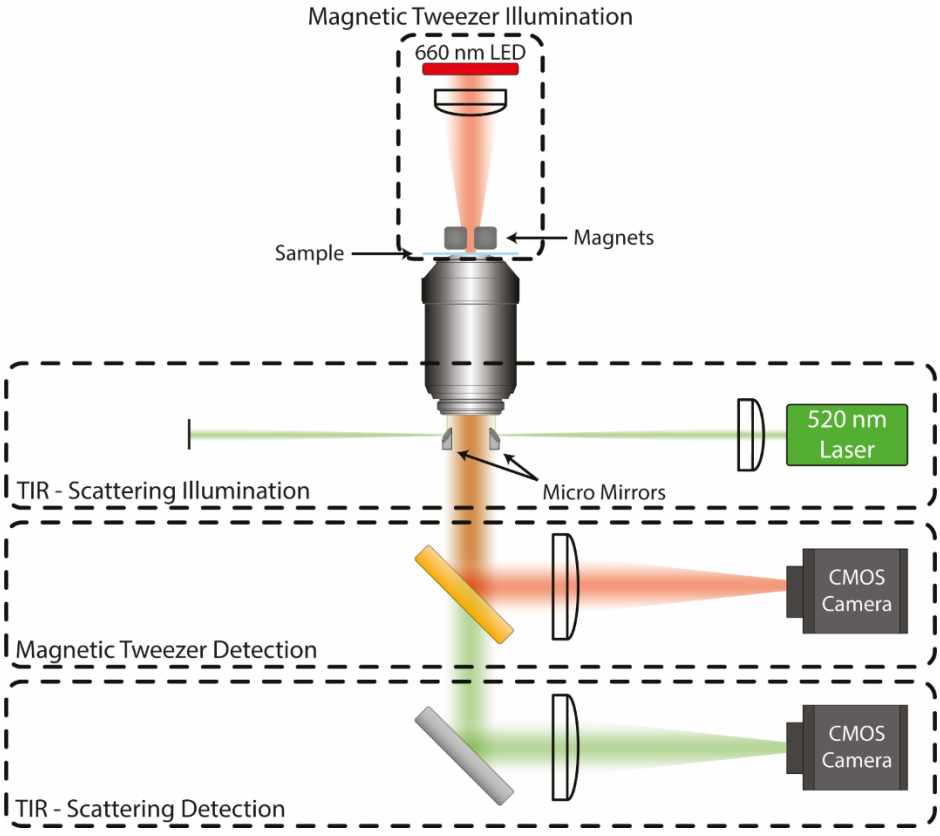

Figure 4.1. **Schematic of combined magnetic tweezer and total internal reflection microscope.** A magnetic bead tethered DNA substrate is added to the sample chamber. Permanent magnets located above the sample chamber can be used to apply force and rotate the magnetic bead. The magnetic bead is illuminated by a 660 nm LED. Motion of the magnetic

bead is captured by a CMOS camera. A 520 nm laser is used to create an evanescent field. The location of gold nano-bead can be captured using the evanescent field[73].

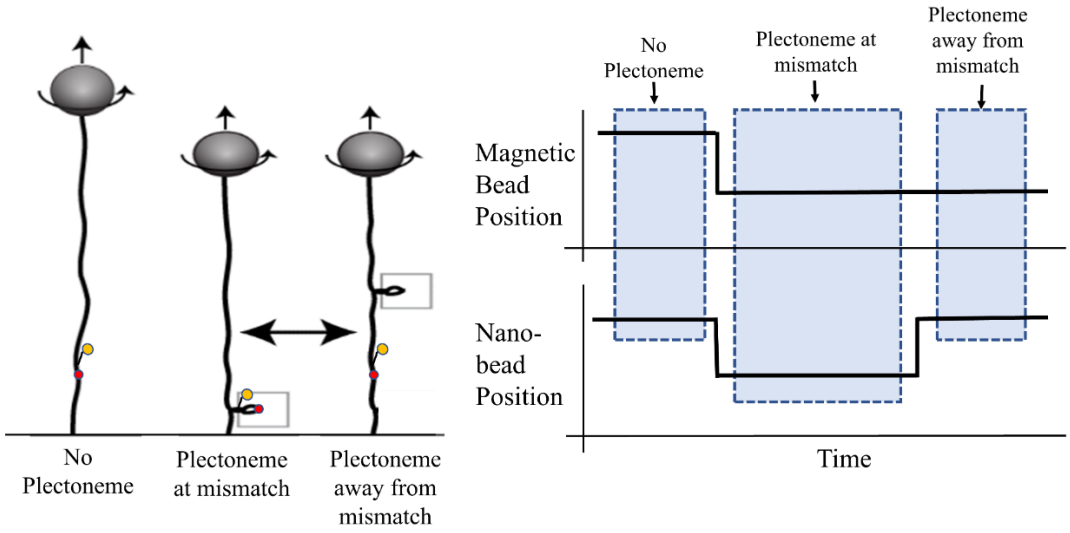

Figure 4.2 Illustration of the proposed experiment. The figure provides a schematic of the proposed substrate and the expected data to be collected. Mismatches (red dot) are introduced close to the surface. A gold nano-bead (gold dot) is attached next to the mismatches (red dot). In the illustration, turns are applied to the magnetic bead until initial plectoneme formation occurs. The formation of plectoneme results in a reduction in the vertical position of the magnetic bead. When the plectoneme is located below the gold nano-bead, we will observe a reduction in the vertical position of the gold nano-bead. If the plectoneme is located above the gold nano-bead, the vertical position of the gold nano-bead remains unchanged. Since the mismatches and gold nano-bead are located close to the surface, we can assume that when the plectoneme is located below the gold bead, the plectoneme center is localized at the mismatches. Simultaneous tracking of magnetic bead and gold nano-bead allows us to study plectoneme localization in conditions when plectoneme is not always pinned at the mismatch. Blue regions are shown to identify the magnetic and gold nano-bead vertical position when plectoneme formation has not occurred, when plectoneme is located at the mismatch, and when plectoneme is away from the mismatch.

In the experiment proposed here, we plan to simultaneously track the location of the magnetic bead (attached to one end of DNA) and the gold nano-bead (attached next to mismatch) (Fig. 4.2). DNA absorbs the applied tortional stress by change in twist. The change in twist does not significantly change the vertical position of the DNA tethered magnetic bead. Once a critical amount of torsional stress has been absorbed through twist, the DNA buckles and forms a plectoneme. The formation of

plectoneme leads to a reduction in the vertical position of the DNA tethered magnetic bead. The plectoneme, once formed, can translocate along the DNA backbone. The vertical position of the DNA tethered magnetic bead does not provide any information on the location of the plectoneme.

We introduce the mismatches close to where one end of the DNA is attached to the surface. The gold nano-bead is attached close to the mismatch. The location of the gold-nanoparticle can provide a direct signal if the plectoneme is present at the mismatch. If the plectoneme is located below the gold nano-bead, we expect the gold-nanoparticle to move close to the surface. Since the mismatches and gold nano-bead are located close to the surface, we can infer that when the plectoneme is located below the gold nano-bead, the plectoneme center is localized at the mismatch. The presence of plectoneme away from the mismatch (plectoneme is located above the gold nano-bead) will not result in a change in the vertical position of the gold nanoparticle. In short, the presence of plectoneme at the mismatch will result in a change in the vertical position of both the magnetic bead and the gold nano-bead. When the plectoneme is not located at the mismatch, we will observe a change in the vertical location of the magnetic bead, but the vertical position of the gold nanoparticle will be unchanged. This method will allow us to study plectoneme localization even if multiple plectonemes are present and the plectoneme doesn't stay pinned at the mismatch. Most importantly the method can be used to study plectoneme pinning in a negatively supercoiled DNA.

In traditional magnetic tweezers experiment, plectoneme pinning could be observed only if the plectoneme remained pinned at the mismatch. High forces and

high salt concentrations were needed to ensure plectoneme pinning at the mismatch. Under high forces, plectoneme formation does not occur in a negatively supercoiled DNA, instead the DNA locally melts. Hence, traditional magnetic tweezers experiments could not be used to observe the plectoneme pinning in a negatively supercoiled DNA. The experiment outlined here can be used to study plectoneme pinning under low forces (when the plectoneme does not always stay pinned at the mismatch) and hence can be used to study plectoneme pinning in a negatively supercoiled DNA.

In the remainder of this chapter, we describe the procedure to generate the DNA substrate used in a rotor bead assay. We use agarose gel electrophoresis to verify that the generated substrate is of the accurate length. We also verify that our protocol can be used to add mismatches to DNA. Finally, we also perform extensive single-molecule magnetic tweezers assay to verify that the presence of the dithiol group does not cause plectoneme pinning in an intact DNA.

4.2 Materials and Methods

4.2.1 Description of DNA substrate

In order to perform the single-molecule rotor bead assay we generate an approximately 4.5 kilo-base long DNA molecule. The substrate is designed to be a modular assembly which can be easily modified to induce mismatches or modify the DNA length.

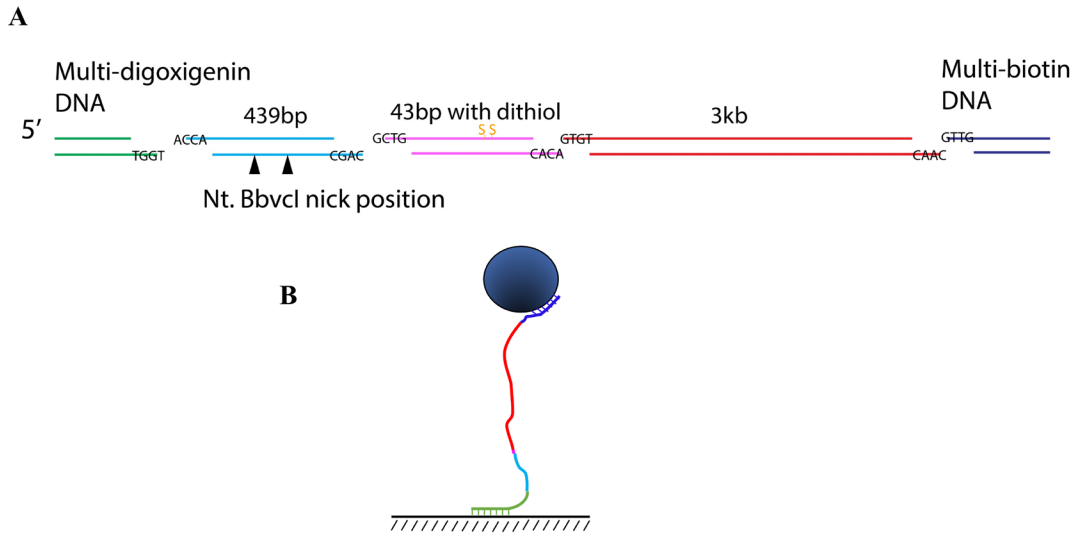

Figure 4.3 Schematic of DNA substrate. (A) Figure shows the schematic of the modular DNA substrate generated in this thesis. The muti-digoxigenin and multi-biotin handle attach to the flow cell surface and the magnetic bead, respectively. The 439bp DNA segment contains two Nt.BbvcI (New England Biolabs) nicking sites that can be used to add mismatches to the DNA substrate. The 43bp DNA contains the di-thiol group needed to attach the gold nanoparticle. Each segment is individually digested to produce unique 4 bp overhangs. The unique overhangs are used to guide ligation of segments. **(B)** Illustration of the final assembled substrate. The multi-biotin handles attached to a streptavidin coated magnetic bead (blue bead). Multi-digoxigenin handles attach to an anti-digoxigenin coated flow cell surface.

The muti-digoxigenin and multi-biotin handles are 500 bp long and attach to the flow cell surface and the magnetic bead, respectively. The 439 bp DNA segment contained two specific nicking sites. The two nicking sites are specifically nicked by the Nt.BbvcI enzyme (New England Biolabs). This nicking will produce a 37 nucleotide (nt) gap in one strand of the DNA. We can then add a single DNA strand that contains the mismatch and anneal it to the 439 bp DNA segment (see section 4.2.3). The 43 bp DNA segment contains a di-thiol group that can be used to attach the gold nanoparticle to the DNA. The protocol to attach the gold particle has been previously developed and validated by our collaborator, Dr. Haksung Jung (NIH, Bethesda). The

3 kilo-base pair (kbp) DNA is designed to be easily swappable with a DNA segment of different length. The simulations results presented previously have shown that DNA length can affect the plectoneme pinning at the mismatch.

4.2.2 DNA substrate preparation

The 3 kbp were generated using PCR amplification of the pET-28b plasmid (EMD4Biosciences). The 439 bp DNA was generated using PCR amplification the pkz2x plasmid. The pkz2x plasmid is similar to the pET-28b plasmid, except it contains an additional 48bp DNA segment at the BamHI site in the pET28b plasmid. This segment contains two Nt.BbvcI recognition sequences that are separated by 37 nt and also PstI restriction recognition sequence between these two nicking sequences. The nicking sites are used to introduce mismatches in DNA (see section 4.2.3). The multi-digoxigenin and multi-biotin handles were generated using PCR amplification of the pBluescript II KS(+) plasmid (Stratagene). Digoxigenin-11-dUTP (Roche) and Biotin-16-dUTP (Roche) were added to the PCR reaction mixtures for multi-digoxigenin and multi-biotin handles, respectively. The PCR product was purified using the QAquick PCR purification kit (QIAGEN). The purified product was then digested with BsaI-HF (New England Biolabs). The digestion produces specific 4 bp overhangs that can be used to ligate the individual segment of the DNA (the specific 4 bp overhangs of each segment are shown in Fig. 4.3 (A)).

The segments are ligated together using the T4 DNA ligase (Promega). The ligation of DNA substrate is carried out in multiple steps (Fig. 4.4 (A)). First, the 439

bp, multi-digoxigenin and 43 bp DNA (containing di-thiol group) segments are ligated together. In Fig. 4.4 (B), we verify that the expected ligation occurred using a 1% Agarose gel. Separately, we ligate the 3 kbp segment and multi-biotin segments. In Fig. 4.4 (C) we show the ligation product of this step in the third column. We observe lower ligation efficiency for this step. Finally, we ligate the two separate parts together. The ligation efficiency of the final step is low, and we extract the completed DNA substrate using gel extraction (QIAquick Gel Extraction Kit, Qiagen). Each segment of the DNA substrate is designed to have a unique 4 bp overhang (Fig. 4.3 (A)). The multi-step ligation process ensure that we only obtain a supercoilable DNA substrate if all the DNA segments have ligated as expected.

The protocols and materials used to perform the reactions are similar to the Seol et al.[125] and are also provided in appendix A. The gel electrophoresis data shown in the thesis is collected using precast 1% Agarose E-gel (Thermo Fisher) stained with SYBR gold. Invitrogen E-Gel Power Snap Electrophoresis device is used to perform and record gel electrophoresis.

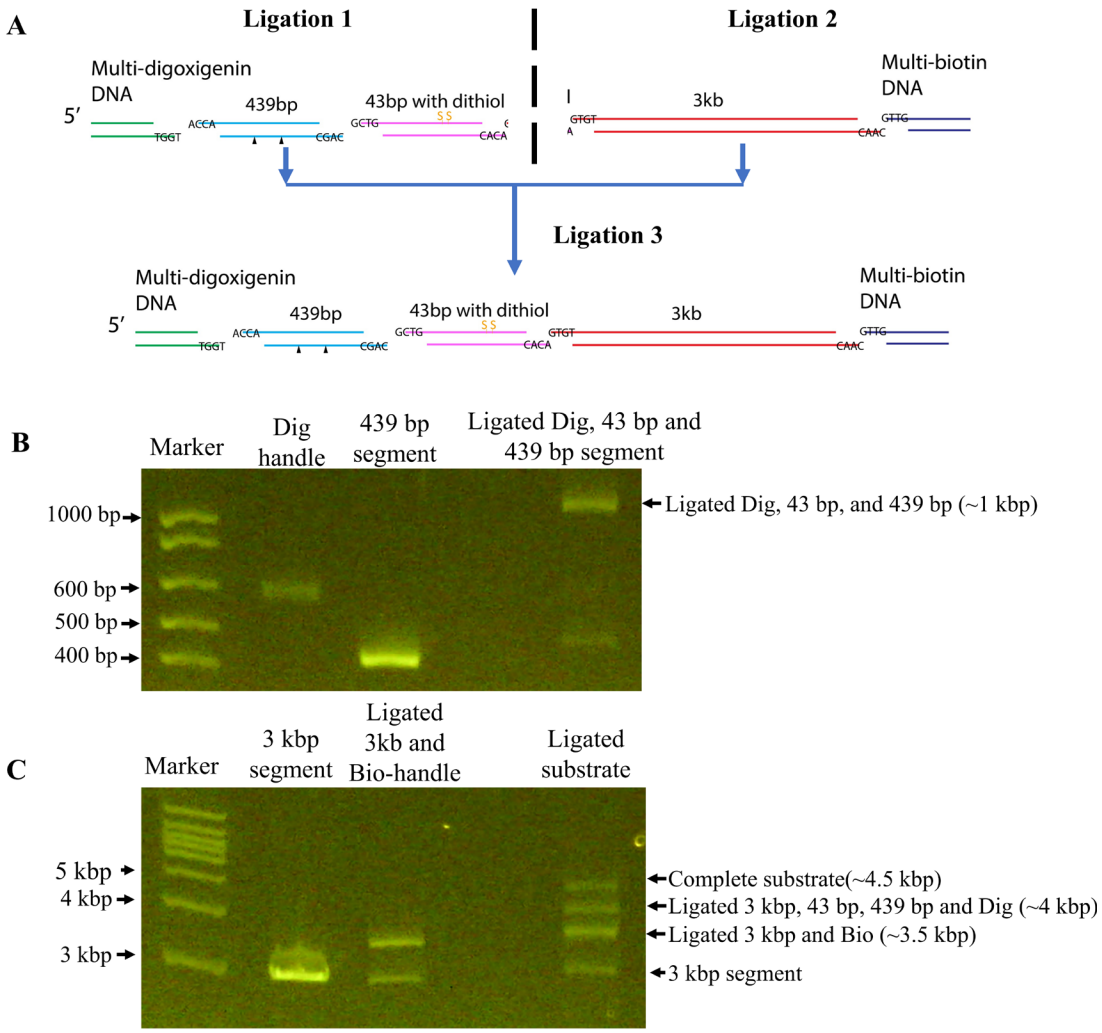

Figure 4.4 Ligation of DNA substrate. (A) Schematic showing the multi-step ligation process. The multistep ligation process ensures that we only obtain a supercoilable DNA substrate if all segments have correctly ligated. First, we ligate multi-digoxigenin handle, 439 bp segment and 43 bp segment together. Separately, we ligate multi-biotin handle with the 3 kbp segment. Finally, we ligate the two separate segments together. **(B, C)** 1% agarose gel showing the ligation product. (B) Ligation 1, The individual DNA segments to be ligated are shown along with the ligated product. Here, "Dig" refers to multi-digoxigenin handle. The expected length of ligated segment is shown in brackets. Samples are run for 10 min. **(C)** Ligation 2 (3 kbp and multi-biotin handle) is shown in Lane 3. Finally, ligation 3 is shown in Lane 4. We observe multiple ligation products for ligation 3. Since each DNA segments ends in a specific 4 bp overhang, we can account for all the different DNA ligation products. Gel extraction is used to obtain the correct ligated substrate. Here, "Bio" refers to the multi-biotin handle. Samples are run for 20 min.

4.2.3 Introduction of mismatched base pairs to DNA.

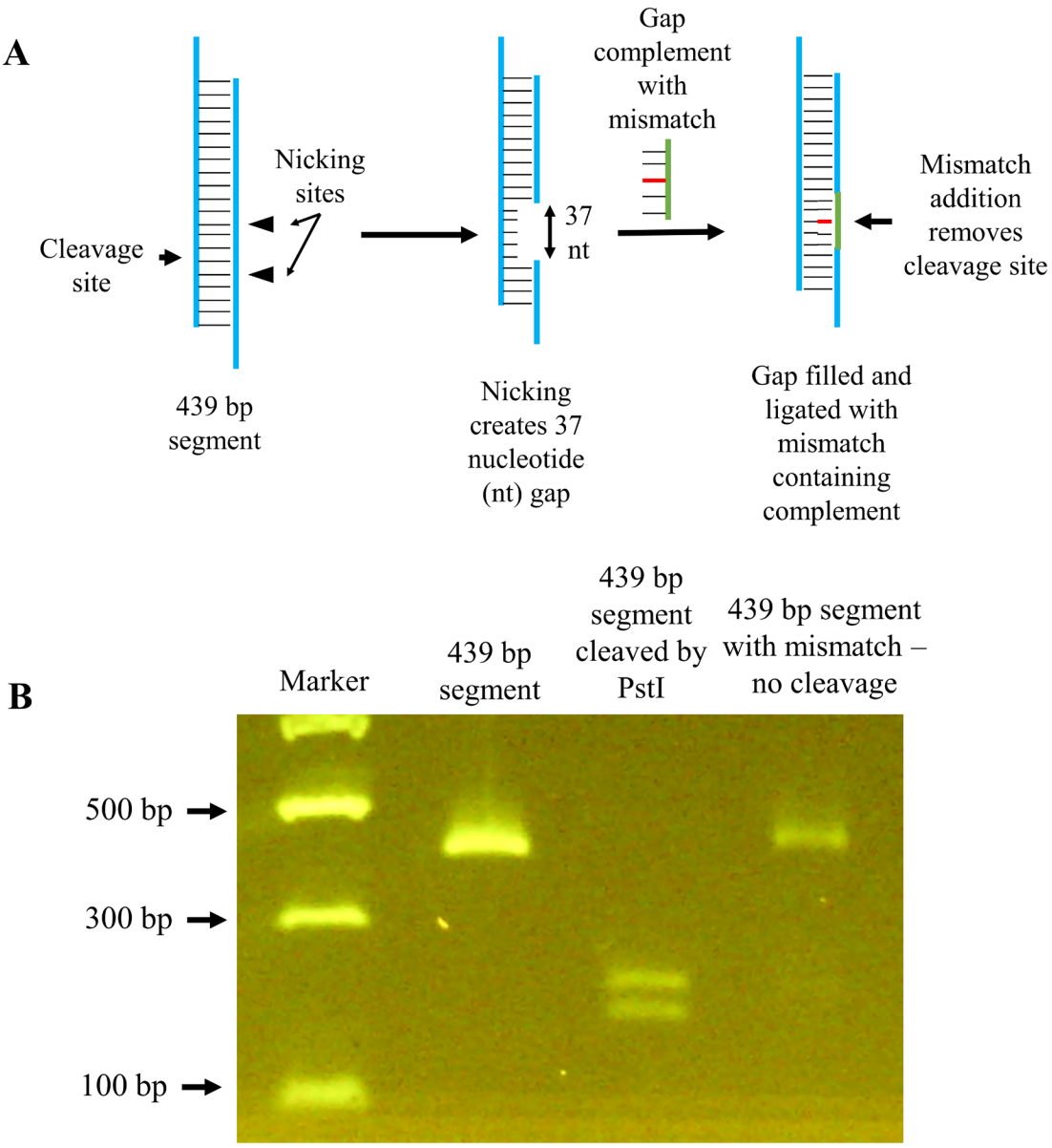

Figure 4.5 Addition of mismatch to DNA. (A) Schematic of the mismatch addition process. The mismatches are introduced in the 439 bp segment. The 439 bp segment contains two specific nicking sites. A specific sequence cleaved by PstI enzyme exists between the nicking sites. Nicking produces a 37 nucleotide (nt) gap in the substrate. A 37 bp single stranded DNA containing mismatches is added to the mixture. The 37 bp segment is otherwise complementary to the gapped substrate. Addition of mismatch removed the PstI cleavage site. Mismatch addition is confirmed if addition of PstI enzyme does not result in cleavage. **(B)** 1% Agarose gel showing cleavage reaction of 439 bp segment with PstI enzyme. Cleavage is observed with

the 439 bp segment with intact DNA sequence. Addition of mismatches removes the cleavage site. No cleavage is observed in a 439 bp segment containing mismatch.

The mismatched bps are introduced in the 439 bp DNA segment. Nt.BbvcI enzyme (New England Biolabs) is used to nick two specific sites in the 439 bp DNA segment (Fig. 4.5 (A)).

We first nick the 439 bp DNA segment using Nt.BbvcI enzyme. Nicking produces a 37 nucleotide (nt) gap in the DNA (Fig. 4.5 (A)). Next, we ligate a mismatch containing single stranded DNA to the nicked 439 bp DNA segment. In order to confirm the mismatch incorporation, the 439 bp DNA segment is designed to contain a specific sequence cleaved by the PstI (New England Biolabs) enzyme. We then cleave the mismatch containing 439 bp DNA with PstI enzyme. Addition of the mismatch removes the PstI cleavage site and the mismatch containing 439 bp DNA segment is no longer cleaved (Fig. 4.5 (B)). This procedure allows us to verify that the mismatch is inserted in the DNA molecule. The protocols used for nicking and cleavage are provided in Appendix A.

4.3 Results and Discussion

4.3.1 Calibration of conventional single-molecule magnetic tweezers

Here, we briefly describe the procedure used to calibrate a conventional single-molecule magnetic tweezers. In this thesis, we plan to observe supercoiling of a DNA molecule containing a di-thiol group. This measurement will allow us to ensure that the presence of the di-thiol group does not cause plectoneme pinning at the thiol group. In

order to accurately capture the force applied to the magnetic bead, we first calibrate the forces in a traditional magnetic tweezers instrument using a 11 kb long DNA.

The procedure to calibrate the magnetic tweezers instrument is similar to the procedure used previously by Dittmore et al.[14]. Briefly, we first attach approximately 500 bp long muti-digoxigenin and multi-biotin handles to the 11 kb DNA segment. We then assemble a flow cell using double sided parafilm sticky tape to glue two cover slips together. Next, we flow in polystyrene beads and incubate the flow cell at room temperature for 30 minutes. We use a microscope to verify that we have a good distribution of poly-styrene beads in the flow cell. We then press the flow cell against a hot plate, at approximately 100 ^{o}C, for around a minute. The heat firmly localizes the polystyrene bead to the cover slip. The localized polystyrene beads are used as a reference for drift correction.

Next, we flow in a 45 µl mixture containing, 3 µl of anti-digoxigenin solution (0.2 mg/ml in 1× PBS), 2 µl of 0.1 nM 11 kbp DNA and 40 µl 10X PBS buffer. The mixture is incubated in the flow cell overnight to allow the formation of a stable bond between the anti-digoxigenin coated surface and the multi-digoxigenin handles of the 11 kb DNA. We then flow through 1 µm diameter magnetic beads (Invitrogen). Once incubated for 10 mins, we use wash buffer (1X PBS, 0.1% w/v BSA, 0.01% v/v Tween-20) buffer to flow out the unbound magnetic beads. We then use this flow cell to calibrate magnetic tweezers instrument.

Similar to a simple pendulum, the force applied to the DNA tethered magnetic bead can be calculated using the Brownian motion of the magnetic bead about its

equilibrium position. The force on the magnetic bead can be approximated using the expression[70]

$$F = \frac{k_b T \langle z \rangle}{\langle \delta y^2 \rangle} \quad (4.1)$$

where, k_b is the Boltzmann constant, T is the absolute temperature, $\langle z \rangle$ is the average DNA extension, and $\langle \delta y^2 \rangle$ is the standard variation in bead position in Y- direction (Fig. 4.6 (A)). We start the measurement with the magnet approximately 10 mm away from the flow cell. The position of DNA tethered magnetic bead provides a reference of the surface of the flow cell. We then move the magnet close to the flow cell. We refer to this position as the zero position of the magnet. Next, we incrementally move the magnets upwards in steps of 0.15 mm. We generally, repeat the measurement for 5 different DNA-tethered magnetic beads and use the average force at each magnet position to calibrate the force.

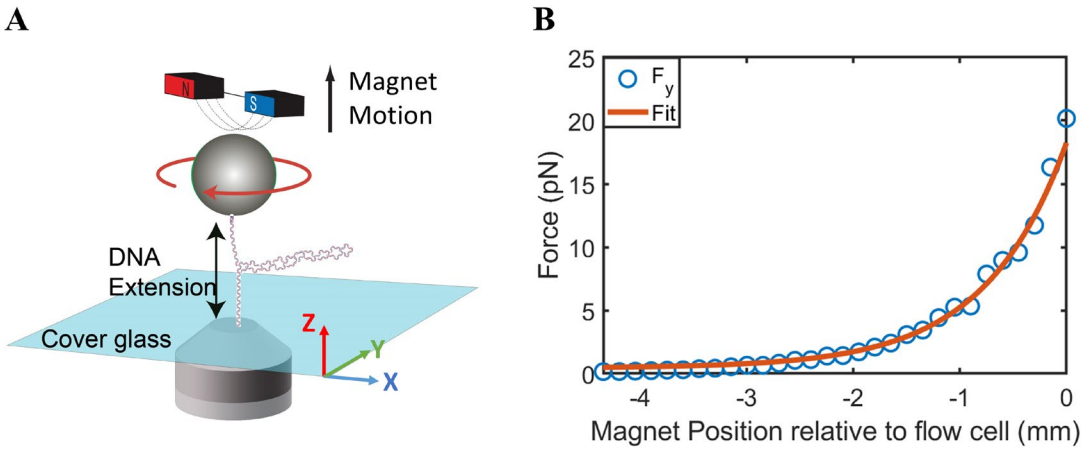

Figure 4.6 Force calibration of Magnetic tweezers. (A) Schematic of magnetic tweezers instrument. Force is calculated from the motion of DNA-tethered bead. The magnets apply the maximum force when it is closest to the DNA (position zero). Magnet is moved away from the flow surface (in Z direction) to capture change in force. **(B)** Variation of the force applied to the magnetic bead as a function of the relative Z-position of magnet with respect to cover glass.

The relative distance is measured in millimeters (mm). The force applied to the magnetic bead is calculated using the variance in Y-direction movement of a magnetic bead. A phenomenological double exponential is fit to the data to allow interpolation of force between the calibrated magnet positions.

Figure 4.6 (B) shows a typical force calibration curve for a single magnetic bead. F_y is the force obtained using the variance in bead position in the Y – direction (using Eq. 4.1 shown above). The position of the magnet is calculated relative to the cover glass. The magnet position closest to the flow cell is assigned a value of 0. Each force measure is performed by moving the magnet a distance of -0.15 millimeters (mm) in Z- direction (Fig. 4.6 (A)).

The forces obtained are fit to a double exponential equation, which can be used to interpolate the forces applied to the magnetic bead. The double exponential is a phenomenological fit, but it has previously been shown to accurately capture the variation of force with magnet position[126].

4.3.2 Presence of di-thiol group does not cause plectoneme localization

We plan to use the DNA substrate, generated here, to study plectoneme pinning at the mismatch in a supercoiled DNA molecule. We have previously shown that plectoneme pinning at the mismatch is caused by the reduction in local bending rigidity at the mismatch. Here, we verify that, under the force (< 2 pN) and salt concentrations (< 1 M NaCl) we plan to use in the rotor bead experiment, the presence of di-thiol group does not result in plectoneme pinning (Fig. 4.7 (A)). We generate a 4.5 kb DNA containing the di-thiol group using the procedure outlined in section 4.2.2 (Fig. 4.4

(A)). We then assemble a flow cell containing the 4.5 kb DNA using a procedure similar to the procedure outlined in section 4.3.1.

Fig. 4.7 (A) shows the response of the di-thiol containing generated DNA substrate to positive supercoiling. A single run is shown in Fig. 4.7 (A). A force of 2 pN is applied to the DNA. A 1 M NaCl buffer is used to observe the supercoiling. We have previously shown that high force and high salt concentrations increase the probability of pinning. The data shown in Fig. 4.7 (A) is for the highest force and salt concentration that will be used in the rotor bead assay. We induce supercoiling by rotating the magnetic bead tethered to the DNA. A torsionally constrained DNA absorbs the applied turns through a change in the DNA twist, until a critical twist is reached. The DNA molecule buckles and forms a plectoneme once this critical twist is reached. The buckling results in a reduction in DNA extension (Fig. 4.7 (A)), which can be observed by recording the vertical location of DNA-tethered magnetic bead. Additional turns applied to the DNA results in elongation of the plectoneme and results in additional continuous reduction in DNA length. In Fig. 4.7 (A) we apply turns at a rate of 1 turn every 10 seconds. Data acquisition is performed at 100 Hertz (Hz). The raw-timeseries data of DNA extension is shown with blue dots in Fig. 4.7 (A). The black curve shows average extension at each turn. In the DNA substrate generated here, we observe a continuous and uniform decrease in DNA extension once the initial buckling occurs (first gray bar in Fig. 4.7 (A)). This continuous and uniform decrease is expected when plectoneme pinning has not occurred. In Fig. 4.7 (A), we overlay the red curve to show the expected DNA extension data if plectoneme pinning was observed at the di-thiol group. If the plectoneme pinning occurs at the di-thiol group,

we will observe a second buckling transition (red curve (Fig. 4.7 (A))) once around 30% of DNA length has been absorbed by the plectoneme. Generally, we carry out multiple supercoiling experiments (> 30 DNA extension curves). We never observe a second buckling indicating that the plectoneme pinning unlikely occur at the di-thiolated DNA site under relatively high force (~ 2pN) and high salt (~ 1M MaCl).. The rotor bead experiment proposed here, will be performed at lower force (< 2 pN) and salt concentrations (< 1 M NaCl). Hence, we do not expect the di-thiol group to cause plectoneme pinning in the rotor bead experiments.

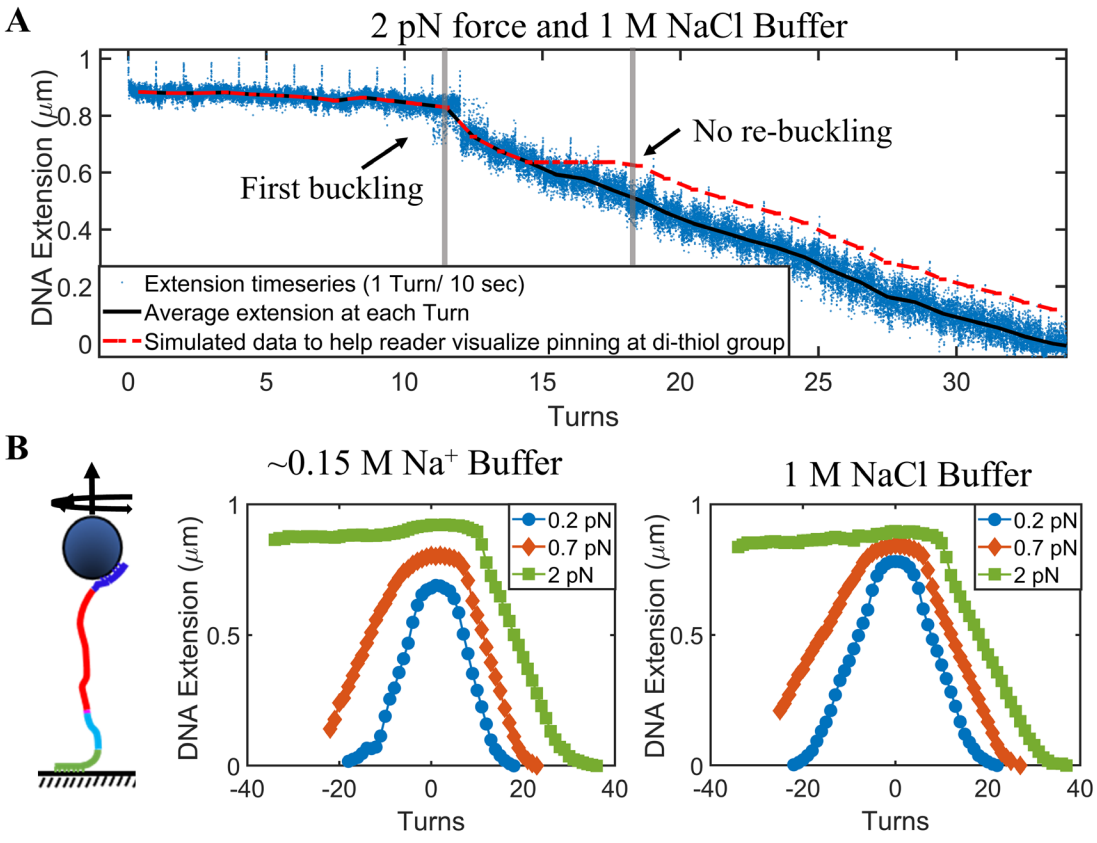

Figure 4.7 Supercoiling of DNA substrate containing a di-thiol group. (A) Data of DNA extension vs turns (excess linking number). DNA is in a 1 M NaCl buffer and 2 pN force is applied to the magnetic bead tethered to the DNA. Blue dots show the raw time series when turns are applied at a rate of 1 turn every 10 seconds. Black curves show the average DNA extension at each turn. A single buckling transition is observed for the DNA substrate

generated here (first gray bar). Plectoneme pinning at the di-thiol group would have caused a second buckling transition (second gray bar). The red curve is simulated data to help the reader visualize a hypothetical second buckling transition.[14] **(B)** Averaged DNA extension vs turns data for the DNA in two different buffers (~ 0.15 M NaCl buffer and 1 M NaCl salt buffer).

In Fig 4.7 (B), we show the average DNA extension change of supercoiled DNA containing di-thiol group. We use two different buffers (a ~0.15 M NaCl buffer and a 1 M NaCl buffer) to observe supercoiling of a 4.5 kb DNA containing di-thiol group. We also use three different tensions (0.2 pN, 0.7 pN and 2 pN). We typically start from a positively supercoiled DNA where the DNA tethered magnetic bead is very close to the surface. We then apply 1 turn every 10 seconds and cycle between positive and negative supercoiling. Generally, we use around 5 different DNA tethered magnetic bead and average the results. The DNA extension curves (Fig. 4.7 (B)) show that, similar to intact DNA, positive and negative supercoiling results in similar reductions in DNA extensions when relatively low forces (~ 0.2 pN, blue curve in Fig. 4.7 (B)) are applied to the DNA. DNA, stretched by intermediate force (~ 0.7 pN, red curve in Fig. 4.7 (B)), shows slight asymmetry in the extension curve. Intermediate forces result in slight local melting when DNA is negatively supercoiled. The local melting can absorb some of the tortional stress applied to the DNA and result in formation of a smaller plectoneme. The smaller plectoneme will cause a smaller reduction in DNA extension, and we will observe the DNA extension to be longer for a negatively supercoiled DNA compared to a positively supercoiled DNA. Application of high stretching forces (~ 2 pN, green curve in Fig. 4.7 (B)) will cause significant melting in a negatively supercoiled DNA. The melting will absorb the applied tortional stress in a negatively supercoiled DNA and we will observe no plectoneme formation.

The magnetic tweezers assays show that the presence of di-thiol group does not result in double buckling in DNA molecule (Fig. 4.7 (A)).

4.4 Conclusions

In this chapter of the thesis, we have provided a detailed procedure to perform single molecule rotor bead assays to provide direct experimental validation of the simulation results presented here (Fig. 4.1, Fig. 4.2 and Fig. 4.3). We have successfully generated the DNA substrate that can be used in the rotor bead assays (Fig 4.4). We have also verified that the procedure outlined in the thesis can be used to successfully add mismatches to the DNA substrate (Fig 4.5). We have used single-molecule magnetic tweezers assay to verify that the presence of di-thiol group does not result in plectoneme pinning (Fig 4.7). The results outlined in this chapter should form the basis of immediate future studies (not to be conducted as a part of this dissertation) that would enable first direct quantification of the localization of the plectoneme at mismatches for negatively supercoiled DNA in low-force and low-salt regime and in the process help to validate the predictions of our simulation results.

Chapter 5: Conclusions and Outlook

In this dissertation we have used MD simulations and single molecule magnetic tweezers assay to understand the effect of mismatches on supercoiled DNA. *In vivo*, mismatches need to be rapidly corrected. The mechanism used by mismatch repair proteins to find the mismatch is not well understood. The goal of this dissertation was to study the effect of mismatches on supercoiled DNA.

We first created a simulation framework using the OxDNA2 model to study plectoneme pinning in a DNA containing mismatches. To the best of our knowledge, we have for the first time shown that the OxDNA2 model can reproduce all atom simulations results on the effect of mismatches on local bending rigidity of DNA (Fig. 2.1). We then used the OxDNA2 model to study plectoneme localization in a positively supercoiled DNA containing mismatches. We found that the simulation framework used here can quantitively reproduce the plectoneme pinning observed in single molecule experiments at relatively high force (~2 pN) and relatively high salt concentrations (~1 M NaCl) (Fig. 2.5 (a)). We also find that in the high force-high salt condition, the simulation framework presented here can quantitatively reproduce theoretical predictions of plectoneme localization in the presence of mismatches (Fig. 2.6 (a)). Next, we use the simulation framework to observe plectoneme localization in a positively supercoiled DNA in lower force (~0.3 pN) and more physiologically relevant salt concentration (~ 0.2 M NaCl). We find that the under these conditions, the probability of plectoneme localization is lower compared to the high force- high salt conditions (Fig. 2.5 (b, c)). We also find that under low force (~0.3 pN) and low salt (0.2 M NaCl) conditions, the MD simulation results no longer quantitively agree with

the theoretical predictions of plectoneme pinning at the mismatches (Fig. 2.6 (b, c)). In the low force-low salt case, we also simulate DNA molecule of two different lengths. The entropic cost associated with plectoneme localization will be higher in a longer DNA. We find that, under similar conditions, the probability of plectoneme localization at the mismatch is lower for a longer DNA (Fig. 2.5 (c)). To the best of our knowledge, the simulations presented here are the first to show that plectoneme pinning can be observed in timescales reached by coarse-grained simulations.

We then use the simulation framework developed here, to study the effects of mismatches on a negatively supercoiled DNA. The presence of mismatches would promote local melting at the mismatches and this local melting could reduce the bending rigidity at the mismatch and this would promote plectoneme localization at the mismatch. Alternatively, instead of forming a plectoneme, the local melting at mismatches could absorb the negative tortional stress. We find that the presence of mismatches enhances plectoneme localization at the mismatch (Fig. 3.2 (A)). We also find that plectoneme pinning is not dependent on the amount of torsional stress applied to a negatively supercoiled DNA (Fig. 3.2 (B)). In order to investigate the competition between plectoneme pinning and local melting in a mismatch containing negatively supercoiled DNA, we also simulate a negatively supercoiled DNA under three different forces. We find that higher forces enhance plectoneme pinning in negatively supercoiled DNA containing 2 mismatches (Fig. 3.4 (A - C)). We don't observe a significant effect of tension on plectoneme pinning in a negatively supercoiled DNA containing 4 or 6 mismatches (Fig. 3.4 (A - C)). We also directly calculate the effect of tension on base flipping, i.e., local melting, in a negatively supercoiled DNA. We find

that tension enhances local melting at the mismatches (Fig. 3.5). We also observe that, under similar conditions, the size of the melted region is bigger in a negatively supercoiled DNA compared to a positively supercoiled DNA (Fig. 3.3). To the best of our knowledge, the simulation results presented here are the first direct evidence of plectoneme pinning in a negatively supercoiled DNA. The simulations of negatively supercoiled DNA containing mismatches can be used to define a phase space (range of tension and salt concentration) to observe plectoneme localization *in vitro*.

Finally, we outline a detailed framework to experimentally verify the simulation results presented here. We outline the procedure to perform a single-molecule rotor bead assay (Fig. 4.1). In the rotor bead assay, a small gold nanoparticle is attached to the DNA molecule. The gold nanoparticle is attached close to the mismatched base pairs; therefore, monitoring the motion of the gold nanoparticle allows us to directly observe plectoneme localization (Fig. 4.2). We generate a 4.5 kilo-base long DNA molecule containing a di-thiol group (Fig. 4.4 (C)). We also perform extensive single-molecule magnetic tweezers experiments to verify that the presence of di-thiol group does not result in pinning in an intact DNA molecule (Fig. 4.7 (A)). We also verify that the protocol suggested here can be used to create mismatched base pairs in a DNA molecule (Fig. 4.5).

This thesis highlights the advances that can be obtained by combining theoretical, simulation and experimental techniques to study a problem. In future, it would be beneficial to use the results of the rotor bead study to improve the simulation and theoretical models. The theoretical model developed by Brahmachari et al.[26] should generally apply to any DNA defect that can reduce the local bending rigidity of

the DNA. Improving the theoretical model, using the simulation and experimental results provided here, will provide an excellent tool to study the plectoneme pinning caused by other DNA defects. The rotor bead assay outlined here can be used to study plectoneme pinning in a DNA under different multivalent salt conditions. Generally, presence of multivalent ions reduces the local bending rigidity of DNA. This reduced rigidity could enhance plectoneme localization. The experimental framework outlined here, can for the first time, provide direct evidence of the effect of multivalent ions on plectoneme pinning in a DNA containing mismatches.

Another exciting application of our experimental framework would be to study the effect of protein binding on supercoiled DNA containing mismatches. The mechanism of MutS binding to DNA with mismatches is not well understood. The method proposed here can be used to study the mechanism of MutS. The magnetic tweezers instrument is combined with single-molecule fluorescence TIRF microscopy that can be used to measure the localization of fluorescently labeled MutS to mismatched DNA. The time to bind and lifetime of MutS binding to the mismatch can be directly probed as a function of supercoiling in this assay. This will shed light on the question that in part motivated this work – *can supercoiling facilitate mismatch detection by MutS and related mismatch repair proteins*?

Appendix A: DNA substrate generation

A.1 Materials

- pET28b plasmid (EMD4Biosciences)
- pBluescript plasmid (Stratagene)
- Phusion DNA polymerase (M0530, New England Biolabs) and buffer
- Taq DNA polymerase (M0273, New England Biolabs) and Taq pol standard buffer
- BsaI-HF (R3535, New England Biolabs)
- Nt.BbvcI (R0632L, New England Biolabs)
- T4 DNA ligase (M1804, Promega) and supplied buffer
- dNTP solution set (N0446, New England Biolabs) to prepare a low dTTP-dNTP mixture containing 10 mM dATP, dCTP, dGTP, and 6 mM of dTTP (N0447, New England Biolabs).
- Biotin-16-dUTP (11093070910, Roche)
- Digoxigenin -11-dUTP (11093088910, Roche)
- PCR Purification Kit (28104, Qiagen)

A.2 PCR amplification to produce DNA segments

A.2.1 Primers needed to perform PCR amplification:

- **3kb segment**

Template: pET28b
Forward primer: 5' GCTGGGTCTCGGTGTCGCCCTGCACCATTATGTTCCGGATCTG 3'
Reverse primer: 5' GCTGGGTCTCGCAAC CCGCTCCTTTCGCTTTCTTC 3'

- **439 bp segment**

Template: pkz2x
Forward primer:5'GCTGGGTCTCGACCA CGGATATAGTTCCTCCTTTC 3'
Reverse primer: 5'GCTGGGTCTCGCAGC TTGTGAGCGGATAACAATTC 3'

- **Multi-biotin handles**

Template: pBluecript II KS (+)
Forward primer: 5' GCTGGGTCTCGGTTG TTCCCTTTAGTGAGGGTTAATTG 3'
Reverse primer: 5' TATAGTCCTGTCGGGTTTCG 3'

- **Multi-Digoxigenin handles**

Template: pBluecript II KS (+)

Forward primer: 5' GCTGGGTCTCGTGGT TTCCCTTTAGTGAGGGTTAATTG 3'
Reverse primer: 5' TATAGTCCTGTCGGGTTTCG 3'

A.2.2 Protocol for PCR amplification of segments

A.2.2.1 PCR amplification of 3kb and 439 bp segment

1. Prepare 1x 1.5 ml eppendorf tube and 6x 200 µl PCR tubes on ice.
2. Add the following dependent on the reaction volume to 1.5 ml tube and mix well. Substitute the appropriate forward primer, reverse primer and template for 3kb and 439 bp segments

Component	50 µl Reaction	300 µl Reaction	Final Concentration
Nuclease-free water	to 50 µl	to 300 µl	
5X Phusion HF Buffer	10 µl	60 µl	1X
100 µM Forward Primer	0.25 µl	1.5 µl	0.5 µM
100 µM Reverse Primer	0.25 µl	1.5 µl	0.5 µM
Template	5 ng	30 ng	<5 ng/µl
10 mM dNTPs	1 µl	6 µl	200 µM
Phusion DNA Polymerase	0.5 µl	3 µl	1.0 units/50 µl PCR

3. Transfer 50 µl of the reaction mixture to each PCR tube.
4. PCR condition

STEP	TEMP	TIME
Initial Denaturation	98°C	30 s
30 Cycles	98°C 63°C 72°C	10 s 15 s 20 s *x kb
Final Extension	72°C	30 min
Hold	4°C	

5. Check the size of PCR products on a 1% DNA agarose gel.
6. After PCR reaction, purify DNA using Qiagen PCR purification kit.
7. Quantify DNA product and its purity.

A.2.2.2 PCR amplification of multi-biotin and multi-digoxigenin handle

1. Prepare 2x 1.5 ml eppendorf tubes and 12 x 200 µl PCR tubes on ice.
2. Add the following to 1.5 ml tube, dependent on the reaction volume and labeling type (one for biotin and the other for digoxigenin) and mix:

Component	50 µl Reaction	300 µl Reaction	Final Concentration
Nuclease-free water	to 50 µl	to 300 µl	

10 X standard Taq Buffer	5 µl	30 µl	1X
100 µM Forward Primer (Either bio or dig)	0.25 µl	1.5 µl	0.5 µM
100 µM Reverse Primer	0.25 µl	1.5 µl	0.5 µM
pBluescript II KS	1 ng	6 ng	
low dTTP-dNTP mixture (10 mM of dATP, dCTP, dGTT and 7.0 mM of dTTP)	1 µl	6 µl	200 µM except for dTTP (140 µM)
Either Biotin-16-dUTP or Digoxigenin -11-dUTP*	3 µl	12 µl	60 µM
Taq DNA pol	0.5 µl	3 µl	1.0 units/50 µl PCR

3. Transfer 50 µl of the reaction mixture to each PCR tube.
   PCR condition

STEP	TEMP	TIME
Initial Denaturation	94°C	2 min
33 Cycles	94°C 60°C 68°C	20 seconds 30 seconds 2 minute
Final Extension	68°C	30 minutes
Hold	4-10°C	

4. Check the size of PCR products on a 1% DNA agarose gel.
5. After PCR reaction, purify DNA using PCR purification kit.
6. Quantify DNA product and calculate in terms of molar scale.

A.3 Restriction Digestion to produce 4 bp overhangs

1. Add the mixture of following components into a 200 µl PCR tube:

Component	50 µl Reaction
Nuclease-free water	to 50 µl
10x Buffer 4	5 µl
PCR product	<1 µg
BsaI-HF	3 µl

2. Restriction digestion condition

STEP	TEMP	TIME
Incubation	37°C	6 hrs
Add additional 2 µl BsaI-HF	37°C	
Incubation	37°C	6 hrs
Add additional 2 µl BsaI-HF	37°C	
Incubation	37°C	Overnight

Stop reaction	80°C	20 min
Hold	4°C	

3. Purify the digestion product using Qiagen PCR product purification kit.

4. Quantify the purified DNA in molar scale.

A.4 Ligation of segments with T4 DNA ligase

1. Use the digested DNA segments
- Add appropriate substrate. Example shown for ligation of 3 kb and multi-biotin segments the mixture of following components into a 200 µl PCR tube:

Component	**50 µl Reaction**	**Final concentration**
Nuclease-free water	to 50 µl	
10x T4 DNA ligase Buffer	5 µl	1x
3 kb segment	~0.10 pmol	~2 nM
Biotinylated DNA handle	~0.10 pmol	~2 nM
T4 DNA ligase	1 µl	1 unit

3. T4 DNA ligation condition

STEP	**TEMP**	**TIME**
Incubation	25°C	3 hrs
Incubation	16°C	15 hrs
Hold	4°C	

4. Purify DNA using PCR purification kit

5. Run the ligated product on 1% DNA agarose gel to check the ligation efficiency.

A.5 Incorporation of mismatch to DNA

A.5.1 Materials required to introduce mismatches in 439 bp segment

- Digested 439 bp segment
- Nt.BbvCI enzyme and buffer (R0632S, New England Biolabs)
- PstI-HF enzyme and buffer (R3140S, New England Biolabs)
- "Top oligo" (complementary oligo for lift-off single-stranded DNA region for gapping)
- Mismatch containing oligonucleotide

A.5.2 Nicking the 439 bp segment

1. Add the mixture of following components into a 200 µl PCR tube:

Component	**50 µl Reaction**
Nuclease-free water	to 50 µl
10x Buffer 4	5 µl
Digested 439 bp segment	~50 nM
Nt.BbvCI	3 µl

2. Nicking condition

STEP	TEMP	TIME
Incubation	37°C	6 hrs
Add additional 2 µl Nt.BbvCI	37°C	
Incubation	37°C	Overnight
add "Top oligos". Add 3x amount of 439 bp segment		
Stop reaction	80°C	20 min
Hold	4°C	

3. Purify the digestion product using Qiagen PCR product purification kit.

A.5.3 Introduce mismatch to nicked 439 bp segment

1. Add the mixture of following components into a 200 µl PCR tube:

	Upto 20 ul
Nicked 439 bp DNA	~20 nM
Mismatch containing Oligonucleotide	3x Nicked Gap DNA
Anneal buffer	4ul

2. Heat mixture to 75 ºC then cool to 4 ºC at 1 ºC/10 sec
3. Add 5 ul ligase buffer, 1 ul ligase and 24 ul water to mixture.
4. Cleave with PstI to verify addition of mismatch

A.7 Single molecule magnetic tweezers instrument

Detailed protocol to perform single molecule magnetic tweezers experiments can be found in Seol et al.[125]

Bibliography

1. Dahm, R. Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. *Hum. Genet.* **122**, 565–581 (2008).
2. Avery, O. T., MacLeod, C. M. & McCarty, M. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Deoxyribonucliec Acid Fraction Isolated from Pneumococcus type III. *J. Exp. Med.* **79**, 137–158 (1944).
3. Franklin, R. E. & Gosling, R. G. Evidence for 2-Chain Helix in Crystalline Structure of Sodium Deoxyribonucleate. *Nature* **172**, 156–157 (1953).
4. Klug, A. Rosalind Franklin and the Discovery of the Structure of DNA. *Nature* **219**, 808–810 (1968).
5. Watson, J. D. & Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **171**, 737–738 (1953).
6. Baranello, L., Levens, D., Gupta, A. & Kouzine, F. The importance of being supercoiled: how DNA mechanics regulate dynamic processes. *Biochim. Biophys. Acta* **1819**, 632–638 (2012).
7. Mielke, S. P., Fink, W. H., Krishnan, V. V, Grønbech-Jensen, N. & Benham, C. J. Transcription-driven twin supercoiling of a DNA loop: A Brownian dynamics study. *J. Chem. Phys.* **121**, 8104–8112 (2004).
8. Liu, L. F. & Wang, J. C. Supercoiling of the DNA template during transcription. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 7024–7027 (1987).
9. Fuller, F. B. Decomposition of the linking number of a closed ribbon: A problem from molecular biology. *Proc. Natl. Acad. Sci.* **75**, 3557–3561 (1978).

10. White, J. H. Self-Linking and the Gauss Integral in Higher Dimensions. *Am. J. Math.* **91**, 693–728 (1969).

11. Purohit, P. K. Plectoneme formation in twisted fluctuating rods. *J. Mech. Phys. Solids* **56**, 1715–1729 (2008).

12. Benham, C. J. & Mielke, S. P. DNA Mechanics. *Annu. Rev. Biomed. Eng.* **7**, 21–53 (2005).

13. Seol, Y. & Neuman, K. C. The dynamic interplay between DNA topoisomerases and DNA topology. *Biophys. Rev.* **8**, 101–111 (2016).

14. Dittmore, A., Brahmachari, S., Takagi, Y., Marko, J. F. & Neuman, K. C. Supercoiling DNA Locates Mismatches. *Phys. Rev. Lett.* **119**, 147801 (2017).

15. Ganji, M., Kim, S. H., van der Torre, J., Abbondanzieri, E. & Dekker, C. Intercalation-Based Single-Molecule Fluorescence Assay To Study DNA Supercoil Dynamics. *Nano Lett.* **16**, 4699–4707 (2016).

16. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715 (1993).

17. Raiber, E.-A., Hardisty, R., van Delft, P. & Balasubramanian, S. Mapping and elucidating the function of modified bases in DNA. *Nat. Rev. Chem.* **1**, 69 (2017).

18. Hsiao, Y.-Y., Fang, W.-H., Lee, C.-C., Chen, Y.-P. & Yuan, H. S. Structural Insights Into DNA Repair by RNase T—An Exonuclease Processing 3′ End of Structured DNA in Repair Pathways. *PLOS Biol.* **12**, e1001803 (2014).

19. Skoko, D., Wong, B., Johnson, R. C. & Marko, J. F. Micromechanical Analysis of the Binding of DNA-Bending Proteins HMGB1, NHP6A, and HU

Reveals Their Ability To Form Highly Stable DNA−Protein Complexes. *Biochemistry* **43**, 13867–13874 (2004).

20. Kunkel, T. A. & Erie, D. A. DNA MISMATCH REPAIR. *Annu. Rev. Biochem.* **74**, 681–710 (2005).

21. Lamers, M. H. *et al.* The crystal structure of DNA mismatch repair protein MutS binding to a G·T mismatch. *Nature* **407**, 711–717 (2000).

22. Obmolova, G., Ban, C., Hsieh, P. & Yang, W. Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature* **407**, 703–710 (2000).

23. Warren, J. J. *et al.* Structure of the Human MutSα DNA Lesion Recognition Complex. *Mol. Cell* **26**, 579–592 (2007).

24. Yang, W. Structure and mechanism for DNA lesion recognition. *Cell Res.* **18**, 184–197 (2007).

25. Goyal, S. & Perkins, N. C. Looping mechanics of rods and DNA with non-homogeneous and discontinuous stiffness. *Int. J. Non. Linear. Mech.* **43**, 1121–1129 (2008).

26. Brahmachari, S. *et al.* Defect-facilitated buckling in supercoiled double-helix DNA. *Phys. Rev. E* **97**, 22416 (2018).

27. Ramstein, J. & Lavery, R. Energetic coupling between DNA bending and base pair opening. *Proc. Natl. Acad. Sci.* **85**, 7231–7235 (1988).

28. Sharma, M., Predeus, A. V, Mukherjee, S. & Feig, M. DNA Bending Propensity in the Presence of Base Mismatches: Implications for DNA Repair. *J. Phys. Chem. B* **117**, 6194–6205 (2013).

29. Pengyu, H. *et al.* Recurrent mismatch binding by MutS mobile clamps on DNA localizes repair complexes nearby. *Proc. Natl. Acad. Sci.* **117**, 17775–17784 (2020).

30. Au, K. G., Welsh, K. & Modrich, P. Initiation of methyl-directed mismatch repair. *J. Biol. Chem.* **267**, 12142–12148 (1992).

31. Fernandez-Leiro, R. *et al.* The selection process of licensing a DNA mismatch for repair. *Nat. Struct. Mol. Biol.* **28**, 373–381 (2021).

32. Brackley, C. A., Cates, M. E. & Marenduzzo, D. Facilitated Diffusion on Mobile DNA: Configurational Traps and Sequence Heterogeneity. *Phys. Rev. Lett.* **109**, 168103 (2012).

33. van den Broek, B., Lomholt, M. A., Kalisch, S.-M. J., Metzler, R. & Wuite, G. J. L. How DNA coiling enhances target localization by proteins. *Proc. Natl. Acad. Sci.* **105**, 15738–15742 (2008).

34. Lee, I., Dong, K. C. & Berger, J. M. The role of DNA bending in type IIA topoisomerase function. *Nucleic Acids Res.* **41**, 5444–5456 (2013).

35. Cournac, A. & Plumbridge, J. DNA looping in prokaryotes: experimental and theoretical approaches. *J. Bacteriol.* **195**, 1109–1119 (2013).

36. Kouzine, F., Sanford, S., Elisha-Feil, Z. & Levens, D. The functional response of upstream DNA to dynamic supercoiling in vivo. *Nat. Struct. Mol. Biol.* **15**, 146–154 (2008).

37. Šponer, J. *et al.* Relative Stability of Different DNA Guanine Quadruplex Stem Topologies Derived Using Large-Scale Quantum-Chemical Computations. *J. Am. Chem. Soc.* **135**, 9785–9796 (2013).

38. Gkionis, K. *et al.* Ion Binding to Quadruplex DNA Stems. Comparison of MM and QM Descriptions Reveals Sizable Polarization Effects Not Included in Contemporary Simulations. *J. Chem. Theory Comput.* **10**, 1326–1340 (2014).

39. Cornell, W. D. *et al.* A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **117**, 5179–5197 (1995).

40. Hart, K. *et al.* Optimization of the CHARMM Additive Force Field for DNA: Improved Treatment of the BI/BII Conformational Equilibrium. *J. Chem. Theory Comput.* **8**, 348–362 (2012).

41. Weiner, S. J. *et al.* A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**, 765–784 (1984).

42. Ivani, I. *et al.* Parmbsc1: a refined force field for DNA simulations. *Nat. Methods* **13**, 55–58 (2016).

43. Zgarbová, M. *et al.* Refinement of the Sugar–Phosphate Backbone Torsion Beta for AMBER Force Fields Improves the Description of Z- and B-DNA. *J. Chem. Theory Comput.* **11**, 5723–5736 (2015).

44. Minhas, V. *et al.* Modeling DNA Flexibility: Comparison of Force Fields from Atomistic to Multiscale Levels. *J. Phys. Chem. B* **124**, 38–49 (2020).

45. Galindo-Murillo, R., Roe, D. R. & Cheatham, T. E. Convergence and reproducibility in molecular dynamics simulations of the DNA duplex d(GCACGAACGAACGAACGC). *Biochim. Biophys. Acta - Gen. Subj.* **1850**, 1041–1058 (2015).

46. Dans, P. D. *et al.* How accurate are accurate force-fields for B-DNA? *Nucleic*

Acids Res. **45**, 4217–4230 (2017).

47. Yin, Y. *et al.* Dynamics of spontaneous flipping of a mismatched base in DNA duplex. *Proc. Natl. Acad. Sci.* **111**, 8043–8048 (2014).

48. Etheve, L., Martin, J. & Lavery, R. Protein–DNA interfaces: a molecular dynamics analysis of time-dependent recognition processes for three transcription factors. *Nucleic Acids Res.* **44**, 9990–10002 (2016).

49. Savelyev, A. & Papoian, G. A. Molecular Renormalization Group Coarse-Graining of Polymer Chains: Application to Double-Stranded DNA. *Biophys. J.* **96**, 4044–4052 (2009).

50. Ouldridge, T. E., Louis, A. A. & Doye, J. P. K. Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *J. Chem. Phys.* **134**, 085101 (2011).

51. Hinckley, D. M., Freeman, G. S., Whitmer, J. K. & de Pablo, J. J. An experimentally-informed coarse-grained 3-site-per-nucleotide model of DNA: Structure, thermodynamics, and dynamics of hybridization. *J. Chem. Phys.* **139**, 144903 (2013).

52. Korolev, N., Luo, D., Lyubartsev, A. P. & Nordenskiöld, L. A coarse-grained DNA model parameterized from atomistic simulations by inverse Monte Carlo. *Polymers.* **6**, 1655–1675 (2014).

53. Maffeo, C. & Aksimentiev, A. MrDNA: a multi-resolution model for predicting the structure and dynamics of DNA systems. *Nucleic Acids Res.* **48**, 5135–5146 (2020).

54. Machado, M. R. & Pantano, S. Exploring LacI–DNA Dynamics by Multiscale

Simulations Using the SIRAH Force Field. *J. Chem. Theory Comput.* **11**, 5012–5023 (2015).

55. Chakraborty, D., Hori, N. & Thirumalai, D. Sequence-Dependent Three Interaction Site Model for Single- and Double-Stranded DNA. *J. Chem. Theory Comput.* **14**, 3763–3779 (2018).

56. Maciejczyk, M., Spasic, A., Liwo, A. & Scheraga, H. A. DNA Duplex Formation with a Coarse-Grained Model. *J. Chem. Theory Comput.* **10**, 5020–5035 (2014).

57. Maffeo, C., Ngo, T. T. M., Ha, T. & Aksimentiev, A. A Coarse-Grained Model of Unstructured Single-Stranded DNA Derived from Atomistic Simulation and Single-Molecule Experiment. *J. Chem. Theory Comput.* **10**, 2891–2896 (2014).

58. Ouldridge, T. E., Louis, A. A. & Doye, J. P. K. DNA Nanotweezers Studied with a Coarse-Grained Model of DNA. *Phys. Rev. Lett.* **104**, 178101 (2010).

59. Ivenso, I. D. & Lillian, T. D. Simulation of DNA Supercoil Relaxation. *Biophys. J.* **110**, 2176–2184 (2016).

60. Coronel, L., Suma, A. & Micheletti, C. Dynamics of supercoiled DNA with complex knots: large-scale rearrangements and persistent multi-strand interlocking. *Nucleic Acids Res.* **46**, 7533–7541 (2018).

61. Snodin, B. E. K. *et al.* Introducing improved structural properties and salt dependence into a coarse-grained model of DNA. *J. Chem. Phys.* **142**, 234901 (2015).

62. Matek, C., Ouldridge, T. E., Doye, J. P. K. & Louis, A. A. Plectoneme tip bubbles: Coupled denaturation and writhing in supercoiled DNA. *Sci. Rep.* **5**,

7655 (2015).

63. Kriegel, F. *et al.* The temperature dependence of the helical twist of DNA. *Nucleic Acids Res.* **46**, 7998–8009 (2018).

64. Vologodskii, A. Determining Protein-Induced DNA Bending in Force-Extension Experiments: Theoretical Analysis. *Biophys. J.* **96**, 3591–3599 (2009).

65. Bryant, Z., Oberstrass, F. C. & Basu, A. Recent developments in single-molecule DNA mechanics. *Curr. Opin. Struct. Biol.* **22**, 304–312 (2012).

66. Greenleaf, W. J., Woodside, M. T. & Block, S. M. High-Resolution, Single-Molecule Measurements of Biomolecular Motion. *Annu. Rev. Biophys. Biomol. Struct.* **36**, 171–190 (2007).

67. Neuman, K. C., Lionnet, T. & Allemand, J.-F. Single-Molecule Micromanipulation Techniques. *Annu. Rev. Mater. Res.* **37**, 33–67 (2007).

68. Neuman, K. C. Single-molecule Measurements of DNA Topology and Topoisomerases. *J. Biol. Chem.* **285**, 18967–18971 (2010).

69. Bustamante, C., Bryant, Z. & Smith, S. B. Ten years of tension: single-molecule DNA mechanics. *Nature* **421**, 423–427 (2003).

70. Strick, T. R., Allemand, J.-F., Bensimon, D., Bensimon, A. & Croquette, V. The Elasticity of a Single Supercoiled DNA Molecule. *Science (80-.).* **271**, 1835–1837 (1996).

71. van Oene, M. M. *et al.* Biological Magnetometry: Torque on Superparamagnetic Beads in Magnetic Fields. *Phys. Rev. Lett.* **114**, 218301 (2015).

72. Neuman, K. C. & Nagy, A. Single-molecule force spectroscopy : optical tweezers , magnetic tweezers and atomic force microscopy. **5**, 491–505 (2008).

73. Lebel, P., Basu, A., Oberstrass, F. C., Tretter, E. M. & Bryant, Z. Gold rotor bead tracking for high-speed measurements of DNA twist, torque and extension. *Nat. Methods* **11**, 456–462 (2014).

74. Lohman, T. M. & Bjornson, K. P. Mechanisms of helicase-catalyzed DNA unwinding. *Annu. Rev. Biochem.* **65**, 169–214 (1996).

75. Kouzine, F. & Levens, D. Supercoil-driven DNA structures regulate genetic transactions. *Front. Biosci.* **12**, 4409–4423 (2007).

76. Dunaway, M. & Ostrander, E. A. Local domains of supercoiling activate a eukaryotic promoter in vivo. *Nature* **361**, 746–748 (1993).

77. Wang, G. & Vasquez, K. Effects of Replication and Transcription on DNA Structure-Related Genetic Instability. *Genes.* **8**, 17 (2017).

78. Postow, L., Hardy, C. D., Arsuaga, J. & Cozzarelli, N. R. Topological domain structure of the Escherichia coli chromosome. *Genes Dev.* **18**, 1766–1779 (2004).

79. Champoux, J. J. DNA Topoisomerases: Structure, Function, and Mechanism. *Annu. Rev. Biochem.* **70**, 369–413 (2001).

80. Argudo, D. & Purohit, P. K. The dependence of DNA supercoiling on solution electrostatics. *Acta Biomater.* **8**, 2133–2143 (2012).

81. Brutzer, H., Luzzietti, N., Klaue, D. & Seidel, R. Energetics at the DNA supercoiling transition. *Biophys. J.* **98**, 1267–1276 (2010).

82. Clauvelin, N., Audoly, B. & Neukirch, S. Mechanical response of plectonemic

DNA: An analytical solution. *Macromolecules* **41**, 4479–4483 (2008).

83. Baumann, C. G., Smith, S. B., Bloomfield, V. A. & Bustamante, C. Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci.* **94**, 6185–6190 (1997).

84. Geggier, S. & Vologodskii, A. Sequence dependence of DNA bending rigidity. *Proc. Natl. Acad. Sci.* **107**, 15421–15426 (2010).

85. Mitchell, J. S., Glowacki, J., Grandchamp, A. E., Manning, R. S. & Maddocks, J. H. Sequence-Dependent Persistence Lengths of DNA. *J. Chem. Theory Comput.* **13**, 1539–1555 (2017).

86. Marko, J. F. & Siggia, E. D. Statistical mechanics of supercoiled DNA. *Phys. Rev. E* **52**, 2912–2938 (1995).

87. Moroz, J. D. & Nelson, P. Torsional directed walks, entropic elasticity, and DNA twist stiffness. *Proc. Natl. Acad. Sci.* **94**, 14418–14422 (1997).

88. Jian, H., Vologodskii, A. V. & Schlick, T. A combined wormlike-chain and bead model for dynamic simulations of long linear DNA. *J. Comput. Phys.* **136**, 168–179 (1997).

89. Mitchell, J. S., Laughton, C. A. & Harris, S. A. Atomistic simulations reveal bubbles, kinks and wrinkles in supercoiled DNA. *Nucleic Acids Res.* **39**, 3928–3938 (2011).

90. Fathizadeh, A., Schiessel, H. & Ejtehadi, M. R. Molecular Dynamics Simulation of Supercoiled DNA Rings. *Macromolecules* **48**, 164–172 (2015).

91. Krajina, B. A. & Spakowitz, A. J. Large-Scale Conformational Transitions in Supercoiled DNA Revealed by Coarse-Grained Simulation. *Biophys. J.* **111**,

1339–1349 (2016).

92. Lepage, T., Képès, F. & Junier, I. Thermodynamics of Long Supercoiled Molecules: Insights from Highly Efficient Monte Carlo Simulations. *Biophys. J.* **109**, 135–143 (2015).

93. Henrich, O., Gutiérrez Fosado, Y. A., Curk, T. & Ouldridge, T. E. Coarse-grained simulation of DNA using LAMMPS: An implementation of the oxDNA model and its applications. *Eur. Phys. J. E* **41**, 57 (2018).

94. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics* vol. 117 1–19 (1995).

95. Nomidis, S. K., Kriegel, F., Vanderlinden, W., Lipfert, J. & Carlon, E. Twist-Bend Coupling and the Torsional Response of Double-Stranded DNA. *Phys. Rev. Lett.* **118**, 217801 (2017).

96. Suma, A. & Micheletti, C. Pore translocation of knotted DNA rings. *Proc. Natl. Acad. Sci.* **114**, E2991–E2997 (2017).

97. Harrison, R. M., Romano, F., Ouldridge, T. E., Louis, A. A. & Doye, J. P. K. Identifying Physical Causes of Apparent Enhanced Cyclization of Short DNA Molecules with a Coarse-Grained Model. *J. Chem. Theory Comput.* **15**, 4660–4672 (2019).

98. Kingsland, A. & Maibaum, L. DNA Base Pair Mismatches Induce Structural Changes and Alter the Free-Energy Landscape of Base Flip. *J. Phys. Chem. B* **122**, 12251–12259 (2018).

99. Fiorin, G., Klein, M. L. & Hénin, J. Using collective variables to drive molecular dynamics simulations. *Mol. Phys.* **111**, 3345–3362 (2013).

100. Roux, B. The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun.* **91**, 275–282 (1995).

101. Kim, S. H. *et al.* DNA sequence encodes the position of DNA supercoils. *Elife* **7**, e36557 (2018).

102. Davidchack, R. L., Ouldridge, T. E. & Tretyakov, M. V. New Langevin and gradient thermostats for rigid body dynamics. *J. Chem. Phys.* **142**, 144114 (2015).

103. Skoruppa, E., Laleman, M., Nomidis, S. K. & Carlon, E. DNA elasticity from coarse-grained simulations: The effect of groove asymmetry. *J. Chem. Phys.* **146**, 214902 (2017).

104. Stukowski, A. Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool. *Model. Simul. Mater. Sci. Eng.* **18**, 015012 (2010).

105. Harrison, R. M., Romano, F., Ouldridge, T. E., Louis, A. A. & Doye, J. P. K. Coarse-grained modelling of strong DNA bending I: Thermodynamics and comparison to an experimental 'molecular vice'. **arXiv**, 1506.09005 (2015).

106. Curuksu, J., Zakrzewska, K. & Zacharias, M. Magnitude and direction of DNA bending induced by screw-axis orientation: influence of sequence, mismatches and abasic sites. *Nucleic Acids Res.* **36**, 2268–2283 (2008).

107. Curuksu, J., Zacharias, M., Lavery, R. & Zakrzewska, K. Local and global effects of strong DNA bending induced during molecular dynamics simulations. *Nucleic Acids Res.* **37**, 3766–3773 (2009).

108. Sankararaman, S. & Marko, J. F. Formation of loops in DNA under tension.

Phys. Rev. E **71**, 021911 (2005).

109. Matek, C. Statistical mechanics of nucleic acids under mechanical stress. *Ph.D Thesis, Oxford University* (2014).

110. Pyne, A. L. B. *et al.* Base-pair resolution analysis of the effect of supercoiling on DNA flexibility and major groove recognition by triplex-forming oligonucleotides. *Nat. Commun.* **12**, 1053 (2021).

111. Rossetti, G. *et al.* The structural impact of DNA mismatches. *Nucleic Acids Res.* **43**, 4309–4321 (2015).

112. Van Loenhout, M. T. J., De Grunt, M. V. & Dekker, C. Dynamics of DNA supercoils. *Science (80-.).* **338**, 94–97 (2012).

113. Kolodner, R. D. & Marsischky, G. T. Eukaryotic DNA mismatch repair. *Curr. Opin. Genet. Dev.* **9**, 89–96 (1999).

114. Buermeyer, A. B., Deschênes, S. M., Baker, S. M. & Liskay, R. M. Mammalian DNA Mismatch Repair. *Annu. Rev. Genet.* **33**, 533–564 (1999).

115. Gupta, S., Gellert, M. & Yang, W. Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops. *Nat. Struct. Mol. Biol.* **19**, 72–79 (2012).

116. Kim, J., Zwieb, C., Wu, C. & Adhya, S. Bending of DNA by gene-regulatory proteins: construction and use of a DNA bending vector. *Gene* **85**, 15–23 (1989).

117. V., K. S., Sawako, S., Paula, V., M., C. D. & Anjum, A. Direct observation of DNA bending/unbending kinetics in complex with DNA-bending protein IHF. *Proc. Natl. Acad. Sci.* **103**, 18515–18520 (2006).

118. Law, S. M. & Feig, M. Base-Flipping Mechanism in Postmismatch Recognition by MutS. *Biophys. J.* **101**, 2223–2231 (2011).

119. LeBlanc, S. J. *et al.* Coordinated protein and DNA conformational changes govern mismatch repair initiation by MutS. *Nucleic Acids Res.* **46**, 10782–10795 (2018).

120. Charvin, G., Bensimon, D. & Croquette, V. Single-molecule study of DNA unlinking by eukaryotic and prokaryotic type-II topoisomerases. *Proc. Natl. Acad. Sci.* **100**, 9820 LP – 9825 (2003).

121. Petrushenko, Z. M. *et al.* Real-Time Detection of Single-Molecule DNA Compaction by Condensin I. *Curr. Biol.* **14**, 874–880 (2010).

122. Cui, Y., Petrushenko, Z. M. & Rybenkov, V. V. MukB acts as a macromolecular clamp in DNA condensation. *Nat. Struct. Mol. Biol.* **15**, 411–418 (2008).

123. Rybenkov, V. V, Vologodskii, A. V & Cozzarelli, N. R. The effect of ionic conditions on the conformations of supercoiled DNA. II. equilibrium catenation11Edited by D. E. Draper. *J. Mol. Biol.* **267**, 312–323 (1997).

124. Vologodskii, A. V & Cozzarelli, N. R. Conformational and Thermodynamic Properties of Supercoiled DNA. *Annu. Rev. Biophys. Biomol. Struct.* **23**, 609–643 (1994).

125. Seol, Y. & Neuman, K. C. Combined Magnetic Tweezers and Micro-mirror Total Internal Reflection Fluorescence Microscope for Single-Molecule Manipulation and Visualization BT - Single Molecule Analysis: Methods and Protocols. in (ed. Peterman, E. J. G.) 297–316 (Springer New York, 2018).

doi:10.1007/978-1-4939-7271-5_16.

126. Yu, Z. *et al.* A force calibration standard for magnetic tweezers. *Rev. Sci. Instrum.* **85**, 123114 (2014).