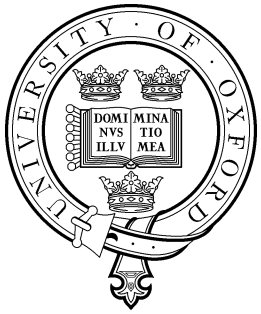
Statistical Mechanics of Nucleic Acids under Mechanical Stress



Christian Matek
Merton College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2014

The isolated knowledge obtained by a group of specialists
in a narrow field has in itself no value whatsoever,
but only in its synthesis with all the rest of knowledge
and only inasmuch as it really contributes in this synthesis
toward answering the demand, "Who are we?"
Erwin Schrödinger, Science and Humanism (1951)

Le secret d'ennuyer est celui de tout dire.
Voltaire, Sept Discours en Vers sur l'Homme (1738)

Acknowledgements

Many people have supported and encouraged me in many ways during the work on this thesis; it is a pleasure to thank them all.

During the process of working on this thesis, I have greatly benefited from many interactions with colleagues who shared their experience and ideas, both in Physics and Biology. Even if as brief listings, acknowledgements necessarily fail to do justice to the entirety of insights gained from the community, I would like to highlight the most important ones.

Firstly, I am indebted to my supervisors Ard Louis and Jonathan Doye for their guidance and support, both in fruitful suggestions and indispensable resources, during my time as a graduate student.

Thanks extend to many other colleagues at Oxford for stimulating discussions both on the science of physics and the art of programming. In particular, it was a pleasure to collaborate with Tom Ouldridge, Flavio Romano and Petr Šulc, who shared much of their knowledge and intuition. Valuable insights came from discussions with fellow workers in the field of DNA biophysics and nanotechnology, in particular Ben Snodin, John Schreck, Majid Mosayebi, Lorenzo Rovigatti, Agnes Noy, Ryan Harrison, Ferdinando Randisi and Pedro Fonseca, as well as Sarah Harris at Leeds and the other collaborators in the minicircle project.

Outside my own field of research, I have learned much about the manifold aspects of physics in conversations with fellow students, in particular with Andrei Constantin, Simon Davenport, Michael Klaput, Adam Nahum, Neil Robinson, Steffen Schaper, Lukas Witkowski and Irwin Zaid.

Over the years, I had the opportunity to attend several international meetings and schools, during which conversations with John Marko, Tim Liedl and Jan Lipfert were particularly fruitful and rewarding.

Special thanks go to Jean-François Allemand, David Bensimon and Vincent Croquette for admitting me as a visitor in their lab and allowing me to gain some first-hand experience with the fascinating single molecule experiments carried out in the Laboratoire de Physique Statistique at the École Normale Supérieure in Paris.

I am grateful to Erwin Frey and Ulrich Gerland at LMU München for fostering my interest in Biophysics from a very early stage of my studies, and to the Conceptual Foundations of Systems Biology group at Balliol College, Oxford headed by Denis Noble, for providing a communicative and open environment for discussions on the big questions of the field.

Finally, I would like to thank Julia Yeomans and Terence Strick for agreeing to be my examiners.

Beyond the professional domain, friends and family have been a constant and invaluable source of support, for which I am more grateful than I could hope to express on these pages.

I wish to thank Studienstiftung des deutschen Volkes and the German Academic Exchange Service (DAAD) for financial and material support, as well as Merton College, Oxford for generous travel grants.

Abstract

In this thesis, the response of DNA and RNA to linear and torsional mechanical stress is studied using coarse-grained models. Inspired by single-molecule assays developed over the last two decades, the end-to-end extension, buckling and torque response behaviour of the stressed molecules is probed under conditions similar to experimentally used setups. Direct comparison with experimental data yields excellent agreement for many conditions. Results from coarse-grained simulations are also compared to the predictions of continuum models of linear polymer elasticity.

A state diagram for supercoiled DNA as a function of twist and tension is determined. A novel confomational state of mechanically stressed DNA is proposed, consisting of a plectonemic structure with a denaturation bubble localized in its end-loop. The interconversion between this novel state and other, known structural motifs of supercoiled DNA is studied in detail. In particular, the influence of sequence properties on the novel state is investigated. Several possible implications for supercoiled DNA structures *in vivo* are discussed. Furthermore, the dynamical consequences of coupled denaturation and writhing are studied, and used to explain observations from recent single molecule experiments of DNA strand dynamics.

Finally, the denaturation behaviour, topology and dynamics of short DNA minicircles is studies using coarse-grained simulations. Long-range interactions in the denaturation behaviour of the system are observed. These are induced by the topology of the system, and are consistent with results from recent molecular imaging studies. The results from coarse-grained simulations are related to modelling of the same system in all-atom simulations and a local denaturation model of DNA, yielding insight into the applicability of these different modelling approaches to study different processes in nucleic acids.

Contents

List of Figures

List of Tables

Publications

This thesis is based on work contained in the following publications:

C. Matek, T. E. Ouldridge, A. Levy, J. P. K. Doye, and A. A. Louis

DNA cruciform arms nucleate through a correlated but asynchronous cooperative mechanism.

J. Phys. Chem. B, 116(38):11616 – 11625, 2012.

C. Matek, T. E. Ouldridge, J. P. K. Doye, and A. A. Louis

Plectoneme tip bubbles: Coupled denaturation and writhing in supercoiled DNA.

submitted, 2014

T. Sutthibutpong, C. Matek, C. Benham, G. Gouvea-Slade, A. Noy, C. Laughton, E. Dringo, J. P. K. Doye, A. A. Louis and S. A. Harris.

Sequence Dependent Denaturation in Small DNA Circles: A Combined Theoretical and Computational Study.

to be submitted, 2014

J. P. K. Doye, T. E. Ouldridge, A. A. Louis, F. Romano, P. Sulc, C. Matek, B. E. K. Snodin, L. Rovigatti, J. S. Schreck, R. M. Harrison, and W. P. J.Smith.

Coarse-graining DNA for simulations of DNA nanotechnology.

Phys. Chem. Chem. Phys., 15:20395 – 20414, 2013.

Chapter 1

Introduction

In this thesis, coarse-grained models of DNA and RNA are used to study the behaviour of these molecules when they are subjected to mechanical stresses such as supercoiling and tension. The response of nucleic acids to such external stresses is not only important to study their basic physical properties and understand the results of single molecule experiments, but also has implications for the behaviour of these molecules in a biological context. The thesis is structured as follows:
In chapter 1, basic features of nucleic acids and their biological implications will be introduced, and a survey of recently developed experimental assays that permit study of their properties on a single-molecule level will be given. The coarse-grained models and a simple linear elastic model, as well as methods to determine the topology of nucleic acids will be introduced in chapter 2. The basic behaviour of nucleic acids under external twist and tension as probed in the context of typical single-molecule experiments will be discussed in chapter 3. In chapter 4, a novel conformational state of DNA will be proposed, and its implication on the large-scale configuration of DNA double strands explored. Experimental probes of the novel state are suggested, and possible biological implications discussed. Chapter 5 is concerned with the dynamical and displacement properties of plectoneme structures. Different modes of displacement will be described and compared to experimental findings. Finally, chapter 6 describes the combined denaturation and writhing behaviour of short DNA minicircles.

1.1 Nucleic acids: Chemical and structural properties

The recognition of nucleic acids as information-storing molecules is one of the cornerstones of modern biology. Historically, work on the constituents of the cell nucleus dates back to the 19$^{\text{th}}$ century, when in 1869 the physiologist Miescher first isolated a substance from the nuclei of human white blood cells, and called it "nuclein" [7]. Studying its chemical composition, he found that it was rich in phosphate, but contained no sulfur, concluding that the substance was not a protein. Building on this early work, Kossel could show that nuclein contained four nucleobases and sugar [8]. Later, Levene identified the structure of nucleotides as a unit consisting of sugar, phosphate and a base [9] (see Fig. 1.1(a)), and recognised DNA and RNA as two different forms of nucleic acids. However, he hypothesized that DNA contained equal amounts of adenine (A), guanine (G), cytosine (C), and thymine (T), and thus could not carry information. This assumption was falsified by Chargaff and coworkers, who discovered that the relative content of A and T as well as that of G and C was equal in DNA [10]. The advent of X-ray diffraction techniques enabled more detailed structural studies of nucleic acids, which were pioneered by the group of Astbury in the 1930s [11]. While these studies hinted at a periodic organization of the DNA molecule, an interpretation of the diffraction data became possible only when Franklin found that DNA existed in two allomorphs, the 'A' and the 'B' form [12] (see Fig. 1.1(b)). Building on these results, research in the molecular structure of DNA culminated in 1953, when Watson and Crick presented their model of DNA as a right-handed double helix of nucleotides (Fig. 1.1(b)) [13]. While the individual nucleotides on a single strand are linked via covalent bonds between their ribose and phosphate groups, canonical base-pairing is mediated by two hydrogen bonds in A-T base pairs and three hydrogen bonds in G-C base pairs. Fig. 1.1 shows the chemical structure of the canonical base pairs A-T and G-C, which are the most important motifs underlying the secondary structure of the DNA molecule [1]. They

[1] In addition to these canonical motifs, other binding patterns such as Hoogsteen base pairs are known [16, 17].

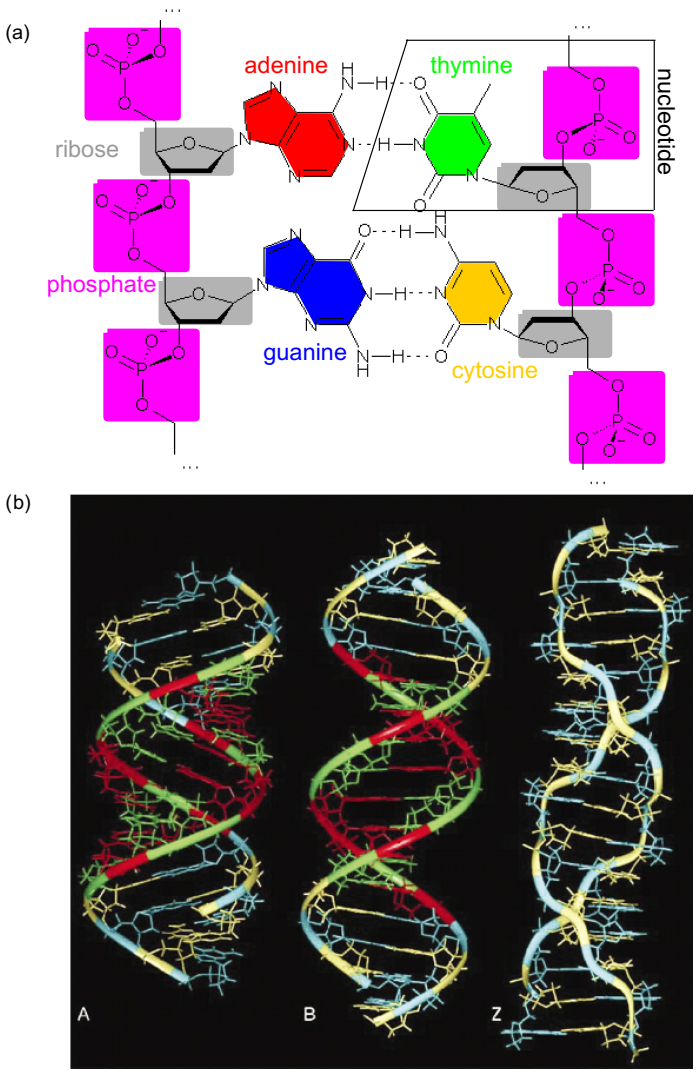


Figure 1.1: Molecular structure of DNA. Adenine residues are colored in red, thymine in green, guanine in blue and cytosine in yellow. (a) Chemical structure of the DNA double strand, showing the most common secondary structure motifs, an AT and a GC base pair. Covalent bonds are represented as solid lines, and hydrogen bonds as dashed lines. Figure adapted from Ref. [14]. (b) Stick-and-ball representation of the most common DNA tertiary structures, A-DNA, B-DNA and Z-DNA. Figure reproduced with permission from Ref. [15].

provide the molecular basis for the semi-conservative replication mechanism of DNA in biological systems to be described in Sec.1.2. The three most common double helical tertiary structures of DNA in a biological context are shown in Fig. 1.1(b). B-DNA is the prevailing configuration *in vivo*, while Z-DNA is assumed to have a biological function in particular in the presence of high supercoiling, which can occur, for example, in transcription [18], and the biological relevance of A-DNA is

still unclear[2].

In RNA, the second carbon in the ribose ring is linked to a hydroxyl group instead of a single hydrogen, and the base thymine (T) is replaced by uracil (U). In spite of the very similar chemical structure of DNA and RNA, the latter possesses a much richer structural repertoire [17], and a more diverse biological function, which will be briefly described in chapter 1.2. Most biologically occurring RNA molecules are single-stranded, but they can also occur in a predominantly A-helical double stranded form [17, 19]. Furthermore, RNA-DNA hybrid structures, occurring e.g. during transcription, are thought to mainly adopt an A-helical conformation [20, 21].

1.2 Key biological functions of nucleic acids

1.2.1 Inheritance and gene expression

Although theories of inheritance have existed since ancient times [22], only the work of Mendel on pea plants established a rigorous foundation of genetics [23]. The role of nucleic acids for inheritance was first recognised due to experiments by Avery *et al.* [24], who found that DNA was responsible for the induction of bacterial virulence. Motivated by the theoretical expectation that the carrier of biological information should be an "aperiodic crystal", as formulated by Schrödinger [25], Watson and Crick immediately realised the significance of the molecular structure of DNA for a copying mechanism of the genetic material [13]. Due to the specificity of base pairing described in Sec. 1.1, both strands carry complementary versions of the full base sequence. This allows for the copying of DNA by a semiconservative replication mechanism: The two single strands are separated, and each is used as a blue print to reconstruct the respective complementary strand, a task performed by DNA polymerases *in vivo* [26].

The DNA sequence stores the information necessary to produce proteins, which are of paramount importance for the structural and functional organization of the

[2]Next to these most important configurations, a large number of other tertiary structures have been reported. A comprehensive survey of these can be found in Ref. [15].

cell. During gene expression, the base sequence of DNA is accessed by RNA polymerase, an enzyme that transcribes the genetic information into a copy of messenger-RNA molecules, which itself are translated into the amino acid sequences constituting proteins by ribosomes [26]. Apart from being a mediator between DNA and ribosomes, RNA is also known to fulfill a large number of regulatory roles in the cell [27].

While the storage of genetic information for protein synthesis is the key function of DNA[3], the biological importance of non-coding sequences, which are not involved in protein synthesis, has recently been increasingly appreciated [29]. At the same time, the role of chemical modifications of the DNA molecule by the cellular environment, such as methylation, in sequence readout has been recognised [30]. These developments suggest a more richly faceted picture of biological DNA as an active constituent embedded in the complex environment of the cell, rather than a passive repository of genetic information [31, 32]. To understand this dynamical role of nucleic acids *in vivo*, a one-dimensional description of their sequence has to be extended to include the physical context in which the sequence is stored, namely the polymeric properties of DNA and RNA molecules. The importance of these physical properties of nucleic acids, and their interplay with the encoded sequence, is a guiding theme of many results presented in this thesis.

1.2.2 DNA topology

Immediately after the discovery of the double-helical structure of DNA, its implications for readout and replication mechanisms were considered [33, 34]. An early problem was the question how strand separation could proceed over long distances given the highly intertwined structure of the double helix. It was expected that melting of the double strand over large regions would torsionally stress adjacent parts of the strand, and cause the system to bend back on itself, building up mechanical resistance against further progression of the replication fork[4] [35]. These

[3]The unidirectional information flow from DNA to proteins via RNA has been characterized as the "Central Dogma of Molecular Biology" by Crick [28].

[4]This problem was widely discussed as the "DNA replication problem" in the 1950s [34].

early theoretical considerations lead to a first recognition of the importance of topological properties of DNA. Further evidence for the relevance of supercoiling in DNA came from experiments by VINOGRAD and collaborators, which suggested that the genome of the Polyoma virus possessed a "twisted circular" microscopic configuration [36], allowing for the existence of topoisomers, which only differ in topological properties of their DNA strand, but not in their sequences.

Closed circular genomes were found to exist in an underwound state in most bacterial species (see Fig. 1.2). In eukaryotes, association of the genomic DNA with the nuclear matrix forms topological domains, within which supercoiling is conserved [37]. Much subsequent work [37, 38] established that topological properties considerably influence the biological activity of the sequence, and can control many key processes that involve separation of the DNA single strands, such as transcription, replication and recombination. Not only can supercoiling affect the stability of the double helix, and thus determine the free energy barriers that need to be overcome to access the encoded sequence. It is also important for organising the genome into looped domains (see Fig. 1.2), and therefore aids compaction of the DNA molecule. Another indication of the biological importance of topological properties of DNA is the existence of topoisomerases, a highly conserved enzyme class whose function it is to regulate the interconversion between different topoisomers of DNA by over- or underwinding the double strand [38].

1.3 The single molecule perspective

Early results on DNA structure and function were obtained using traditional biochemical bulk assays, such as gel electrophoresis, optical absorption studies, electron microscopy, or X-ray diffraction. These experimental techniques infer the average behaviour of an ensemble of molecules, and give only indirect information on individual constituents of the population. Due to ensemble averaging, it can be difficult to obtain kinetic information on the molecule studied, which is of particular interest for molecules which, such as DNA, occur in small copy numbers *in vivo*. Moreover, the

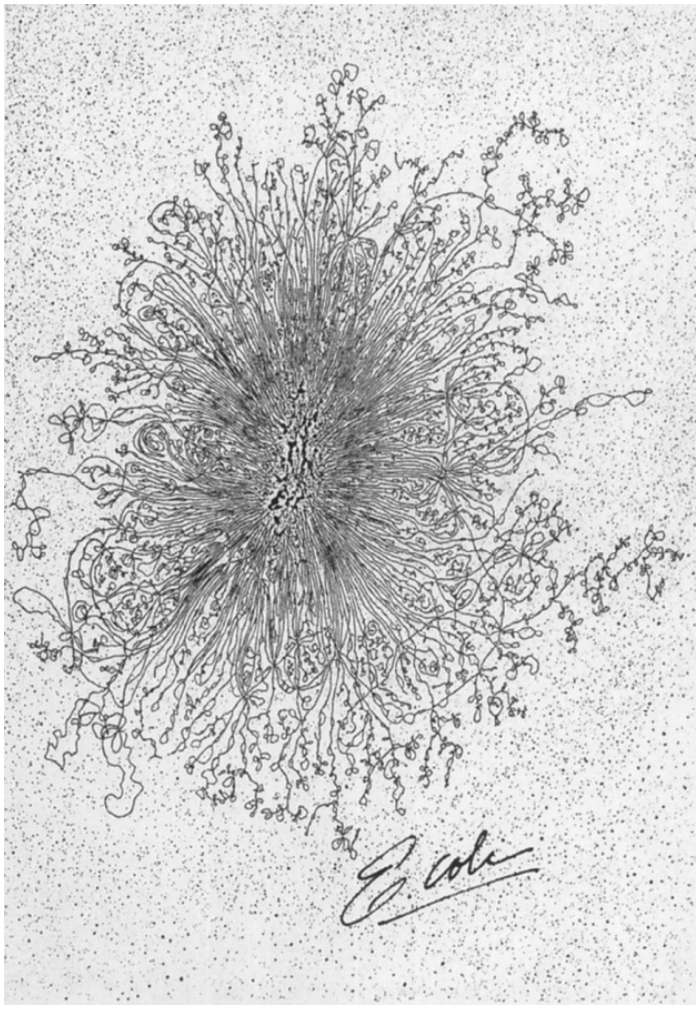


Figure 1.2: Electron micrograph of the closed circular *E. coli* chromosome of length 4.7Mbp, showing multiple supercoiled loops emerging from a central region. Bacterial genomes are typically underwound by about 6% *in vivo* [38]. Figure taken from Ref. [39].

difference between static and dynamical disorder in a molecular population cannot be extracted from the bulk, which is often key to understanding the conformational dynamics of a biomolecule [40].

By contrast, single molecule experiments provide a much more detailed picture of the constituents of the system studied. They give access to studying statistical distributions rather than just average values for a given property of a molecule. Apart from extracting information on structure and dynamics of their microscopic state, single molecule assays also allow to actively manipulate the molecules studied, thus making it possible to directly determine mechanical properties of individual

biopolymers *in vitro*.

1.3.1 Atomic Force Microscopy

Atomic Force Microscopy (AFM) was originally developed as a method to probe the atomistic structure of surfaces, but it has been applied to the study of DNA from its early stages of development [41]. The method is based on a cantilever with an ultra-thin tip that is scanned over a probe. Its movement is detected by the deflection of a laser beam. This allows one to image single molecules of DNA or RNA on a sub-nanometer scale [40, 42]. Due to the high spatial resolution of the technique, defects of the DNA double strand such as denaturation bubbles [43, 44] or cruciforms [45], as well as recruitment of DNA-binding proteins such as RecA [46] can be visualized directly. At the same time, AFM can also be used for force spectroscopy of single molecules [40, 47]. A drawback of most AFM setups is the relatively long image acquisition time necessitated by the scanning process, although recently time-resolution has been considerably improved [42]. Additionally, AFM studies can be performed in solution only at the expense of a worse temporal resolution [40]. For studies performed on molecules attached to a surface, results have to be interpreted with attention to possible artifacts introduced by the deposition process [43].

1.3.2 Molecular tweezers

Since their development and first appication to the study of biomolecules in the early 1990s, molecular tweezers have developed into a key tool for studying the mechanical response of DNA to external forces or torques. The most common basic designs are magnetic and optical tweezers, although many variants and combinations of both techniques have been developed for particular applications [48, 49, 50, 51].

Optical Tweezers

In optical tweezer assays, DNA is attached to dielectric beads, which are then trapped in a highly focussed laser beam, as shown in Fig. 1.3 (b). Beads and

trap positions are monitored, and the force determined from the displacement from the trap center. Shifting the focal point of the lasers then allows direct positioning of the molecule ends, imposing a constant end-to-end distance, or a constant force by including a feedback mechanism on the trap position [40, 48]. In their native form, optical tweezers are best fit to exert linear forces, and have been used to probe the stretching behaviour of DNA [52, 53], or the disruption of secondary structure elements such as hairpins in DNA [54] and RNA [55], as well as the denaturation behaviour of RNA [56]. Recently, extensions of the original setup have been developed to angularly trap optical beads [57], and directly measure torque exerted on the molecule [58], which complement the *magnetic* tweezer assay as an experimental technique to measure the response of nucleic acids to force and torsion.

Magnetic Tweezers

The basic setup of a magnetic tweezer assay is shown in Fig. 1.3 (a). A molecule is attached to a coverslip with one of its ends, and to a magnetic bead with the other using antibody-antigen tethers. The magnetic bead can be manipulated by magnets situated above the molecule. In this way, the stretching force on the molecule can be set by varying the distance between magnets and bead, while the imposed twist can be controlled by rotating the magnets. The diffraction pattern of the bead is monitored, from which the distance between bead and coverslip can be determined. The force acting on the bead is inferred from its transversal fluctuations [59]. Magnetic tweezers were first introduced by Strick *et al.* to study the elastic behaviour of supercoiled phage λ DNA [60, 61, 62]. They have since been used by different groups to probe the mechanical response of DNA to supercoiling for different strand lengths, salt concentrations and pH conditions [6, 4, 3, 2, 63]. Of particular interest for the topic of this thesis are extensions of the classical magnetic tweezer setup that allow direct measurement of the twist and torque response of the strand by angular bead tracking, as they permit direct comparison to theoretical results [64, 65, 66, 5, 67].

In addition to accessing the mechanical properties of dsDNA, magnetic tweezers have proven a versatile tool for studying a large class of systems that involve the change of local twist on a nucleic acid molecule. Intrinsic structural transitions of the DNA strand addressed with magnetic tweezers include the kinetics of cruciform formation [68], the transition of B-DNA to Z and L forms [67], and the denaturation behaviour of the double strand [3, 2, 63]. Double stranded RNA has been probed in magnetic tweezers as well [69], but only very recently Lipfert and coworkers have performed a systematic investigation of its mechanical properties [70].
Apart from studying the intrinsic behaviour of nucleic acids, magnetic tweezers have also been applied to microscopically investigate reactions with nucleic acids processing enzymes, such as in mechanisms of transcription [71], or the action of helicases [72] and topoisomerases [73, 74].

Although they have proven to be versatile tools for the study of many nucleic acid systems, molecular tweezer assays have some obvious limitations. In itself, a molecular tweezer assay can only provide information on the end-to-end extension of the molecule, which is a global property of the molecule. Local structural modifications that do not considerably influence end-to-end distance are difficult to study. Furthermore, the achievable time resolution may be limited by the inertia of the molecule attachments, or viscous drag by the solvent. Therefore, molecular tweezer data have to be complemented with results obtained using other experimental techniques, or interpreted using theoretical models to extract information on the local microscopic behaviour of the system. Experimentally, fluorescence-based probes can provide more detailed spatial information about the molecule.

1.3.3 Fluorescent techniques

Over the last two decades, single molecule fluorescent techniques have been widely applied to study conformations of nucleic acids in isolation or when interacting with other biomolecular species [76, 40, 51]. As opposed to molecular tweezers, these assays provide spatially resolved information on molecule configurations.

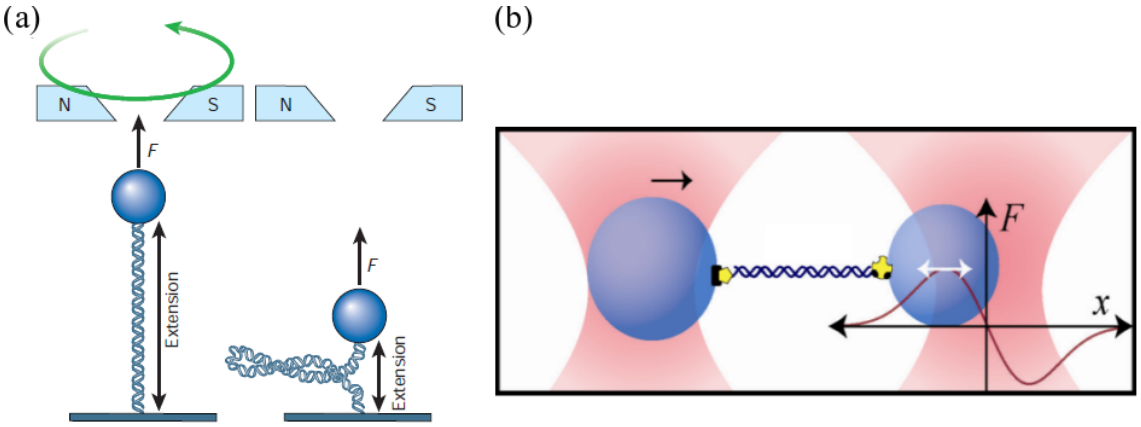


Figure 1.3: Schematic depiction of the basic design of molecular tweezer experiments. (a) In a magnetic tweezer, a DNA molecule is attached to a coverslip and a magnetic bead, which is used to manipulate the force and torsion on the molecule. Figure taken with permission form Ref. [48]. (b) In optical tweezers, DNA is attached to beads which are trapped by a highly focused laser beam, allowing to stretch the molecule. Figure from Ref. [75].

Fluorescence microscopy

Fluorescence microscopy is a common method used to monitor the overall molecular conformation of individual DNA strands, and can be combined with single-molecule manipulation techniques able to exert force or torque, such as molecular tweezers described in Sec. 1.3.2. In typical experimental setups, the molecule is marked by either covalently linked or intercalating dye, which is then excited by a light source, and monitored microscopically. This allows tracking the global configuration of the entire molecule in a time-resolved manner, and hence one can follow both its intrinsic dynamics, as well as the kinetics of reactions it is involved in. Recent applications to probe the behaviour of nucleic acids include studies on the local denaturation behaviour of DNA upon overstretching [77, 78], and the dynamics of supercoiled DNA molecules [1]. Fluorescent labelling has also been used to microscopically monitor protein-DNA interactions at high spatial resolution, such as the kinetics of RecA assembly on DNA [79] studied by the Kowalczykowski lab, or the detailed dynamics of the bacterial replication machinery recently investigated by van Oijen

and collaborators [80, 81].

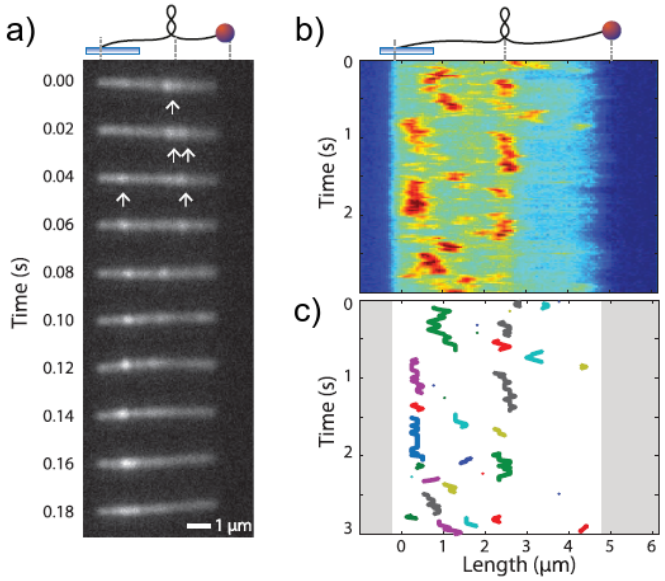


Figure 1.4: Plectoneme dynamics observed in experiments by VAN LOENHOUT *et al.* [1]. A DNA strand of length 21kbp was fluorescently labelled by Cy3 and held under twist and tension using a magnetic tweezer. (a) Real-time microscopic dynamics of DNA plectonemes. (b) Dye density kymograph of the trajectory in (a). (c) Inferred plectoneme displacement behaviour. Figure reproduced with permission from Ref. [1].

FRET

Single molecule FÖRSTER Resonance Energy Transfer (sm-FRET) [76] is another important fluorescence-based technique, and allows measurements of distances on the molecular scale with a nanometer accuracy. This local geometric information can be used to probe the dynamics of conformational changes in detail, and reveal the structure of particular microscopic states. For DNA, HA and colleagues investigated a number of systems using the sm-FRET method, such as the cyclisation properties of short DNA strands [82] and the structure of human telomeric DNA [83]. Also FRET can be combined with single-molecule manipulation techniques [76, 49, 84, 51]. These combined assays have been used to study systems such as

four-way junctions under stretching force [85], or the energetics of the B-Z transition in DNA [86]. Cordes *et al.* have used FRET techniques to study the denaturation of the DNA double strand during transcription initiation in the *lac* promoter [87]. Other applications include the study of DNA bending upon HMGB protein binding [88], or the probing of the reaction pathway of RNA enzymes [89], and detailed investigations of various nucleic acid translocases [90]. The FRET technique can be extended to monitor multiple fluorescent interactions at the same time [86, 91], allowing one to concurrently study several structural changes in the probed molecule, thus providing even more detailed local information on its conformation.

Next to being widely applied in the context of *in vitro* studies, fluorescent techniques hold the potential to be used in living cells [92]. Taken together, the availability of experimental techniques to both manipulate nucleic acids on a single molecule scale and at the same time monitor nanoscopic changes in their molecular configuration holds great promise for testing predictions made on the basis of theoretical and computational results as presented in this thesis, as well as to guide progress towards more detailed and physically informed models.

1.4 Modelling Nucleic Acids under twist and tension

As discussed in Sec. 1.3, single molecule experiments provide a growing set of tools to probe the mechanical behaviour of individual nucleic acids, offering a great potential for elucidating the details of many important biological processes.
At the same time, all individual assays are subject to limitations in terms of spatial and temporal resolution, the possibility to manipulate individual molecules, and the external experimental conditions required in a specific setup. Theoretical models provide the opportunity to unify results from different experimental approaches, and elucidate the physical origins of the observed behaviour. In particular, models can be of use to obtain a quantitatively accurate description of the molecular structures

underlying the mechanical response of single molecules to twist and tension that is being tested in experiments. Knowledge of the microscopic behaviour of molecules under external stress can then suggest mechanisms for particular biochemical reactions involving these molecules. Obtaining a complementary microscopic picture is of particular importance for interpreting the results of molecular tweezer experiments discussed in Sec. 1.3.2, which accurately probe the mechanics of a single molecule, but resolve its spatial configuration if at all only on a relatively coarse scale. For typical mechanical conditions set in single-molecule experiments and relevant *in vivo*, formation of double-helical coils, so-called plectonemes[5] and double strand denaturation represent the most common structural responses. A number of modelling approaches have been developed to describe these responses.

1.4.1 Atomistic models

Atomistic models, such as the AMBER [93] and CHARMM [94] force fields, provide a representation of nucleic acid molecules at the highest level of microscopic detail [95, 96]. From simulations at this resolution, precise structural information on the nucleic acids studied can be obtained, which can be relevant for a detailed mechanistic understanding of the interaction of DNA with other molecular species. For example, systematic studies of sequence-dependent structural properties of short B-helical DNA strands represented in the AMBER force field have been performed by the Ascona B-DNA Consortium [97]. Other atomistic work has investigated the elastic and chiral properties of DNA on the level of few base pairs, and considered the effects of explicit solvent properties [98, 99]. However, the main drawback of atomistic simulations is their high cost in terms of computational resources, although methods have been developed in recent years aimed at increasing the sampling efficiency [100, 101]. Typical DNA structures simulated on the atomistic scale are considerably shorter than one persistence length, and time scales can only be addressed up to the order of few μs with currently available computers [102]. This limits the mechanical response effects that can be reproduced *in silico* with atomistic models to fast modes,

[5] derived from Greek πλέκτος "twist, coil", and νῆμα "thread, yarn"

and does not allow the determination of molecular properties that reach equilibrium on a longer time scale. Comparing to non-equilibrium properties determined in experiment can also be nontrivial, as the simulated loading rates typically lie orders of magnitude above the experimental ones[6] [104].

Atomistic simulations of local structural deformations upon twist stress [105, 106] and linear stretching forces [107, 108, 104] have been performed in B-DNA of length 12bp. Similarly, the supercoiling and denaturation behaviour of DNA minicircles has been studied in atomistic simulations, and revealed a rich microscopic behaviour [109, 110]. These systems represent a case in which models on several degrees of resolution are available, and the study of a similar system on a coarse-grained level is the subject of chapter 6 of this thesis. In spite of these recent developments, physical processes dominating the mechanical behaviour of longer molecules as studied in single-molecule experiments, such as buckling or the thermodynamics of double strand denaturation, remain outside the purview of atomistic simulations.

1.4.2 Continuum models

Continuum models represent nucleic acids as media without structure on the microscopic level, and are therefore on the opposite end of atomistic simulations in terms of resolution. Most continuum models focus on the behaviour of DNA as a semiflexible polymer, and ignore the physics of the molecule on length scales smaller than its persistence length. This very coarse level makes continuum models computationally very efficient. The most popular continuum model to describe the behaviour of double stranded DNA and RNA is the worm-like chain (WLC) model first introduced by Kratky and Porod [111], which describes the molecule as an isotropic rod with continuous linear bending flexibility. In order to capture the behaviour of DNA under typical experimental conditions, the WLC has been extended to capture not only the bending, but also the stretching [112] and torsional [113] elasticity.

[6]Recently, this time scale gap has been narrowed from the experimental side by Rico and collaborators, using high-speed force microscopy [103].

With these physical ingredients, the model remains analytically tractable at least in limiting cases, and many good approximations exist for the extension [112, 114, 115] and torsional [113, 116] response behaviour that allow for a direct comparison to single molecule measurements. In particular, Marko, Neukirch and co-workers have used results from this class of continuum model to very successfully describe the phase-behaviour of DNA under twist and tension [117, 118, 119, 120]. Continuum models have also been adapted to include the effects of solution electrostatics [121, 120].

In spite of yielding important insights into the behaviour of DNA on long scales, the ability of continuum models to capture the structural deformations, or consequences of the chiral structure of the molecule is intrinsically limited. Although extensions exist to capture the chiral effect of twist-stretch coupling [122], non-B-DNA phases [119] or non-homogeneous stiffness [123], these additions do not emerge naturally from the continuum description of the polymer, and have to be parametrized using results from other, more microscopic modelling approaches. They also cannot give a realistic account of strand geometry at the level of individual nucleotides, which is the scale most relevant for biochemical reactions that involve access to the genetic sequence.

Therefore, in order to capture both the large-scale behaviour of nucleic acids as semi-flexible polymers, and its coupling to local properties set by the base sequence and relevant to biological processes, a description is necessary that naturally retains the structural and thermodynamic properties of a chiral helical molecule, and is at the same time efficient enough to access systems large enough for comparison with experimental systems. Such intermediate-resolution representations belong to the domain of coarse-grained models.

1.4.3 Coarse-grained models

Coarse-grained models retain only a limited number of effective degrees of freedom compared to atomistic models, and seek to model the physical behaviour of a molecule in terms of this reduced set of variables. At the expense of microscopic

resolution, this strategy gives access to phenomena at a larger time and length scale than fully microscopic models, while still providing more detail than continuum models. Coarse-graining strategies therefore always reflect a trade-off between detail and efficiency, which has to be evaluated with regard to the particular problem addressed. To study the mechanical response behaviour of DNA under mechanical stresses, different coarse-grained models have been developed that address the denaturation, twist and bending response of DNA to supercoiling.

For the study of double-helix denaturation, several models in reduced dimensionality are available that predict the probabilities of double strand melting [124]. Among the most common of these models is the one introduced by Poland and Scheraga [125], a one-dimensional statistical mechanical model which represents the melting transition in polynucleotides by imposing an entropy contribution to the free energy of denatured regions. It has found wide application in the study of critical exponents of the melting transition [126, 127, 124, 128], and the statistical mechanics of denatured regions, so called denaturation "bubbles" [129, 130].

For oligonucleotides, the nearest-neighbour model by SantaLucia and coworkers, in which a free energy model is fit to a large set of denaturation data obtained under a wide range of experimental conditions, has been very successfully applied to study the sequence-dependent melting transition [131].

A very successful model of the calculation of denaturation profiles in DNA strands is the model developed by Benham and co-workers [132, 133, 134, 135]. Just like the Poland-Scheraga and the SantaLucia models, it describes denaturation on the level of two discrete states, but additionally takes into account explicit energy contributions from single-strand twisting of denatured regions, and residual superhelical stress in non-denatured parts of the strand. Efficient approximation methods exist to apply the model for long sequences on the genomic level [136].

As opposed to two-state models, the Peyrard-Bishop-Dauxois (PBD) model represents base-pair denaturation using a continuous coordinate. The PBD model has been extended to capture the response of the double strand to external torque and overstretching [137, 138, 139], though bending modes are still neglected [138].

All these effective models of DNA melting are geared towards reproducing local denaturation properties, in terms of melting of a particular base pair or short stretch. However, none of them has a detailed enough three-dimensional representation of the nucleic acid molecule to at the same time include the geometry of double strand bending. Thus, none of these approaches is powerful enough to permit a comprehensive picture of the effects of supercoiling in single molecules.

Coarse-grained models have also been developed to study the bending and writhing behaviour of supercoiled DNA. Approaches primarily concerned with the overall geometry of long double strands have chosen a representation on the basis of stiff segments with torsional rigidity, which are typically discretised variants of the WLC model [140, 141, 142, 143, 144, 145], or elastic ribbon models [146, 147]. While those models represent overall topological and geometrical features of the supercoiled structures and can be extended to study the influence of electrostatic interactions or external forces relatively easily, they are too coarse to permit study of local structural effects. Such a comparison requires coarse-graining on the level of individual nucleotides.

Coarse-grained models that permit the study of an interplay between strand writhing and denaturation have been proposed in the literature[7]. Mielke *et al.* introduced a coarse-grained model that absorbs three nucleotides into one coarse-grained site [150], and studied the out-of-equilibrium response of 141 bp minicircles to rapid changes in supercoiling. However, this modelling approach was limited to a non-equilibrium scenario, as it only allows breaking but not re-formation of base pairs, and is therefore of only limited thermodynamic consistency, although it could observe transient coupling of writhing with weak parts of the modelled sequence [150]. The model developed by Trovato and Tozzini [151] was used to study the response of minicircles to supercoiling and observed denaturations, but it does not include sequence-dependent interaction strengths and can therefore only reproduce an average-base denaturation behaviour. A model by Sayar and

[7] A more detailed comparison of existing coarse-graining approaches to DNA can be found in Refs. [148] and [149].

collaborators was used to study the twist-writhe partitioning in minicircles, but was not parametrised to reproduce the thermodynamics of melting, and therefore did not observe base pair breaking at all during the simulations performed. All of these studies were performed for small, closed-circular systems, so that the effect of stretching forces could not be directly observed. No systematic study of nucleic acids under torsional and extensional stress has been performed so far using a model that can at the same time capture the single-nucleotide behaviour of these molecules.

Chapter 2

Methods

Coarse-grained models of nucleic acids which provide a nucleotide-level representation of the molecule and include consistent thermodynamics of melting are a prerequisite for studying the effects of sequence on the global configuration of DNA. At the same time, the spatial and temporal scales of relevance are only accessible when the methods used to study the models' properties are sufficiently efficient. The models used for the studies reported in this thesis will be briefly introduced in Sec. 2.2. Furthermore, the molecular dynamics techniques used to simulate the models are discussed in Sec. 2.3.

When resolution and system size conditions are fulfilled, it is possible to directly compare the results from the coarse-grained models to predictions of continuum theories. Elastic properties of nucleic acids observed in experiment are normally reported in terms of the parameters of a such continuum models. Global, topological properties of supercoiled polymers are customarily quantified using concepts from differential geometry. In Sec. 2.1, a linear elastic model is presented, which defines the most important parameters measured in simulations. Concepts to describe the topology of the simulated molecules are also introduced.

Finally, in Sec. 2.4, the boundary conditions used for the simulations are outlined, and their relation to boundary conditions in experimental assays discussed.

2.1 Quantifying properties of semiflexible polymers

Generically, semi-flexible polymers can be defined as polymers which are characterised by a finite persistence length, which is large enough to be comparable with other relevant lengths scales of the system. As double stranded nucleic acids are twist-storing polymers, they have persistence lengths associated not only with bending, but also with twist modes. Moreover, the molecules are extensible by linear force, and therefore posess a finite YOUNG'S modulus. The interplay between these different length scales generates the rich configurational behaviour observed in supercoiled DNA.

2.1.1 Elasticity

Consider the configuration of DNA given by a space curve $\mathcal{C}$ consisting of small segments ds. Then, local elastic stresses can be described to linear order by three quantities [152]:

- Changes in the local direction of the strand result in **bending** deformations, which are proportional to the local change rate of the tangent vector $\hat{\mathbf{t}}$, $\boldsymbol{\beta} = d\hat{\mathbf{t}}/ds$

- Deviations of the twist angle per base pair $\Delta\phi$ from the equilibrium configuration, $\omega = \Delta\phi/r_0 - \omega_0$, result in **torsional** deformations. Here, the relaxed local rotation of the strand is given by $\omega_0 = 2\pi/(p_0 r_0)$, $p_0 \approx 10.4$bp is the helical pitch of the relaxed DNA double helix and $r_0 \approx 0.34$nm/bp is its eqilibrium rise. Consequently, $\omega = 2\pi/(r_0 p_0)\ \sigma$, where σ is the dimensionless superhelical density (see Sec. 2.1.2.1) commonly used to quantify torsional stress.

- Relative length extensions lead to **stretching** deformations, and are measured by the local quantity $u = \Delta(ds)/ds$

Taken together, the effective elastic free energy of the system is then given by the following functional [152]:

$$\frac{F_{\text{elastic}}}{k_B T} = \frac{1}{2} \int_{\mathcal{C}} \left[A_0 \boldsymbol{\beta}^2 + B_0 u^2 + C_0 \omega^2 + 2Du\omega \right] ds \tag{2.1}$$

Here, the quantities A_0 and C_0 are the bend and twist persistence lengths respectively, and B_0 is the stretch modulus. The term proportional to D describes a coupling of the local twist of the molecule to its extension, which physically is a consequence of the anisotropy of the double helix. It has been found in experiments that $D < 0$ for DNA, which means that the stretched molecule will overwind, somewhat in contrast to macroscopic intuition [153, 154]. However, recent results indicate that $D > 0$ for double-stranded RNA [155, 156]. In the presence of external forces and torques, additional global terms are added to Eq. 2.1 to yield

$$F_{\text{tot}} = F_{\text{elastic}} - \mathbf{f} \cdot \hat{\mathbf{z}} - 2\pi\tau\Delta Lk, \tag{2.2}$$

where $\mathbf{f}$ is an external stretching force, $\hat{\mathbf{z}}$ is the global setup direction of the strand, τ is the applied torque and ΔLk the excess linking number of the strand, i.e. $\Delta Lk = 0$ for the relaxed double helix (see Sec. 2.1.2).

While the full elastic model defined in Eq. 2.2 has been shown to fit experimental results well over a large range of conditions, it is mathematically fairly complex [157, 122, 116]. The effect of twist-stretch coupling is small relative to the other effects included in the model, so that it can be neglected in most cases. In addition, the finite extensibility of the molecule only becomes significant for stretching forces $F \geq$ 10pN [122], and the behaviour of DNA can be described by an inextensible model for lower forces. The values of the bending and torsional persistence lengths A_0 and C_0 in Eq. 2.1 are bare microscopic quantities. By contrast, the values measured in experiments are renormalized by thermal fluctuations. For intermediate forces F small enough that $B_0 = 0$ can be assumed, but large enough that twist dominates over bending, the effect of bending fluctuations on the measured effective torsional persistence length can be determined by an approximate analytical calculation first

performed by Moroz and Nelson [113, 116]. The effective torsional stiffness of the molecule is given by

$$C_{\text{eff}} = C_0 \left[1 - \frac{C_0}{4A_0} \sqrt{\frac{k_B T}{A_0 F}} + \mathcal{O}\left(F^{-3/2}\right) \right]. \tag{2.3}$$

This relation allows the determination of the microscopic parameters A_0 and C_0 at the same time by fitting to the measured twist rigidity as a function of force. It has been used to interpret experimental results [6], and to construct a model of supercoiled DNA representing different microscopic conformations of mechanically stressed DNA as distinct phases [117].

2.1.2 Supercoiling

2.1.2.1 Definitions

Next to the energetics of elasticity associated with the deformation of the molecule, it is desirable to quantify the geometry of a molecule under torsional stress. Rigorous expressions exist to describe the spatial conformation of topologically closed systems such as circular DNA strands. The fundamental quantities associated with the topology of a system are twist (Tw) and writhe (Wr). For a system consisting of two interwound space curves $\mathbf{r}_1$ and $\mathbf{r}_2$, as the backbones of a closed-circular double strand, the intuitive meaning of these quantities is illustrated in Fig. 2.1: Tw equals the number of times both curves wrap around each other, while Wr represents the number of times the double strand axis crosses itself, averaged over all different aspect angles. Let $\mathbf{r}(s) = \left[\mathbf{r}_1(s) + \mathbf{r}_2(s)\right]/2$ be the centre line of the double strand, parametrised by the contour length s of the strand. Then, twist and writhe can be defined by [158]

$$Tw = \frac{1}{2\pi} \oint_0^L ds\, \hat{\mathbf{t}}(s) \cdot \left[\hat{\mathbf{u}} \times \frac{d\hat{\mathbf{u}}(s)}{ds} \right] \tag{2.4}$$

and

$$Wr = \frac{1}{4\pi} \oint_0^L ds \oint_0^L ds'\ \left[\hat{\mathbf{t}}(s) \times \hat{\mathbf{t}}(s')\right] \cdot \frac{\mathbf{r}(s) - \mathbf{r}(s')}{|\mathbf{r}(s) - \mathbf{r}(s')|^3}, \tag{2.5}$$

where L is the total strand length, $\hat{\mathbf{t}} = d\mathbf{r}/ds$ the tangent vector field, and $\hat{\mathbf{u}} = \left[\mathbf{r}_1(s) - \mathbf{r}_2(s)\right]/|\mathbf{r}_1(s) - \mathbf{r}_2(s)|$ is the vector field connecting both single strands. As

can be seen intuitively from Fig. 2.1, Tw and Wr can be interconverted for a given topologically closed configuration of the system. A rigorous formulation of this intuition is given by the FULLER-WHITE-CĂLUGĂREANU relation [159, 160, 161, 162]

$$Lk = Tw + Wr, \tag{2.6}$$

where the linking number Lk is a topological invariant[1] constrained to be integer in closed circular systems. Note that by definition, Tw is a local quantity, with the kernel of the integral in Eq. 2.4 representing a twist density, whereas the definition of Wr in Eq. 2.5 couples nonlocal contributions due to the double integral. For topologically open systems such as linear strands, Lk can take non-integer values. The definition of Tw can be applied in such systems, while the continuation of Wr to the open case brings about mathematical subtleties [164, 165]. It is however often possible to close the strand at infinity without significantly changing the value of Wr [166].

A torsionally relaxed double helix has an equilibrium linking number of

$$Lk_0 = \frac{N}{p_0}, \tag{2.7}$$

where N is the length of the molecule in bp and p_0 is the average equilibrium helix pitch, with $p_0 \approx 10.4$bp/turn being a typical value for DNA. The supercoiling of the double helix is therefore often given in terms of the excess linking number $\Delta Lk = Lk - Lk_0$. A common measure independent of the molecule length is the superhelical density σ defined as

$$\sigma = \frac{Lk - Lk_0}{Lk_0} = \frac{\Delta Lk}{Lk_0} = \frac{Lk}{Lk_0} - 1. \tag{2.8}$$

In this thesis, overall supercoiling is mostly given in terms of the superhelical density σ, as this quantity allows direct comparison to experimental results which are normally obtained for longer molecules.

[1] In the mathematical literature Lk is sometimes referred to as Invariant of GAUSS, who first introduced it to study the orbits of asteroids [163].

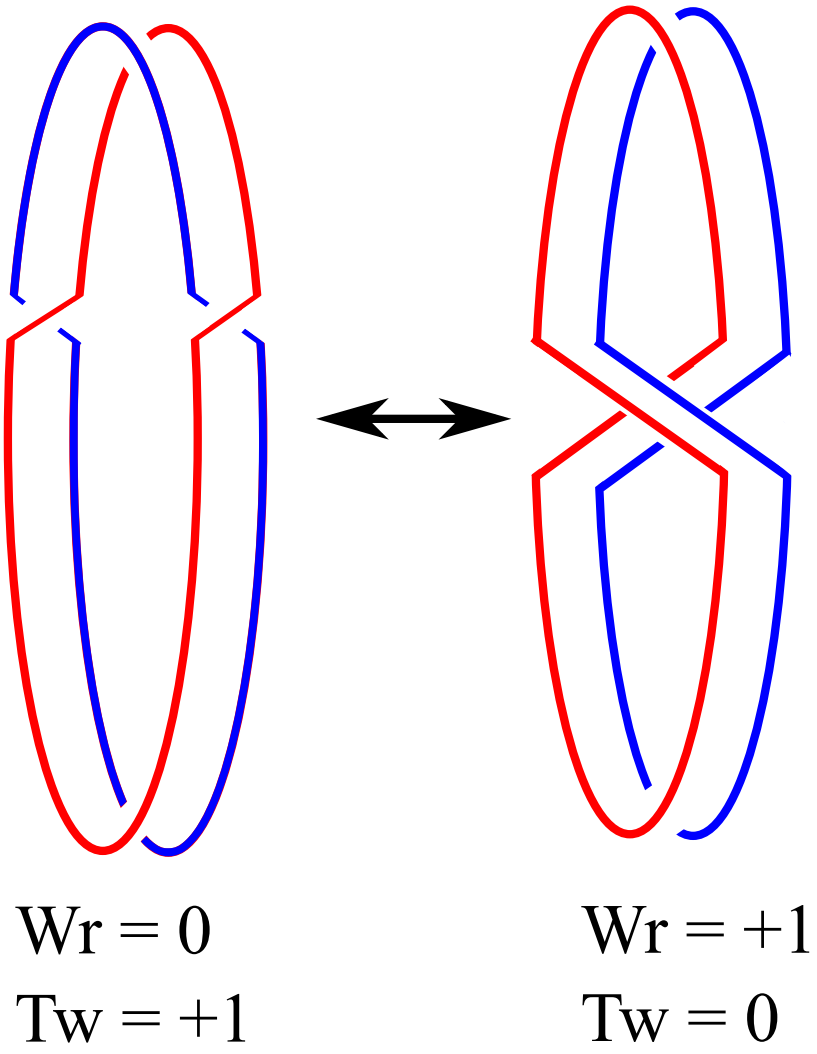


Figure 2.1: Schematic depiction of the interconversion of Twist and Writhe for a closed circular ribbon with conserved linking number $Lk = 1$.

2.1.2.2 Approximation schemes

While Eqs. 2.4 and 2.5 in Sec 2.1.2.1 define twist and writhe rigorously and ensure that Eq. 2.6 is fulfilled for all configurations, they are of limited use in practice. Firstly, they assume $\mathbf{r}_1$ and $\mathbf{r}_2$ to be smooth, twice differentiable space curves, which is only approximately true for the discrete curves obtained from a nucleotide-level model. Secondly, direct evaluation of the double integral in Eq. 2.5 can prove numerically difficult, both in terms of accuracy and computation time. Therefore, several approximations to the definitions of Eqs. 2.4 and 2.5 have been developed [167, 168, 169]. In this work, only one particular choice is considered, and

the subtleties connected to different approximation schemes are neglected. To check the consistency of the approximation methods used, the validity of the FULLER-WHITE-CĂLUGĂREANU relation (Eq. 2.6) is tested in applications on closed circular systems (see chapter 6).

Approximate measurement of Twist

For the work in this thesis, the following algorithm was used to determine the value of twist on a base-pair level (see Fig. 2.2 for illustration):

- Connect the centre of mass sites of corresponding nucleotides on both strands. This defines base pair vectors $\mathbf{r}_{bp,i}$.
- Define the basepair centres as the middle between two corresponding centre of mass sites.
- Project $\mathbf{r}_{bp,i}$ and $\mathbf{r}_{bp,j}$ of two consecutive base pairs into a plane perpendicular to the line connecting the base pair centres.
- Measure the local twist angle $\tau_{ij} = \arccos(r_{bp,i} \cdot r_{bp,j})$ as the angle between the projections of $\mathbf{r}_{bp,i}$ and $\mathbf{r}_{bp,j}$ in the plane.

The total twist of a structure is then given by

$$Tw = \frac{1}{2\pi} \sum_{\langle i,j \rangle} \tau_{ij}, \tag{2.9}$$

where $\langle i, j \rangle$ indicates the sum over nearest neighbours.

Approximate measurement of Writhe

Several approximate ways of efficiently calculating writhe have been proposed in the literature [167, 168, 170]. For the work presented in this thesis, only one approximate scheme for the calculation of writhe was used. This measure was only

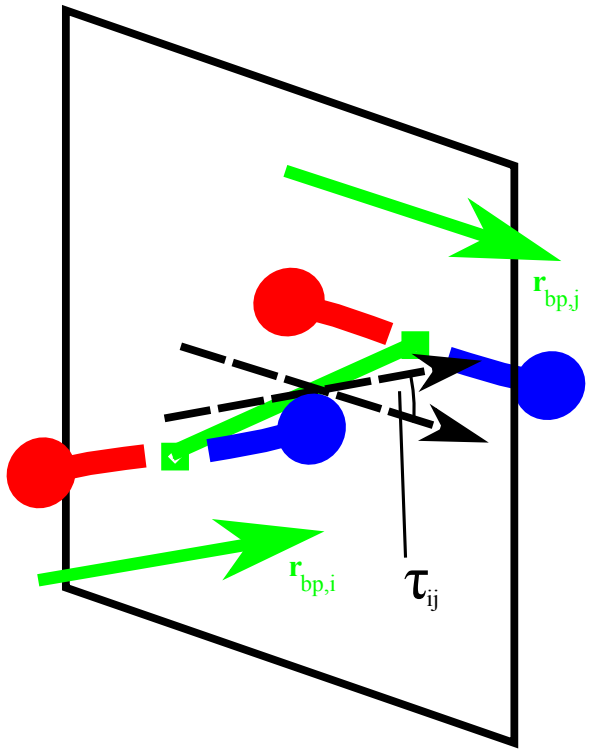


Figure 2.2: Measurement of a local twist angle: Consecutive base pair vectors are projected into the plane perpendicular to the line connecting the base pair centres. Dashed vectors indicate projections.

applied for closed-circular structures, where the quantity Wr is defined unambiguously (see chapter 6). To determine the value of writhe for a given closed-circular system, a discretization of Eq. 2.5 was used. Let $\mathbf{r}_{m,i}$ be the centre points of the base pairs of the system of length Nbp, where i is the index of a given base pair on the strand. Then, the tangent vectors $\hat{\mathbf{t}}(s)$ of the system can be approximated by

$$\hat{\mathbf{t}}_i = \frac{\mathbf{r}_{m,i+1} - \mathbf{r}_{m,i}}{|\mathbf{r}_{m,i+1} - \mathbf{r}_{m,i}|}. \tag{2.10}$$

With the $\mathbf{r}_{m,i}$ and $\hat{\mathbf{t}}_i$ defined in this way, Eq. 2.5 can be discretized as

$$Wr = \frac{1}{2\pi} \sum_{i=1}^{N} \sum_{j=1}^{j<i} [\hat{\mathbf{t}}_j \times \hat{\mathbf{t}}_i] \cdot \frac{\mathbf{r}_{m,j} - \mathbf{r}_{m,i}}{|\mathbf{r}_{m,j} - \mathbf{r}_{m,i}|^3}. \tag{2.11}$$

This discretization leads to numerically stable predictions for most configurations considered. However, for partially denatured systems, ambiguities can arise in the definition of the centre line in very rare cases. Treatment of these pathological configurations is described in more detail in Appendix B.

2.1.3 Structural properties

2.1.3.1 Denaturation

In several sections of this work, it has to be determined if a given base pair is formed or broken. As the status of a basepair is evaluated from a continuum of interaction energies, a cutoff value for the interaction energy has to be imposed which separates the formed from the unformed state. Throughout this work, a base-pair was counted as formed if the energy contribution from hydrogen-bonding was below $-4.13 \times 10^{-21}\,\mathrm{J}$, corresponding to approximately 15% of the typical energy of a fully formed hydrogen bonding interaction[2]. This criterion has been applied in previous applications of oxDNA [172, 173, 174]. It was found to provide a satisfactory picture of the denaturation pattern of a simulated system, and to be robust against small changes in the precise value of the energy cutoff.

Unless explicitly stated otherwise, double stranded molecules studied in this thesis possess fully complementary sequences. In these cases, hydrogen bonding was only permitted between each nucleotide and its direct counterpart on the other single strand, excluding shifted bonds as shown in Fig. 2.3. This restriction was imposed to avoid non-generic secondary structure effects, and increase simulation efficiency. It has been used successfully before for other studies with oxDNA [174, 175, 176, 148] and generally should have a fairly small effect on the properties studied in this thesis.

2.1.3.2 Plectonemes

As discussed in Sec. 2.1.2.1, the definition of the writhe number Wr in Eq. 2.5 reflects a non-local property of the entire topologically closed double strand. It can therefore not be used directly to characterise the localization of plectonemic supercoils. Because in several parts of this work, a value characterising its spatial position has to be assigned to a plectoneme, an algorithm was developed that allows

[2] At $T = 300\mathrm{K}$, the well depths of the hydrogen bonding interaction in oxDNA are $8.9\,kT$ and $12.4\,kT$ for AT and GC base pairs respectively. This compares to experimentally determined hydrogen bonding enthalpies of $12.3\,kT$ for AT and $15.9\,kT$ for GC base pairs [131]. A reason for the difference in the values is that the energies in coarse-grained models are generally *free* energies that contain entropic components via implicit degrees of freedom [171].

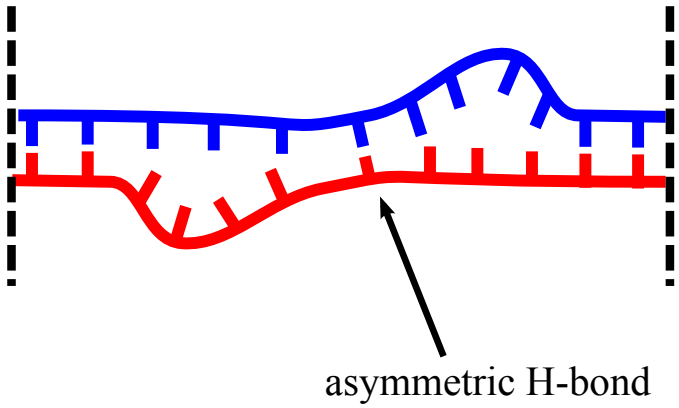


Figure 2.3: Schematic depiction of an asymmetric hydrogen bond. These bonds were excluded in simulations of plectoneme structure.

the detection of position and size of a plectoneme. The method relies on the fact that in plectonemes, the spatial distance d_{lin} between two sites on the molecule is smaller than their separation along the strand (Fig. 2.4). The algorithm only considers the midpoints between the centers of mass of corresponding bases on the single strands. It proceeds as follows:

- Start from strand end, loop over all base pair midpoints
  - If any part of the remaining strand with a distance of more than N_c bp along the contour has a distance $d_{\mathrm{lin}} < d^0_{\mathrm{lin}}$, where d^0_{lin} is a constant cutoff, record the current bp index as the beginning of a plectonemic region, if the beginning of a plectoneme has not yet been detected.
  - If $d_{\mathrm{lin}} > d^0_{\mathrm{lin}}$ and a plectoneme beginning has been detected before, record the current bp index as the end of a plectonemic region and continue searching with the next bp
- The plectoneme position is the mean between the bp indices of the beginning and end of a plectonemic region
- The plectoneme size is the difference between the bp indices of the beginning and end of a plectonemic region

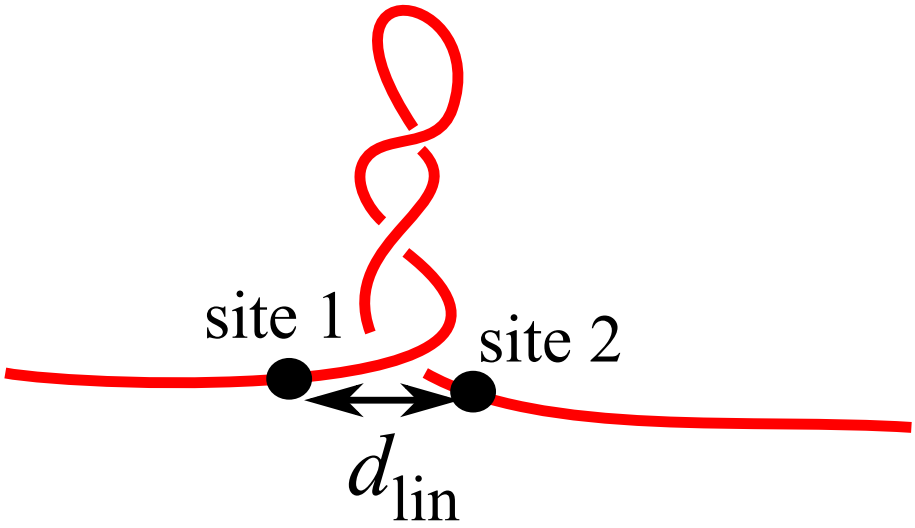


Figure 2.4: In plectonemic regions, the direct spatial distance d_{lin} between two sites is shorter than the distance along the double-strand contour.

In this thesis, plectoneme detection was performed using $d^0_{\mathrm{lin}} = 7.24\mathrm{nm}$ and $N_c = 40\mathrm{bp}$. The results are not very sensitive to the precise choice of these parameters, as long as $d^0_{\mathrm{lin}} < N_c r_0$, where $r_0 \approx 0.34\mathrm{nm}$ is the approximate rise of one base pair. N_c hence imposes a cutoff on the minimum size of plectonemes that can be detected with the search algorithm. The parameter choice made here ensures that writhed bubbles are reliably not counted as a plectonemic state, thus avoiding false positive detections. The algorithm is able to detect multiple plectonemes along the DNA double strand. However, for the salt conditions and strand length used in this work, only one plectoneme occurs in the simulated system. Two simultaneous plectonemic regions were only detected transiently during initial formation of the plectoneme. In order to obtain a single-valued plectoneme coordinate in these rare cases, only the largest plectoneme structure is considered. Extension of the algorithm to detect multiple coexisting plectonemes is straightforward.

2.2 Coarse-grained models of DNA and RNA

To study properties of nucleic acids under mechanical stresses such as twist and tension, coarse-grained models of the molecules were used. These models follow a "top-down" coarse-graining philosophy, which seeks to introduce force fields that

can heuristically capture mesoscale and thermal properties of the molecules, such as duplex melting temperatures or persistence lengths. In this approach, less emphasis is placed on the representation of microscopic structural details, such as local steric effects or sequence-specific mechanical properties of a molecule. This is in contrast to "bottom-up" coarse-graining, which starts from a microscopic representation of the molecule to which a systematic coarse-graining techniques is the applied. In this section, the properties of the models used in this work are briefly reviewed.

2.2.1 oxDNA

The oxDNA model was first introduced by Ouldridge *et al.* to study processes in DNA nanotechnology [177, 172, 149]. It represents DNA on the level of individual nucleotides, which are represented as three-dimensional rigid, classical bodies that interact over effective interactions with an *ad hoc* functional form. The interaction potential has the following structure:

$$\begin{aligned} V_{\mathrm{oxDNA}} = &\sum_{\langle i,j\rangle} \left[V_{\mathrm{backbone}} + V_{\mathrm{stack}} + V'_{\mathrm{exc}}\right] \\ &+ \sum_{i,j\notin\langle i,j\rangle} \left[V_{\mathrm{HB}} + V_{\mathrm{cross\_stack}} + V_{\mathrm{exc}} + V_{\mathrm{coax\_stack}}\right], \end{aligned} \tag{2.12}$$

where $\langle\cdot,\cdot\rangle$ represents the set of neighbouring nucleotides on a strand. V_{backbone} models the covalent bonds connecting the backbone, while V_{stack} represents the stacking of neighbouring nucleotides, $V_{\mathrm{cross\_stack}}$ the cross-stacking of a nucleotide in a base pair with the nearest-neighbour on the opposite strand, and $V_{\mathrm{coax\_stack}}$ the stacking of nucleotides on different single strands. Excluded volume interactions are associated with each nucleotide, and mediated by the potentials V_{exc} and V'_{exc}. The B-helical form of the DNA double strand is set by the difference in the equilibrium distance of backbone and stacking interaction. As a coarse-grained interaction, the potential in Eq. 2.12 is in principle a free-energy rather than a potential energy, and is therefore generally state-point dependent. In particular, the stacking interaction depends explicitly on temperature. Furthermore, interactions between nucleotides depend on their relative orientation, which permits to include anisotropic interactions and enforce right-handedness of the formed double helix. A schematic representation of

the interaction energies in oxDNA is shown in Fig. 2.5[3]. Explicit forms of the interaction potentials are given in Refs. [172, 149]. oxDNA has been parametrised to reproduce the melting behaviour of the nearest-neighbour model by SantaLucia and collaborators [131], which is thought to provide a very accurate representation of experimental melting temperatures. Both an average-base and a sequence-dependent parametrisation of the force field are available [149, 173], which allows one to study both generic and sequence-specific behaviour of the molecule. The model has been successfully tested against experimental data obtained in several different assays [173, 179, 176, 180, 175, 148], including DNA nanotweezers [177], toehold-mediated strand displacement [181] and DNA cruciform extrusion [174]. As the model provides a full 3-dimensional representation of DNA while at the same time including a thermodynamically consistent description of base-pair denaturation, it can be used to study the interplay of sequence effects on the global configuration of the molecule. In the version used throughout this work, oxDNA is limited to an ionic strength of $[Na^+] = 500\text{mM}$, as the effective potentials of the model were parametrised at that salt concentration. Some experimental data obtained at these ambient conditions are published in the literature, allowing the direct testing of the predictions obtained with oxDNA.

2.2.2 oxRNA

oxRNA is a model of RNA, developed by Šulc *et al.* [182] following a similar coarse-graining methodology as oxDNA. In particular, the interactions present in the coarse-grained molecule have the same physical motivation as those present in oxDNA, and can be subsumed in the same overall mathematical form, given in Eq. 2.12. The model can represent double-stranded RNA molecules in an A-helical conformation, and was parametrised to include sequence-dependent thermodynamics of duplexes and hairpins on the level of Watson-Crick and wobble base pairs. For the version of the model used in this thesis, the ambient monovalent salt con-

[3]Three-dimensional representations of configurations in oxDNA shown in this thesis were visualised using the Chimera package [178].

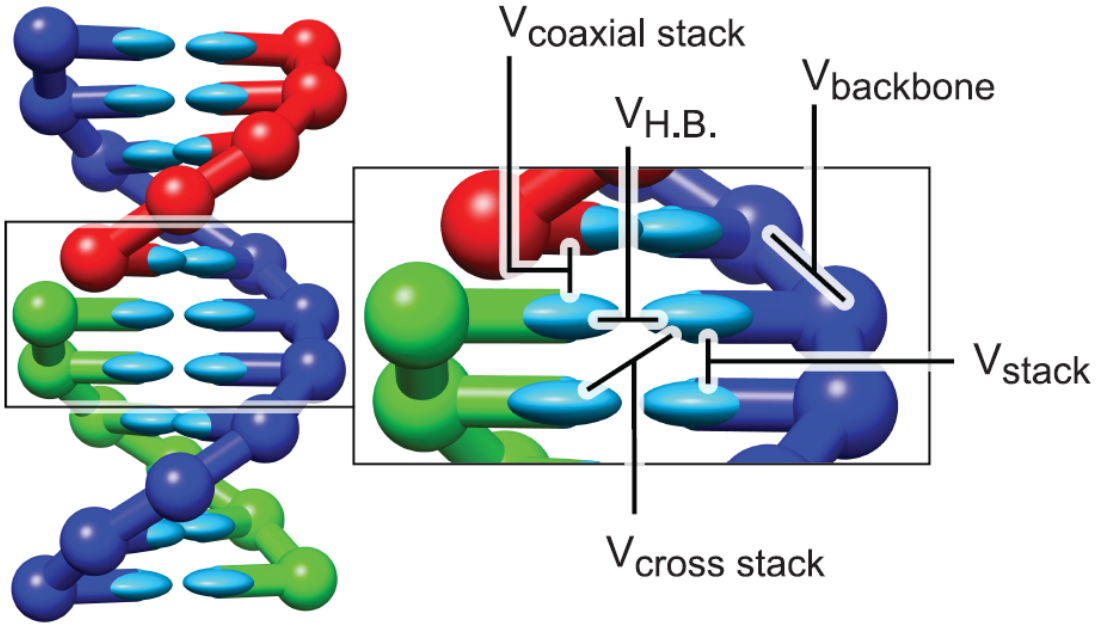


Figure 2.5: Schematics of interactions included in the potential in the oxDNA model. Excluded volume interactions are not explicitly shown. Figure reproduced with permission from Ref. [180].

centration underlying the parametrisation was $[Na^+] = 1M$. The explicit forms of interaction potentials and values of the parameters used are given in Refs. [182, 183]. In comparison to DNA, fewer direct probes for mechanical properties of double stranded RNA molecules are available. Providing a microscopic picture of mechanically stressed RNA in a coarse-grained model can therefore highlight similarities and differences in the behaviour of double stranded DNA and RNA, and suggest further experimental probes.

2.3 Simulation techniques

As oxDNA and oxRNA are implicit solvent models, it is inappropriate to directly integrate the equations of motion resulting from their interaction potentials, which would lead to ballistic particle motion.

Instead, the ANDERSEN-like thermostat that is described in more detail in Ref. [184] was used to generate the results in this thesis. It propagates the system for a set number of time steps N_{Newt} according to NEWTON's equations of motion, using the

VERLET integrator with a time step of 12.2fs. A more detailed description of the VERLET algorithm is given in Appendix A. Linear and angular velocities in the system are then assigned new values drawn from a MAXWELL-BOLTZMANN distribution with probabilities p_{lin} and p_{ang}. Unless explicitly states otherwise, all results in this thesis were obtained for $N_{\mathrm{Newt}} = 103$, $p_{\mathrm{lin}} = 0.0204$ and $p_{\mathrm{ang}} = 0.0068$.

The time-scales reported in this thesis were set by mass, energy and length scales used in the integrator. However, making direct comparisons of time-scales between coarse-grained simulations and experiments is complex [185].

Similar simulations settings have been successfully used in other applications of oxDNA (see e.g. Refs. [176, 175, 148]).

2.4 Boundary Conditions

To impose superhelical stress in simulations of a linear double strand, the ends of the strand were constrained in stiff traps, and passing of the double strand across the strand ends was disallowed. Similar constraining boundary conditions have been successfully used before in simulations of cruciform extrusion [174]. A schematic overview of the boundary conditions applied is shown in Fig. 2.6.

Fixation is implemented by adding five boundary base pairs to the strand ends, which are trapped in harmonic potentials. These potentials acting on the n-th trapped nucleotide have the form

$$V_{\mathrm{trap}}(\mathbf{r}_n; \mathbf{r}_{n,0}) = \frac{1}{2} \sum_{i=1}^{3} k^i_{\mathrm{trap}} (r^i_n - r^i_{n,0})^2, \tag{2.13}$$

where $\mathbf{r}_n = (r^1_n, r^2_n, r^3_n)$ is the centre-of-mass position of the n-th trapped nucleotide and the corresponding trap position is $\mathbf{r}_{n,0} = (r^1_{n,0}, r^2_{n,0}, r^3_{n,0})$, chosen initially to fix a given twist angle of the strand. To keep the overall linking number of the system fixed, but ensure free extensibility of the strand along the setup axis $\hat{\mathbf{x}}_\mathbf{3}$, $k^{\mathrm{trap}}_1 = k^{\mathrm{trap}}_2 = 57.1\mathrm{N/m}$ and $k^{\mathrm{trap}}_3 = 0$ was chosen. The traps defined in this way only constrain the end nucleotides in a co-moving 2-dimensional plane perpendicular to the strand setup axis, while not hindering the strand in the $\hat{\mathbf{x}}_\mathbf{3}$ direction. The high

trap stiffness in that co-moving plane ensures that fluctuations of the linking number Lk in the course of a simulation are negligible.

Strands studied in simulations have finite lengths, which means that more distant parts of the system can pass around the strand ends. Such a process would involve self-crossing of the double helical axis, and therefore change the linking number Lk and thus the superhelical density σ of the system. This process was prevented by introducing repulsion planes oriented perpendicular to the setup axis $\hat{\mathbf{x}}_\mathbf{3}$ which co-move with the first boundary nucleotide of the two single strands in the system. Repulsion planes generate a potential

$$V_{\text{plane}}(\mathbf{r};\mathbf{R}) = \frac{1}{2}k^{\text{plane}}\left((\mathbf{r}-\mathbf{R})\cdot\hat{\mathbf{o}}\right)^2\theta(-(\mathbf{r}-\mathbf{R})\cdot\hat{\mathbf{o}}), \tag{2.14}$$

where $\mathbf{r}$ is the centre-of-mass position of an affected particle, $\mathbf{R}$ and $\hat{\mathbf{o}}$ are anchor point and orientation of the plane, and $\theta(x)$ is the HEAVISIDE step function. We choose $\hat{\mathbf{o}} = \hat{\mathbf{x}}_\mathbf{3}$ and $\hat{\mathbf{o}} = -\hat{\mathbf{x}}_\mathbf{3}$ for the lower and upper repulsion planes respectively, and set $\mathbf{R}$ to the instantaneous positions of the the first and last double strand boundary base pair. If the repulsion planes would act on the boundary nucleotides in the same way as on the other particles in the system, an effective compressing force would act on the system along its setup axis, and make moves towards smaller end-to-end distances irreversible. In order to avoid hindering free strand extensibility in the $\hat{\mathbf{x}}_\mathbf{3}$ direction, the repulsion planes do not interact with the next-to-last boundary base pairs at both strand ends. In all simulations, a plane stiffness $k^{\text{plane}} = 28.5\text{pN/nm}$ was chosen, which prevented the duplex from passing over its ends during the entire simulation.

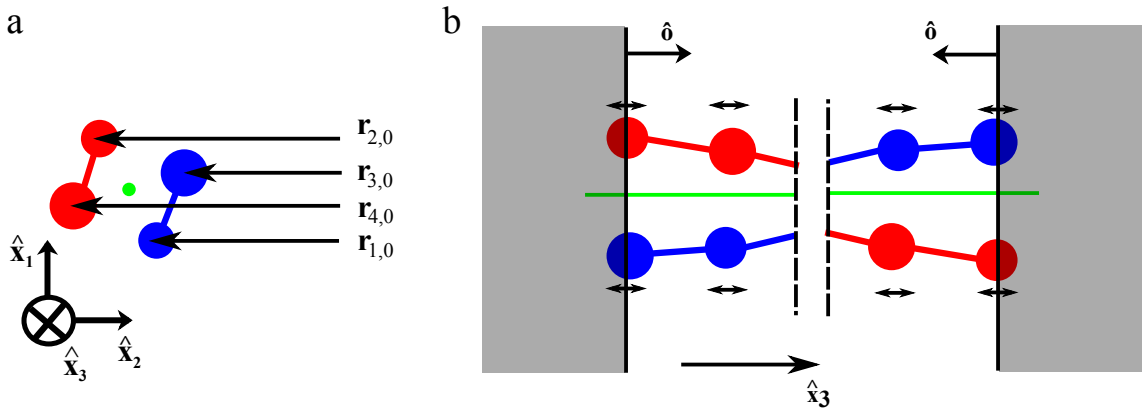


Figure 2.6: Schematic depiction of the applied boundary conditions illustrated for the last 2bp at each end of the strand: (a) View along the strand axis. 5 nucleotides at each strand end are constrained by 2-dimensional harmonic traps, which fix boundary nucleotides to positions $\mathbf{r}_{n,0}$ in planes perpendicular to the strand axis (green). (b) View perpendicular to the double strand. Due to the 2-dimensional traps, nucleotides are unconstrained only in the strand-axis direction. A repulsion plane perpendicular to the strand axis is tagged to the last base pair. Movement of nucleotides into the area below the end base pair (shaded grey) is excluded. In order to allow unconstrained strand extensibility, the repulsion plane does not act on the first two base pairs along the strand.

Chapter 3

Mechanical response of nucleic acids under twist and tension

In this chapter, the mechancial response behaviour of DNA to twist and tension is probed using the coarse-grained model introduced in Sec. 2.2. The equilibrium end-to-end extension of the molecule will be described in Sec. 3.1, and compared to a variety of datasets obtained in molecular tweezer experiments, as introduced in Sec. 1.3.2. Experimental results are available for a range of external conditions, allowing one to extensively check the results from coarse-grained simulations.

In Sec. 3.2, the response of supercoiled DNA to external torque is studied, which has recently also become directly accessible in molecular tweezer experiments [58, 65, 66]. Parameters for higher-level continuum models as described in Sec. 2.1 are determined from coarse-grained simulations, which allows one to make contact between the coarse-grained and the continuum-elastic level of description. In addition, the twist-stretch coupling term in elastic models arising from the chiral structure of the molecule is investigated in Sec. 3.3. In Sec. 3.4, the study is extended to the mechanical properties of double stranded RNA using the recently developed oxRNA model.

3.1 Extension behaviour of DNA

Typical molecular tweezer experiments (see Sec. 1.3.2) yield direct measurements of the end-to-end extension of a DNA molecule, under given values of superhelical density σ and stretching force F. In this section, results from analogous simulations

in oxDNA are presented, and compared directly to experimental data. Superhelical densities in the range $-0.1 \leq \sigma \leq +0.1$, and stretching forces $0.25 \leq F \leq 7.8$pN were imposed[1]. All simulations presented in this chapter were performed at temperature $T = 300$K. Equilibrium data for end-to-end extension was obtained within about two weeks of runtime on a current CPU for the 600bp systems studies presented in this section.

In Secs. 3.1.1 and 3.1.2, the extension behaviour will be studied using the average-base parametrisation, allowing one to focus on generic properties of the systems. The effects of particular sequences on the strand extension behaviour will be described in Sec. 3.1.3.

Experimental measurements have been made for a variety of different DNA strand lengths. Most of the published data was obtained for dsDNA lengths in the range of $1 - 20$kbp [4, 1, 6, 60]. In order to remove the explicit dependence of results on system size and facilitate comparison to experimental data, results in this section are mostly given as *relative* extensions $l = L/L_0$, which is normalized by the strand contour length L_0.

3.1.1 Generic extension behaviour as a function force and supercoiling

The relative end-to-end extension of the double strand consisting of a 600bp system as a function of imposed supercoiling density σ and stretching force F is shown in Fig. 3.1.

The extension "hat curves" show a nearly flat regime at low values of $|\sigma|$, followed by abrupt shortening of the end-to-end distance at a characteristic values σ_b, where *buckling* of the double strand takes place, and a plectonemic structure is formed. The buckling transition is in turn followed by a regime in which the end-to-end extension of the double strand shortens linearly with $|\sigma|$, corresponding to the growth of the plectoneme structure stem (see Fig. 3.2).

[1]The zero point of σ is determined by requiring the overall torque on the system to vanish, as discussed further in Sec. 3.2.

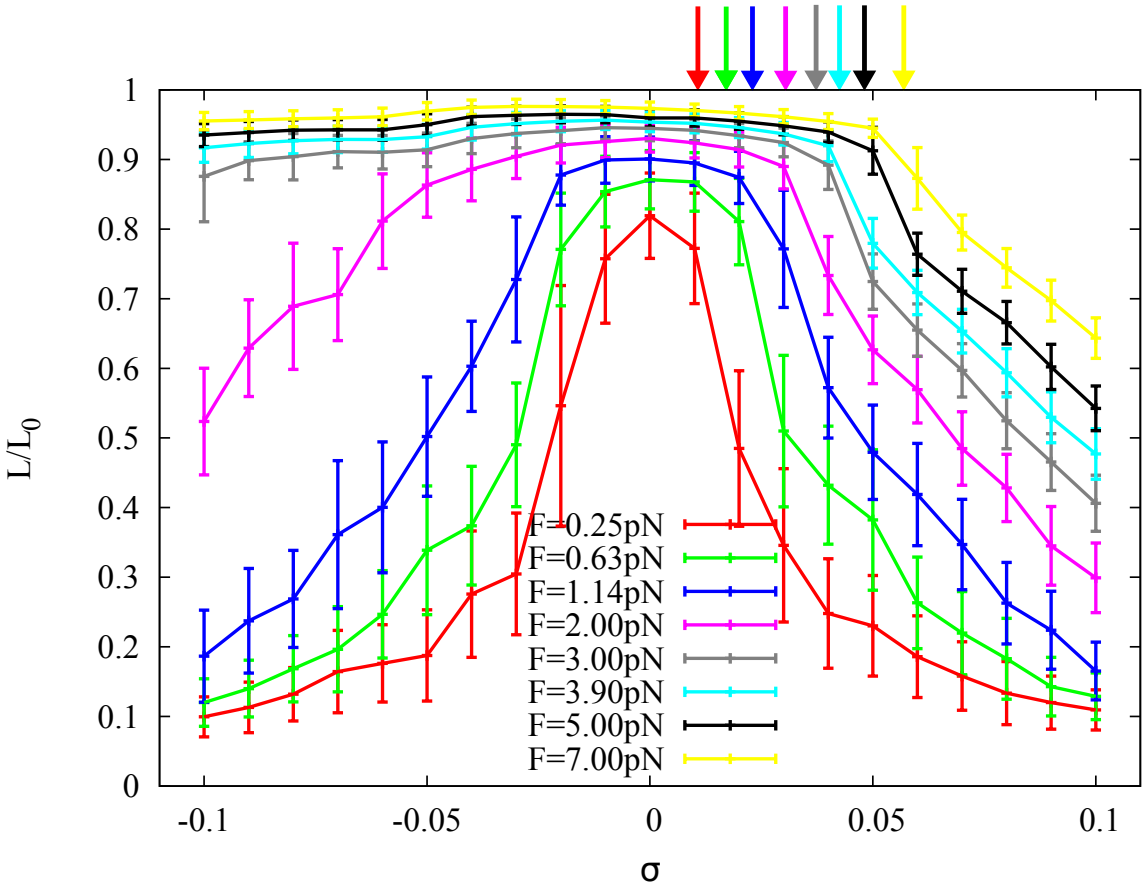


Figure 3.1: Strand extension "hat-curves" show the mean relative extension of a 600-bp duplex against superhelical density σ for various applied stretching forces F. Error bars indicate thermal fluctuations in the end-to-end distance, rather than sampling uncertainties. For stretching forces $F \lesssim F_c$ smaller than a characteristic crossover force $F_c \approx 2\text{pN}$, extension curves are symmetric in σ. For higher forces, denaturation takes place, causing a uniformly extension for $\sigma < 0$. Nonlinearities in the post-buckling slopes of the hat curves for high values of $|\sigma|$ are attributable to finite size effects. Coloured arrows indicate the value of supercoiling at which buckling is predicted to occur by the linear elastic model of Eq. 3.3 for the respective force.

Change in the end-to-end extension of the DNA strand is symmetric for $\sigma \to -\sigma$ at small stretching forces. In contrast, for stretching forces $F \gtrsim 2\text{pN}$, significant length reduction is only observed for positive supercoiling. At negative supercoiling, base pairs can break, and denatured regions (denaturation bubbles) form in the strand. This precludes the formation of plectonemic structures, and therefore does not result in significant change in the end-to-end extension of the strand. Consistently, $l \approx 1$ is observed for negative supercoiling at high stretching force, a

behaviour that is approximately independent of $|\sigma|$.

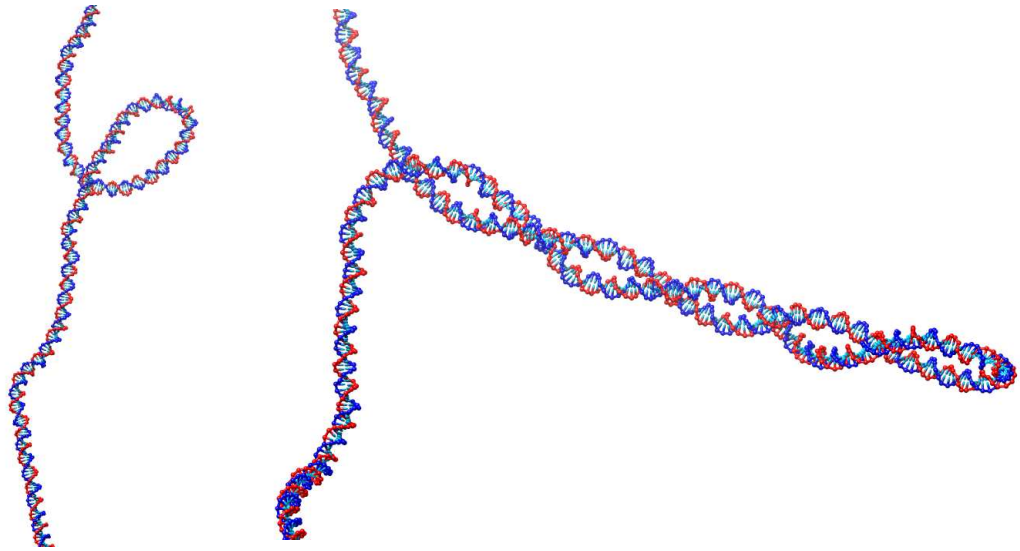


Figure 3.2: Growth of the plectoneme structure upon increasing the superhelical density at $F = 2\text{pN}$. Left: $\sigma = +0.04$, Right: $\sigma = +0.08$

Note that for high values of $|\sigma|$ and low stretching forces F, a deviation from the linear behaviour of the slope in extension is observed. This effect occurs when most of the simulated system is taken up by the plectoneme structure. Further over-winding is hindered by the presence of system boundaries used to keep the overall torsion on the system constant (see Ref. 2.4). Moreover, these system boundaries force an alignment of the plectoneme perpendicular to the setup axis of the system, which can influence the measured end-to-end distance of the strand. Qualitatively similar, somewhat less pronounced finite-site effects are also observed in experimental data (see Sec. 3.1.2).

The extension behaviour observed can be compared directly to experimental data from single molecule assays. Most available data has been obtained using magnetic [60, 4, 5, 6] or optical tweezers [58]. Available datasets differ in molecule length L, ionic strength of the buffer, and sequences properties of the DNA strands used in the respective experiments.

Here, the results obtained for oxDNA in a 600bp system are related to experimental findings. Additionally, the influence of different parameters on the quality of agreement is explored. Note that the version of oxDNA used to determine the

results in this chapter is limited to a monovalent ionic strength of $[Na^+] = 500\text{mM}$, as the effective potentials of the model were parametrised at that salt concentration. This simulated strand length is considerably shorter than the strand lengths typically used in experiment, which will affect the comparison of results. The influences of strand length on the results is explored further in Sec. 5.1, where data for strands of length 1.5kbp is presented, which were obtained using GPU computing.

To facilitate comparison to experimental data, results are given in terms of the relative strand extension $l = L/L_0$ and the superhelical density σ, which are length-independent up to finite-size corrections vanishing in the long-strand limit [118] and interactions with boundary potentials. Some uncertainty exists about the relaxed length L_0, because in typical experimental assays, measurements of end-to-end distance take place on molecules attached to a cover slip and a bead via chemically functionalised ends, such as biotin and digoxigenin [59]. Attachment may slightly modify the effective free length of the double strand, or the mechanical properties of the DNA strand close to the attachment point. The simulated system only uses a very short piece of strand to constrain the duplex ends (see Sec. 2.4) and therefore does not include these chemical attachment effects. Uncertainty about the effective free length of attachment may consequently introduce small systematic differences when comparing experimental data to simulation results, in particular when measuring the overall strand length L. By contrast, experimental measurements of the buckling superhelical densities σ_b and post-buckling slopes $dl/d\sigma$ are expected to be largely unaffected by the absolute length of attachment. Therefore, experimental buckling slopes are easier to compare with simulation results than absolute lengths.

The buckling superhelical density σ_b is determined primarily by the values of bending persistence length and twist persistence length [4], which in oxDNA take the values $A_0 = 42.5$ nm and $C_0 = 114.7$ nm respectively (see Sec. 3.2). As these values lie in the range of values reported for DNA [6, 3], oxDNA might be expected to reproduce experimental values for σ_b. Small deviations with experiment may simply stem from differences in the value of these constants. These differences would occur

for any elastic model such as the one presented in Sec. 2.1.1 parametrised by using these values for A_0 and C_0.

In particular, consider the linear elastic model given by Eq. 2.1. At the relatively low forces considered here, twist-stretch coupling and linear extension elasticity can be neglected, setting $B_0 = D_0 = 0$. Then, the energy cost associated with the formation of the first plectoneme coil can be approximated by

$$E_{\text{coil}}(\alpha, R) = \alpha R \frac{1}{2} \frac{k_B T A_0}{R^2} + \alpha R F, \tag{3.1}$$

where α is the overall bending angle, and R the radius of the coil. The first term in Eq. 3.1 is due to the bending energy of the plectonemic coil approximated as a circle, while the second term stems from the work done against the external stretching force due to a change in extension upon coiling. Minimizing E_{coil} with respect to R yields $E_{\text{coil}} = \alpha\sqrt{2k_B T A_0 F}$. On the other hand, the twist energy associated with the angle α follows from the third term in Eq. 2.1 as

$$E_{\text{twist}} = \frac{k_B T C_0}{2} \frac{\alpha^2}{l_0^2} l_0, \tag{3.2}$$

where l_0 is the overall contour length of the strand. Buckling will take place when $\partial_\alpha E_{\text{coil}} = \partial_\alpha E_{\text{twist}}$, which marks the point at which bending of the double strand becomes energetically more favourable than further twisting. Solving for α and normalising by Lk_0, one finds

$$\sigma_b = \sqrt{\frac{2FA_0}{k_B T}} \frac{r_0 p_0}{2\pi C_0} \tag{3.3}$$

as an estimate for the superhelical density σ_b at which buckling takes place, where $p_0 \approx 10.36\text{bp/turn}$ is the untwisted pitch of DNA, and $r_0 \approx 0.34\text{nm}$ its equilibrium rise. Satisfactory agreement between the simple prediction of Eq. 3.3 and simulation data (see Fig. 3.1) indicates that the coarse-grained model reproduces the buckling behaviour of continuum models on the level of linear elasticity, although Eq. 3.3 was derived from purely energetic considerations and is not based on a consistent treatment of thermodynamic effects. Thermal effects can however be included on the level of a continuum model, as done by Marko and co-workers [119, 118].

The post-buckling slopes are expected to mainly depend on the radius of the plectoneme stem, which is set by the competition of twisting and bending energies, as well as by screened electrostatic interactions between the double strands [186, 120]. Agreement with experimentally determined post-buckling slopes, which will be discussed in more detail in Sec. 3.1.2, can therefore by used to assess if the treatment of strongly screened electrostatics by excluded volume interactions in oxDNA is consistent, and yields plectoneme radii comparable with the ones obtained in experiments.

In addition to reproducing basic elastic properties of the double strand, oxDNA was parametrised to provide a good representation of the melting transition of double-stranded DNA. The model was successfully tested in other systems where breaking of base pairs plays a role [172, 176, 148]. This motivates applying the model to study plectonemes also for negative supercoiling, where both effects of strand elasticity and double strand denaturation are expected to be relevant, and study the mutual influence of denaturation and writhing.

All simulation results presented in this section were obtained using the average-base parametrisation of oxDNA. Details of the mechanical response of DNA to super-helical twist may however depend on the specific sequence used. Firstly, sequence-dependent structural properties of double stranded DNA may yield specific mechanical reponse behaviour. These local effects are expected to be very small for long enough strands with sufficiently heterogeneous sequences, and are not captured by the modelling approach of oxDNA. Additionally, however, the precise value of the crossover force F_{char} between bubbles and plectonemic structures might be expected to depend on the base sequence of the DNA strand, as the enthalpic cost of bubble formation decreases with increasing AT content. Therefore, AT-rich stretches represent preferred nucleation sites for denaturation bubbles, which can facilitate denaturation and therefore decrease F_{char}. The effect of different sequences on "hat curves" is studied in more detail in Sec. 3.1.3.

3.1.2 Comparison to experimental data

Measurements of strand extension as a function of applied force F and superhelical density σ are very reproducible. They have have been measured by several labs in many setups for a variety of different ambient conditions [4, 6, 2, 3, 60, 5].

In this section, the computational results are compared to measured data from a set of recent single-molecule experiments. From published experimental results in the literature, data was chosen from assays performed at conditions close to those used in the parametrisation of oxDNA. Strand extension results are plotted together with experimental data in Figs. 3.3 - 3.5.

Fig. 3.3 shows a comparison to experimental data of Tempestini *et al.* [2] on a stand of length 7kbp. From the length of the attachment described in Ref. [2], an effective free strand length of 5.8kbp can be inferred. The salt concentration used in these experiments is $[Na^+] = 500$mM, the value at which oxDNA was parametrised. Post-buckling slopes are in good agreement with simulation data, whereas buckling occurs at slightly lower experimental superhelical density σ_b. This is consistent with additive finite-size corrections on the order of $L^{-1/2}$ predicted in continuum models of the buckling transition [118].

Fig. 3.4 shows a comparison to data by Salerno *et al.* [3], obtained at $L = 6$kbp and a lower salt concentration $[Na^+] = 150$mM. Lower salt is expected to destabilise plectonemes and make bubbles more favourable, as electrostatic interactions of the backbone are less screened by ions in solution. In particular, this effect is expected to become more significant for $\sigma < 0$, as has been shown before experimentally [2]. Comparing to simulation results, good agreement is still observed in the post-buckling slopes for all forces when $\sigma > 0$, and no significant denaturation effects are expected. In contrast, the crossover to the extended bubble state for $\sigma < 0$ happens at higher forces for oxDNA. That the crossover force is higher is expected because of the difference in salt concentration between simulation and experiment mentioned before. Although appropriate hat curves are unavailable in the current literature at higher salt for $\sigma < 0$, data on the fluctuation of end-to-end

lengths from Ref. [2] suggest that the transition to the extended bubble state occurs at roughly 1pN, about 1pN below the oxDNA estimate for $[Na^+] = 500mM$. These differences are discussed in more detail in Sec. 5.1 on end-to-end fluctuations, and Sec. 3.1.3 on sequence dependence. In this context, it is shown that taking sequence into account with a sequence-dependent model leads to lower forces for the crossover to bubble formation than using an average-base model. This might be expected, as AT-rich regions are more likely to form bubbles, which is why the overall onset of denaturation in a strand is expected to happen at lower forces when weak AT-rich regions are present and provide preferred start sites for base pair breaking.

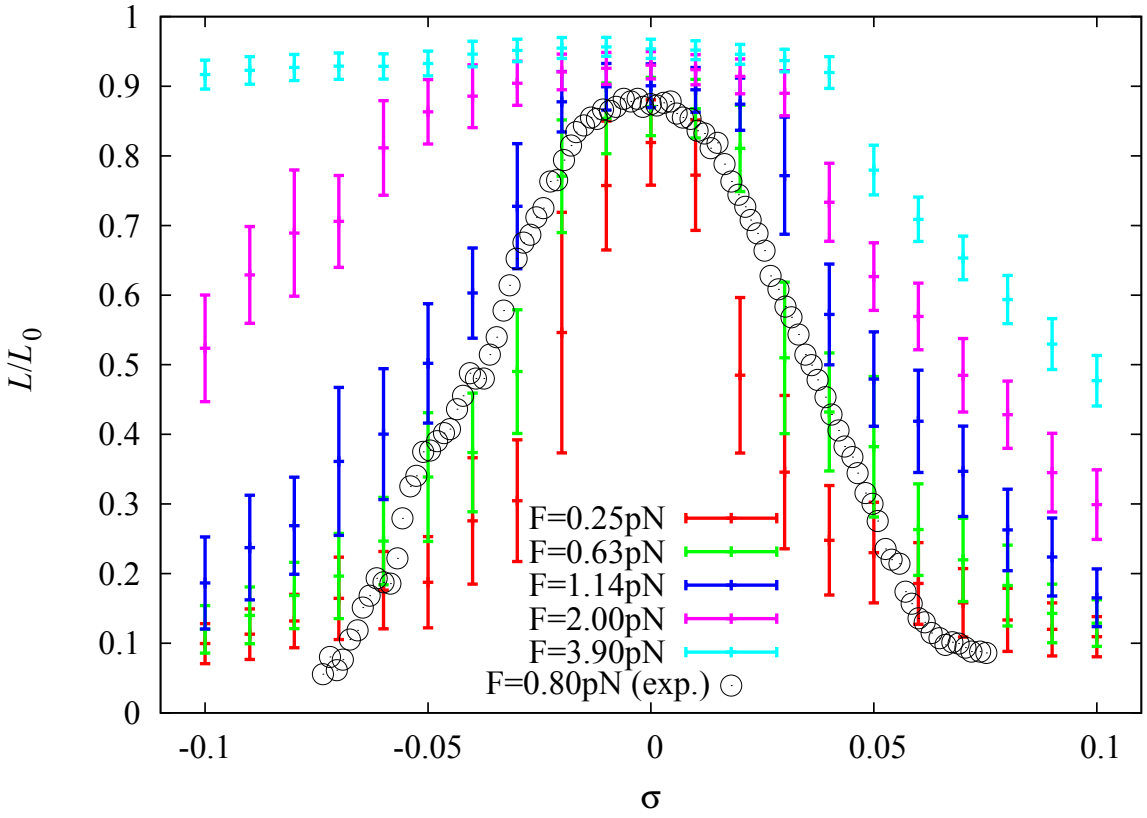


Figure 3.3: Experimental strand extension data of Ref. [2] at $[Na^+] = 500mM$ and $L \approx 5.8kbp$ for $F = 0.8pN$, compared to data from simulations using oxDNA ($[Na^+] = 500mM$ and $L = 600bp$). Error bars on simulation results indicate thermal fluctuations in the end-to-end distance, rather than sampling uncertainties. Good agreement is observed, while nonlinear effects in the post-buckling slopes for high $|\sigma|$, caused by end effects, are somewhat more pronounced for the shorter simulated system.

In the data of Refs. [3] and [2], deviations of the post-buckling slopes from

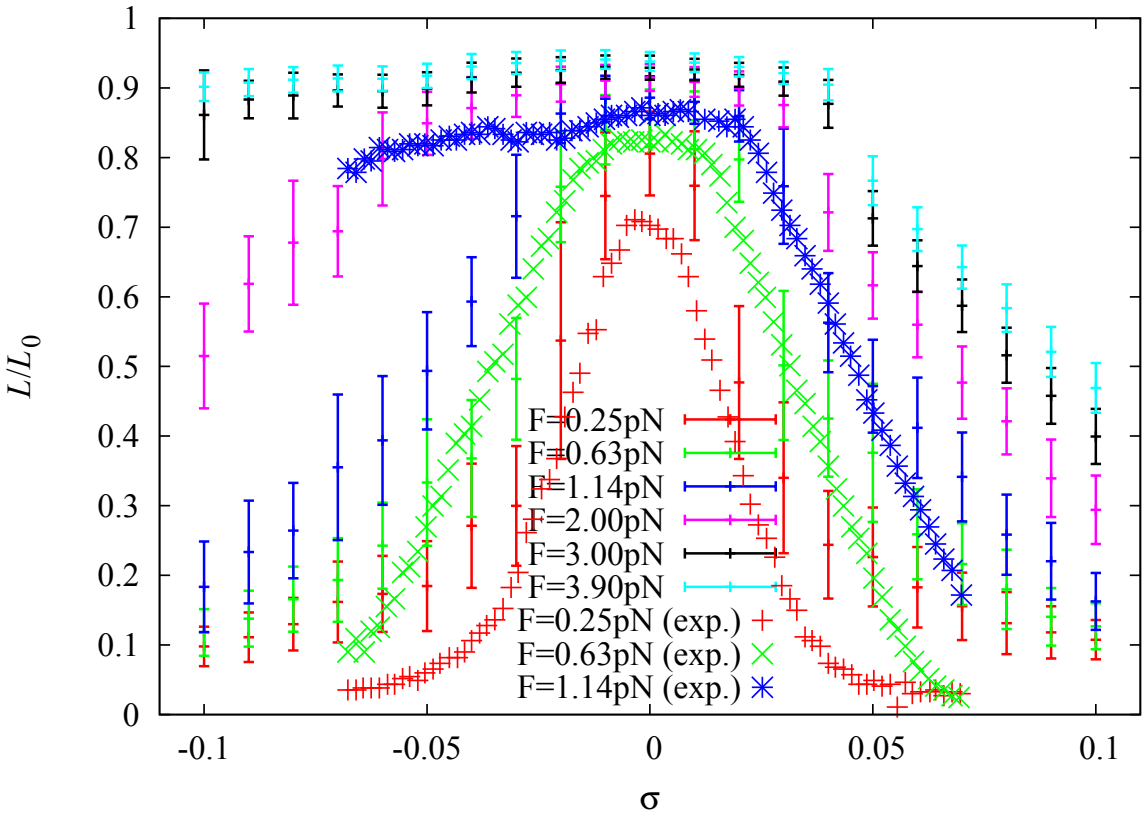


Figure 3.4: Experimental strand extension data of Ref. [3] at $[Na^+] = 150$mM and $L \approx 6$kbp for $F = 0.25$pN, 0.63pN and 1.14pN compared to data from simulations using oxDNA ($[Na^+] = 500$mM and $L = 600$bp). Data for equal stretching force is represented by the same color. Error bars on simulation results indicate thermal fluctuations in the end-to-end distance, rather than sampling uncertainties. At lower salt concentration, plectoneme structures are enthalpically penalised as electrostatic interactions are screened less, making them less favourable compared to bubble configurations. Lower salt concentration also reduces the free-energy cost of base pair breaking. Hence, the crossover from tip-bubble plectonemes to extended bubble states occurs at lower forces than would be the case for $[Na^+] = 500$mM. Sequence dependence may additionally decrease the crossover force F_{char} by providing AT-rich bubble nucleation sites, as discussed in Sec. 3.1.3. Non-linear regions in the post-buckling slope are somewhat more pronounced for the shorter simulated system.

linearity were observed at low stretching forces and high values of $|\sigma|$. Qualitatively similar, more pronounced deviations are observed in simulation results due to finite-size effects stemming from the interaction of the coiled DNA strand with the system boundaries, as described in Sec. 2.4. Boundary interactions become relevant when the diameter of the plectonemic region is comparable to the end-to-end extension of the double strand. This is expected to be the case at somewhat lower values of $|\sigma|$ in the 600-bp system studied in simulations, as compared to the approximately 10

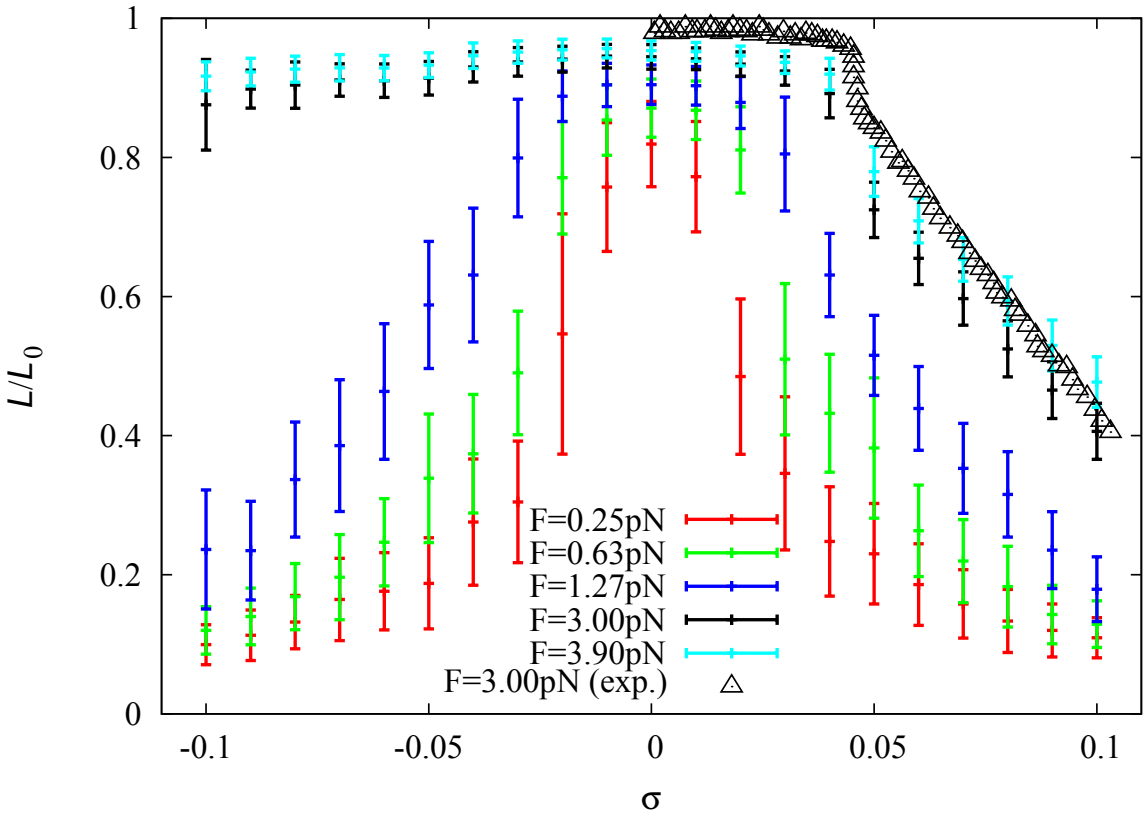


Figure 3.5: Experimental strand extension data of Ref. [4](black triangles) for $[Na^+] = 320$mM and $L = 1.9$kbp at $F = 3.0$pN compared to data from simulations using oxDNA ($[Na^+] = 500$mM and $L = 600$bp). For simulation data, error bars indicate standard deviations due to thermal fluctuations, rather than sampling uncertainties. Good agreement of buckling point and post-buckling slope is observed. Note also the abrupt length reduction upon buckling, present in both experiment and simulations.

times longer experimental systems of Refs. [3] and [2], shown in Figs. 3.3 and 3.4. Additionally, lower forces correspond to a lower enthalpic penalty on deviations of strand configurations from the setup axis of the system, which is why coiled structures with less localised plectonemic structure are accessible to the system at small values of F. These looser coiled structures may deviate from the response of plectonemes to supercoiling, thus adding to the observed nonlinear behaviour.

Fig. 3.5 depicts data obtained by Brutzer *et al.* [4] at $[Na^+] = 320$mM, $L = 1.9$kbp and $\sigma > 0$. Close agreement is observed concerning both the buckling superhelical densities σ_b and the post-buckling slopes $dl/d\sigma$ (Fig. 3.5). In this study, abrupt strand shortening at the buckling point was observed which was consistent

with earlier observations using optical tweezers [58]. The physical origin of abrupt length reduction upon buckling was ascribed to the energetic cost of forming the plectoneme end-loop. The energetic barrier towards formation of the plectoneme was shown to become more pronounced with increasing ionic strength, and decrease in the limit of long strands. This behaviour might be expected, as formation of the plectoneme end-loop is a local effect, whose influence on the overall extension of the strand should become washed out for long systems. Based on these experimental findings, the observed marked shortening of the strand at similar force around the buckling point observed in the coarse-grained simulations is consistent (see Fig. 3.5. A finer sampling of the buckling behaviour as a function of superhelical stress around σ_b is depicted in Fig. 3.6, showing good agreement with the experimentally observed buckling behaviour.

3.1.3 Sequence dependence

Fig. 3.6 shows "hat curves" obtained for both the average-base and sequence-dependent parametrisation of oxDNA. Sequence-dependent simulations were performed for a fully random sequence which was also used to study plectoneme localisation. The full sequence is given in Appendix C. Outside the regime in which stable denaturation bubbles can form, results show no significant difference. However, the sequence-dependent model exhibits a systematically larger extension when forces become big enough to allow denaturation. This indicates that denaturation occurs more easily in the sequence-dependent case than for the average-base parametrisation.

Such behaviour is expected, as weak AT base pairs represent preferred sites of base-pair breaking, making denaturation of the strand less enthalpically costly, while not affecting the bending energy cost of the strand much. From simple considerations of energy scaling for the competition between denaturation and plectoneme formation, the crossover force can be estimated to depend on the free energy of base pair breaking α as $F_{\mathrm{char}} \propto \alpha^2/B_0$ [3]. Hence, sequence-dependent differences in denaturation energy are expected to have a noticeable effect on strand extension properties, and in particular the crossover force F_{char}. As denaturation bubbles can

grow to significantly larger sizes than 1 bp, not only the overall AT content of the strand, but also the base distribution may influence extension properties. For example a base sequence showing long AT-rich stretches may allow stable formation of fewer bubbles than sequences with shorter AT-rich stretches, leading to a slight further decrease in F_{char}.

Next to the sequence-dependent features of mechanical properties in supercoiled DNA that are due to the thermodynamics of double strand denaturation, additional sequence-dependence due to variations in structure may be expected. For example, particular sequences such as A-tracts or poly(dA)-poly(dT) sequences are known to induce intrinsic bending in double strands [187, 188, 189]. These particular properties are not captured by the coarse-graining approach used in oxDNA, which includes sequence-dependence on the level of the thermodynamics of melting, but not on the level of microscopic elasticity. However, while these properties are important for local DNA-protein interactions [188], there is no indication that they significantly influence the overall mechanical behaviour of longer strands. Therefore, the sequence effects presented here can be expected to dominate these mechanical effects for sufficiently long strands.

3.2 Torque response of DNA

While extension curves can be determined relatively easily in experiments and are available from a large variety of molecular tweezer assays, direct measurement of the torque response of DNA has proven more challenging [190]. In usual setups, torque is determined by tracking the twist angle of the strand [64, 65, 5]. Alternatively, optical traps [58] or integration of extension curves [6] allow to access the torque response more indirectly. Typically, the torque response consists of a linear regime for low $|\sigma|$ until the buckling point, which is followed by a constant post-buckling torque. Recent, careful measurements [58, 5] have revealed that the linear and constant regimes are separated by a torque overshoot which has been related to formation of the plectoneme end-loop at the buckling transition [4].

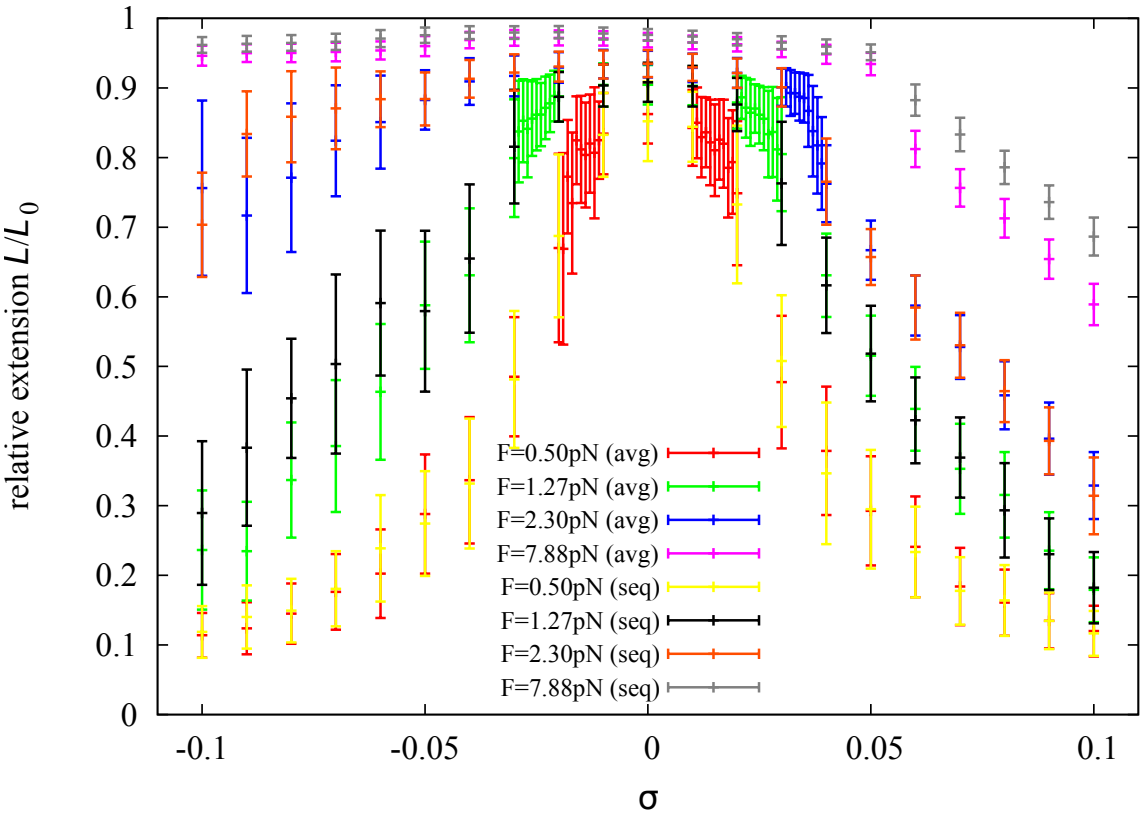


Figure 3.6: Comparison of extension curves for the random sequence given above in the average-base and sequence-dependent parametrisation of oxDNA. For parameter values for which tip-bubble plectonemes occur, sequence-dependent simulations show a systematically higher extension, as weak regions in the sequence provide preferred sites of base-pair breaking, thus enhancing overall strand denaturation. Error bars indicate thermal fluctuations in the end-to-end distance, rather than sampling uncertainties.

For the determination of torque both in experiment and simulations, the zero point of supercoiling is usually determined by requiring the overall torque on the system to vanish, i.e $\Gamma(\sigma = 0, F) = 0$ [6]. The same convention was applied to determine the zero point of σ for the simulation results presented in this chapter, leading to a helical pitch of $p_0 = 10.36\,\mathrm{bp/turn}$. Generally, this value might be expected to depend on F due to non-zero twist-stretch coupling (see Sec. 3.3), as well as the specific sequence and system temperature used. However, for the parameter values and sequences studied in this thesis, no significant deviations from p_0 were observed, indicating that these effects represent minor corrections.

Corresponding to the difficulty in directly measuring the torque response, fewer experimental measurements are available in the literature than for the end-to-end

distance. The published dataset which is closest to the conditions used in the parametrisation of oxDNA is the one recently measured by Janssen *et al.* [5] at $[Na^+] = 550\text{mM}$ and $L = 7.9\text{kbp}$ using a variant of the magnetic tweezer setup, the electromagnetic torque tweezer. A comparison to the torque response data is shown in Fig. 3.7, exhibiting very good agreement between experiment and simulation for both the size and location of the torque overshoot, as well as the slope of the torque response in the linear regime and the magnitude of the constant post-buckling torque.

Post-buckling torques determined in simulations are compared to data taken by Mosconi *et al.* [6] in Fig. 3.8, again showing good agreement for forces higher than approcimately 1pN. At lower forces, the torques determined in simulations are higher than those determined in experiments. This may be due to finite-size effects which become relevant at low stretching forces, at which almost the entire simulated strand is taken up by an extended, collaped plectoneme structure.

The additional energy imposed on the double-stranded system due to torsional stress excites twisting and bending modes of the DNA strand. This leads to a renormalization of the torsional stiffness C_0 of the double strand, yielding an effective twist presistence length C_{eff} [113, 116]. By using a torsional directed walk model, Moroz and Nelson estimated the effect of fluctuations to lead to an effective twist persistence length, as described in Eq. 2.3 of Sec. 2.1.1. Eq. 2.3 allows one to consistently determine both A_0 and C_0 by measuring the effective torsional stiffness as a function of stretching force F. In oxDNA, these microscopic values have previously been determined as $A_0 = 42.5\text{nm}$ and $C_0 = 114.7\text{nm}$ [172]. Here, the force-dependent effective twist persistence length C_{eff} of the simulated 600-bp system was determined by fitting the slope of the linear regime of the torque response curve (see Fig. 3.7). Good agreement of the measured torsional moduli with the theoretical prediction is observed (Fig. 3.9). If the functional form of Eq. 2.3 was fitted to the data with A_0 and C_0 as free parameters, $A_0 = 40.8\text{nm}$ and $C_0 = 115.6\text{nm}$ are obtained, which are very similar to the previously determined microscopic values.

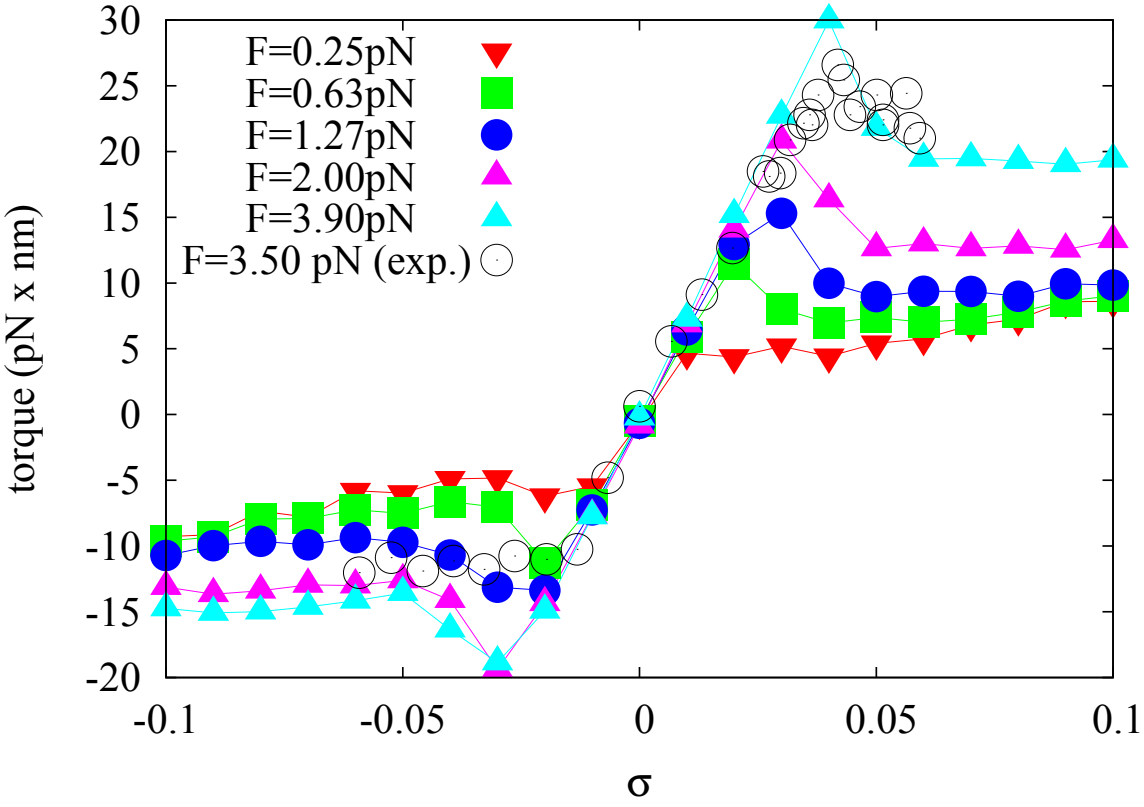


Figure 3.7: Experimental torque response data of Ref. [5] taken at $[Na^+] = 550$mM and $L = 7.9$kbp for $F = 3.50$pN compared to data from simulations using oxDNA ($[Na^+] = 500$mM and $L = 600$bp). Excellent agreement is observed both in the linear regime (see also Fig. 3.9) and for the location and size of the torque overshoot.

Note that the agreement shown in Fig. 3.9 assumes the previously determined values of A_0 and C_0, and is obtained without free fit parameters.

3.3 Twist-stretch coupling

As the DNA double helix is a chiral, anisotropic molecule, the twist and stretching modes are generally not independent [166]. This twist-stretch coupling is reflected to first order in elastic theory by the term proportional to D in the linear elastic model introduced in Eq. 2.1 in Sec. 2.1.1. Experimentally, it is found that the magnitude of twist-stretch coupling in DNA is very small. Somewhat contrary to the intuition derived from twisting macroscopic objects, the twist-stretch coupling constant D was found to be negative in single-molecule assays, meaning that the molecule lengthens when overwound [153, 154].

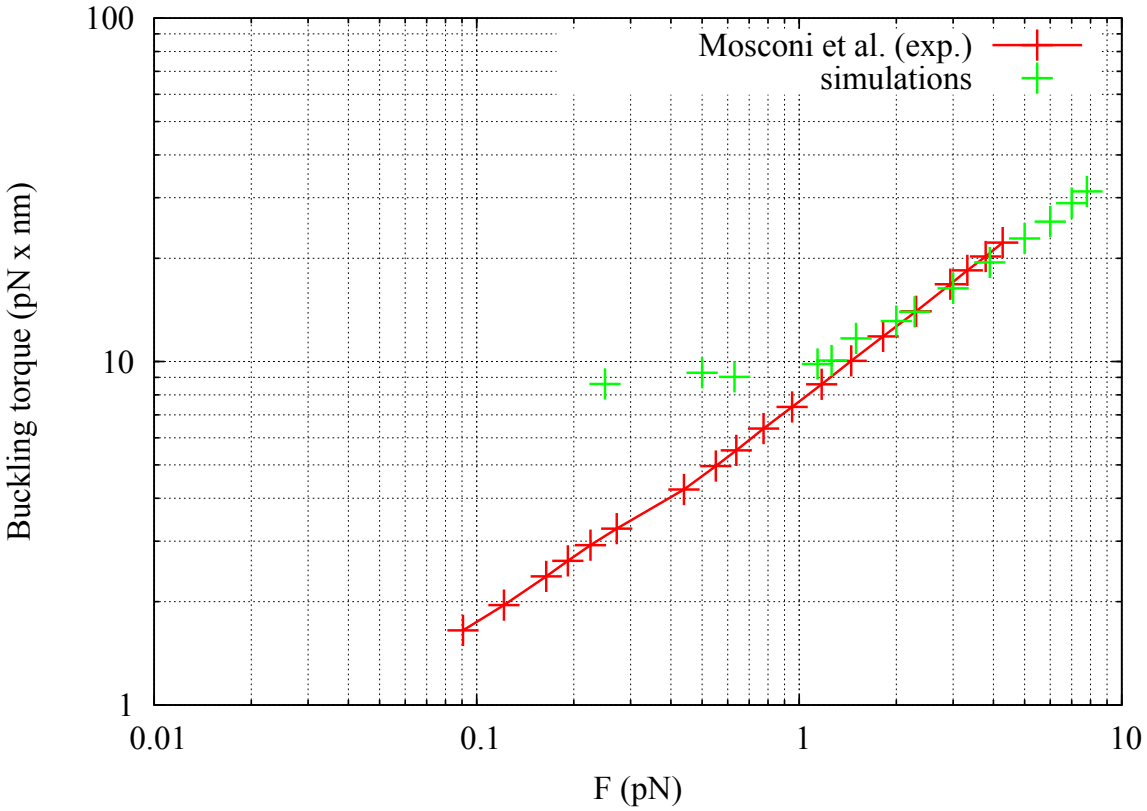


Figure 3.8: Comparison of postbuckling torques determined in simulations to data by Mosconi *et al.* [6], taken at $[Na^+] = 550$mM for a system of length 6kbp. Excellent agreement is observed down to a force of approximately 1pN, when finite-size effects become relevant for the simulated system.

The twist-stretch coupling of in the oxDNA model was determined by fitting a linear function to the of the hat curve for $F = 7.0$pN in the range $-0.01 \leq \sigma \leq +0.02$ (see Fig. 3.1). This measures the change in extension of the B-helical molecule for small variations in twist, and is in analogy with the analysis of experimental data performed by Lionnet *et al.*. The slope was determined as $dL/d\sigma = -58 \pm 6$nm, which corresponds to a value of $D = -\frac{B_0}{2\pi}\frac{dL}{d(\Delta Lk)} \approx +80$ for the dimensionless twist-stretch coupling parameter in the linear-elastic model of Eq. 2.1.

This value compares to the experimental values $D = -22$ found by Gore and collaborators [153] and $D = -16$ reported by Lionnet *et al.* [154]. Therefore, the anomalous twist-stretch coupling found in DNA experimentally cannot be reproduced correctly by the current parametrisation of the oxDNA model. It should be noted that the magnitude of deviation may be related to the value of the stretch

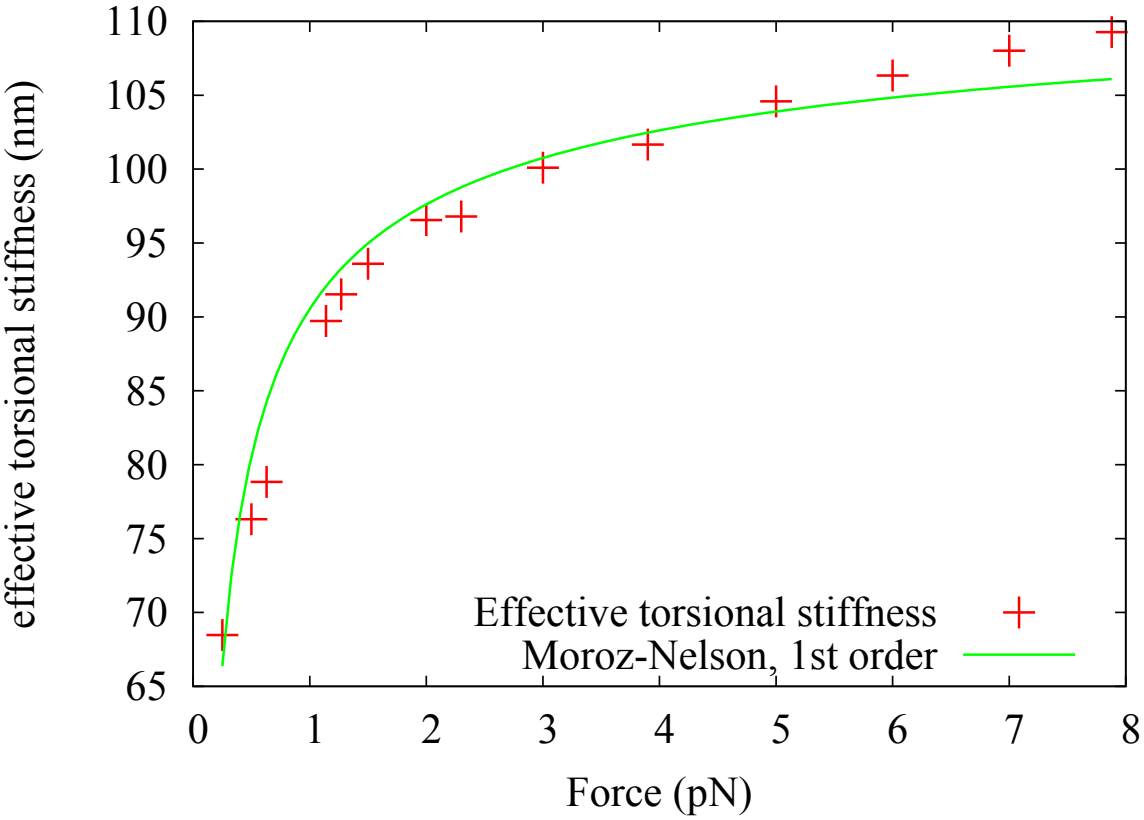


Figure 3.9: Prediction of the effective twist persistence length from microscopic values according to Moroz-Nelson theory (green lines show first-order expansion; no free parameters) and values determined from simulations (red crosses).

parameter $B_0 \approx 523nm^{-1}$ of oxDNA, which has been determined previously [149] and is about two times larger than typical experimentally measured values.

Despite theoretical efforts using both the microscopic [154, 191] and coarse-grained models [192, 155], the precise microscopic structure underlying the observed behaviour is still not known. It is therefore not immediately clear which feature of oxDNA would have to be modified to capture twist-stretch coupling more accurately. Atomistic modelling indicated that a decrease in the double-strand radius and an associated negative inclination of base pairs may contribute to the observed behaviour [154, 191]. It may therefore be that stacking interactions in the current parametrisation of the model is slightly too stiff to reproduce the precise inclination of base pairs under tension. Atomistic simulations also suggest a considerable dependence on sequence-dependent structural properties [191, 98], which may be difficult to include in the coarse-graining approach of oxDNA. A recent bottom-up

coarse-grained model on the level of rigid base pairs was able to capture the correct sign of twist-stretch coupling, although with considerable variance in the magnitude, while at the same time giving a far too low twist persistence length C_0 [155].

These considerations highlight that coarse-grained models reflect a compromise, representing some features of the molecule more accurately than others. For oxDNA, a primary aim is to capture the physics of bending and twisting accurately, while putting less emphasis on linear elasticity and twist-stretch coupling. Relative to the other interactions in supercoiled DNA, twist-stretch coupling is a very small effect, which should only introduce local corrections to the structure of double strand which are on the order of 10^{-4}rad/bp [153]. Therefore, the overall effect on the results presented on the writhing and denaturation behaviour of supercoiled DNA presented in this thesis can be expected to be minor.

3.4 Mechanical properties of RNA

As for the case of DNA, the extension and torque response properties for oxRNA were determined using the same simulation setup and molecule length $L = 600\,\mathrm{bp}$. Extension "hat curves" of the RNA model as a function of σ and F are shown in Fig. 3.10. The overall mechanical behaviour is similar to the one observed for oxDNA, although for $\sigma < 0$, no clear regime of intermediate post-buckling slopes is seen for intermediate force. Instead, irregular and non-monotonic behaviour is seen for $F > 3.0\,\mathrm{pN}$ and $\sigma < -0.07$.

The microscopic configuration corresponding to these states was investigated. It was found that instead of denatured regions, small tightly wound plectonemic structures tended to form. While at positive supercoiling, the common plectonemic structure was observed at all forces, for high stretching forces and undertwist, a four-stranded helical structure commonly developed in the system (see Fig. 3.11). No significant denaturation occurred, even for the highest applied stretching force $F = 6.0\,\mathrm{pN}$.

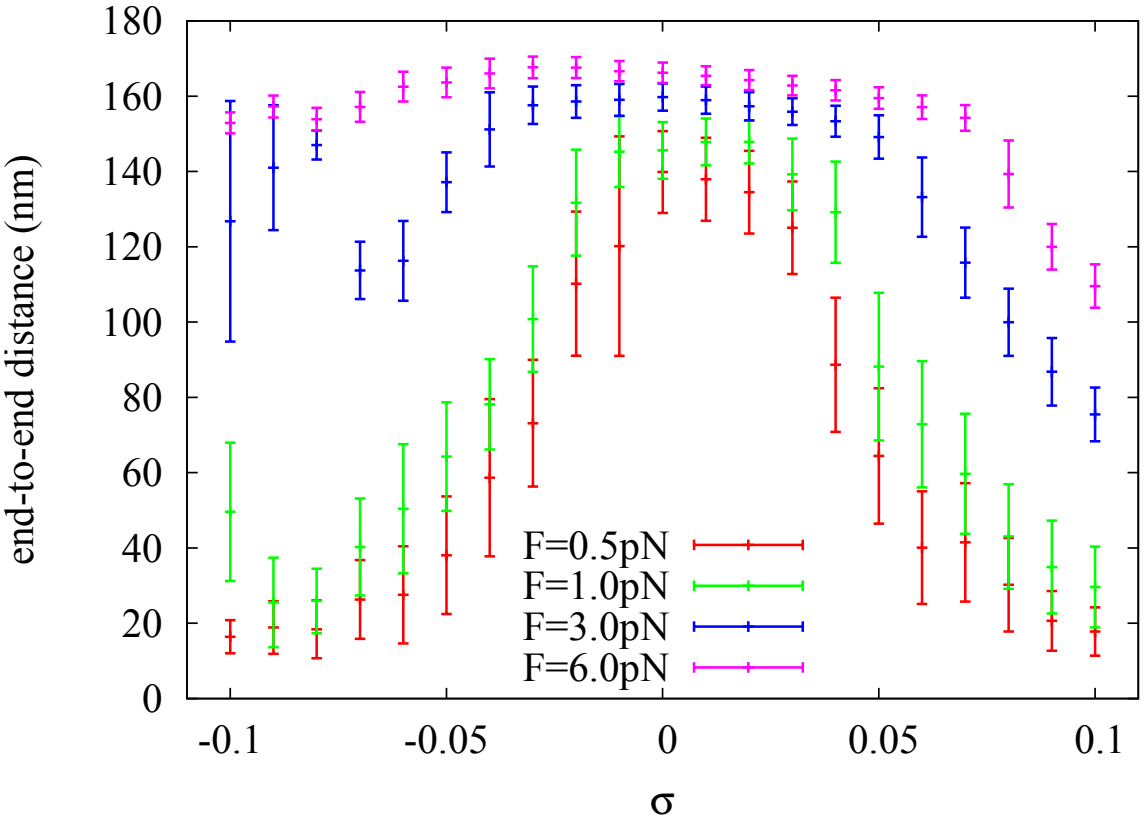


Figure 3.10: End-to-end extension curves for RNA molecules. The overall behaviour is similar to the one observed for DNA, although for negative supercoiling, a densely wound four-stranded structure is observed instead of significant duplex denaturation (cf. Fig. 3.11).

It is unclear if this microstate is a structurally realistic one. It may well be that it is only possible because of the interactions and geometry of the A-helix as represented in oxRNA. In addition, the model is parametrised at a relatively high monovalent salt concentration of 1M [182], at which electrostatic interactions are highly screened. It is likely that for a parametrisation at lower salt concentration, the four-stranded geometry depicted in Fig. 3.11 would be strongly disfavoured due to repulsive interactions of electrostatic origin between the RNA backbone sites. So far, no full hat curves for RNA have been published in the literature, making a direct comparison to experimental data difficult.

The torque responses predicted by the oxRNA model are shown in Fig. 3.12(a). Again, the overall response behaviour is similar to DNA. For small values of $|\sigma|$, a linear regime is observed which leads over to a constant post-buckling torque

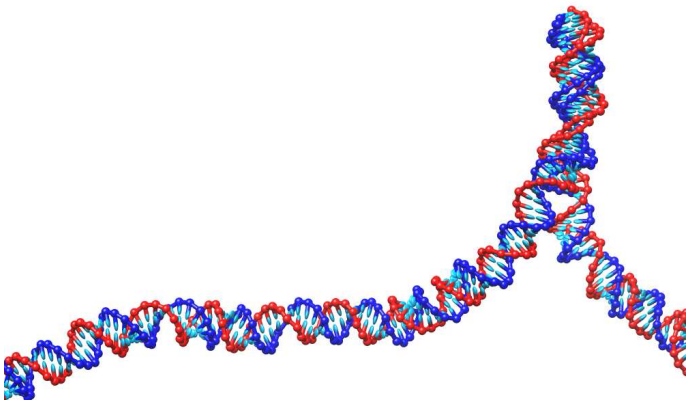


Figure 3.11: Microscopic view of a tightly wound four-stranded microscopic RNA structure, embedded in a long, A-helical double stranded segment. This state is observed in simulations at negative supercoiling and $F \geq 2$pN.

for positive supercoiling. At $\sigma < 0$, deviations from linear behaviour indicate the transition to molecular structures of the type shown in Fig. 3.11.

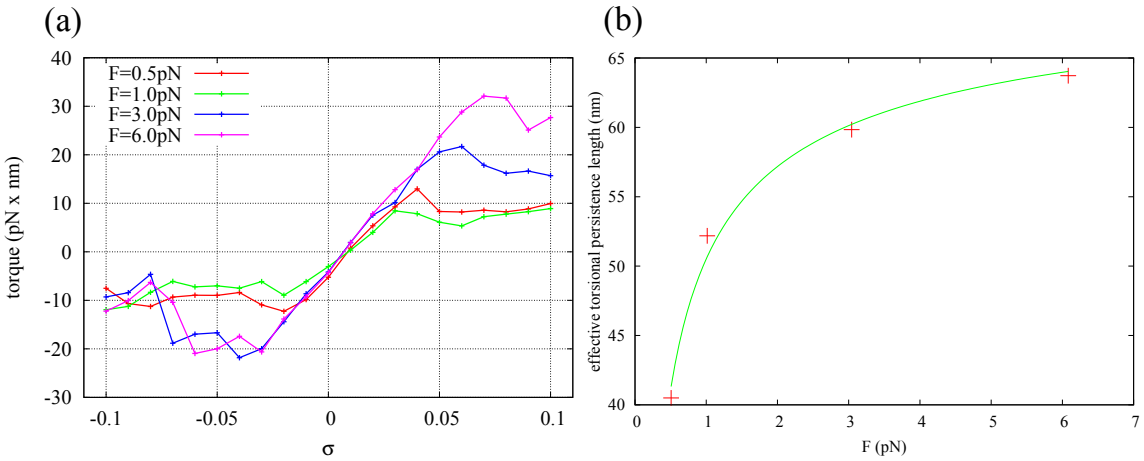


Figure 3.12: Mechanical behaviour of RNA under torsion: (a) Torque response for different values of F. As for DNA, a linear regime is followed by a overshoot and a constant post-buckling torque. (b) Red crosses: Effective torsional persistence lengths at different forces determined from the slopes of the linear regime in (a). Fit according to Eq. 2.3 is shown in green.

The overall similarity of hat curves and torque response in spite of marked differences in the microscopic configurations of the molecule highlights the fact that while hat curves provide important information on the overall behaviour of the system, they do not allow drawing direct conclusions as to the microstructure leading to this global behaviour.

The torque curves shown in Fig. 3.12(a) allow determining the effective torsional stiffness of the molecule, and hence fitting the torsional and bending persistence lengths for RNA using Eq. 2.3 in the same way as described in Sec. 3.2 for DNA (cf. Fig. 3.12(b)). From the forces for which the hat curve was determined, the fitting yields values $A_0 = 24.3\text{nm}$ for the bending persistence length and $C_0 = 73.2\text{nm}$ for the torsional persistence length. This measurement of the bending persistence length A_0 is consistent with a previous determination using the force-extension curve of the torsionally unstressed system, which yielded $A_0 = 26\text{nm}$ [182]. Compared to results from recently published single-molecule experiments by Chou *et al.* [155], which reported $A_0 = 57\text{nm}$ and $C_0 = 100\text{nm}$, both values take lower values in the oxRNA model. Nevertheless, both values are consistent with experimental values which lie in the same order of magnitude up to a factor of 1.5 to 2.

For C_0, slightly closer agreement with experiment is observed for oxRNA compared to the HelixMC model reported in Ref. [155], which yielded a torsional persistence length of $C_0 = 53\text{nm}$. The consequences of these differences for the accuracy of the predicted double-strand buckling behaviour can be estimated in an approximate fashion using Eq. 3.3, which predicts the critical value of σ at which the buckling transition is expected due to the elastic behaviour of the system. As $\sigma_b \propto \sqrt{A_0}/C_0$, the position of the buckling point only depends on the combination of both quantities, leading to cancellation of the deviations in the model properties. Indeed, Eq. 3.3 predicts $\sigma_b = 0.062$ at $F = 6\text{pN}$ using the experimental quantities and $\sigma_b = 0.056$ using the properties of oxRNA determined in the present section, corresponding to a difference of only 10% in σ_b. Hence, even though the individual values of the persistence lengths deviate from the experimental ones, the basic physics of RNA buckling can be expected to be well reproduced by the model.

The value of twist-stretch coupling in dsRNA molecules has recently been determined by single-molecule experiments as $D = +11$ [155], indicating that dsRNA does not exhibit the counter-intuitive negative twist-stretch coupling observed in DNA. As can be seen from the hat curve slopes in the pre-buckling regime at small

$|\sigma|$, oxRNA might be expected to reproduce this feature (see Fig. 3.10). Determining the slope at $F = 6.0$pN by fitting the hat curve for $-0.01 \leq \sigma \leq +0.03$ yields $\frac{dL}{d(\Delta Lk)} \approx -1.7$nm/turn, corresponding to $D \approx 7$ for oxRNA. Hence, oxRNA reproduces the correct sign of the twist-stretch coupling in RNA, while underestimating the experimentally determined magnitude of twist-stretch coupling by about 40%.

3.5 Summary

The oxDNA model was subjected to a systematic test of its mechanical response to twist and tension. Comparison with results from single-molecule experiments show that the model reproduces the supercoiling behaviour well. In particular, buckling of the double strand is observed at levels of supercoiling predicted by continuum theories of linear elasticity. Additionally, significant denaturation increasing with stretching force is observed for negative supercoiling. This allows one to reproduce the full "hat curves" observed in experiment with quantitative agreement, albeit that denaturation takes place at slightly higher force than observed in magnetic tweezer assays. This may partially be due to sequence heterogeneity in experiments which would favour denaturation at AT-rich segments, that are more likely to be present in long strands.

Similarly, the torque response of the model follows the behaviour expected from the Moroz-Nelson model for the interplay of twist and bending modes. Experimental torque curves are reproduced quantitatively, including the torque overshoot at buckling, and post-buckling levels of supercoiling. The twist-stretch coupling for small levels of supercoiling possesses the wrong sign in oxDNA, and the anomalous behaviour observed experimentally is not reproduced. However, the effect is very small compared to the other relevant couplings determining the physics of twisting and bending, and is therefore very unlikely to influence the overall buckling behaviour of the simulated system.

Less experimental data on double-stranded RNA is available to test the supercoiling behaviour of oxRNA [156]. Recent data from magnetic tweezer exper-

iments suggests that oxRNA underestimates both twist and bending persistence length. Similar difficulties were encountered in a recent coarse-grained RNA model by Chou *et al.* [155], where even lower values for the torsional persistence length were reported. However, as the buckling point only depends on the relative size of these two parameters, the effect of this difference on the overall extension behaviour is minor. In contrast to the behaviour in DNA, twist-stretch coupling has recently been found to be negative in RNA. This behaviour is well reproduced by the oxRNA model, up to a quantitative difference in the coupling magnitude of about 40%.

Chapter 4

Plectoneme tip-bubbles

While the overall mechanical behaviour of the molecule, as studied in the previous chapter, can be directly compared to single-molecule measurements, its full geometrical structure remains inaccessible in these assays. Having established that the global response of DNA to mechanical stress is captured to a high degree of accuracy by oxDNA, this chapter investigates the microscopic structure that gives rise to the observed overall behaviour.

A central theme is the interplay between the two most common response modes of dsDNA to mechanical stress, plectonemic supercoils and denaturation bubbles. Co-localisation of both structures is described in Sec. 4.1 as a function of twist and stretching force. The effects of this co-localisation on the distribution of supercoiling are explored in Sec. 4.2. In Sec. 4.3, experimental tests of these effects are proposed. Finally, in Sec. 4.4, several biological implications of this coupled behaviour of supercoiling and denaturation are proposed.

4.1 Tip-bubbles

4.1.1 Co-localisation

For negative supercoiling and intermediate stretching force $1.0\,\mathrm{pN} \lesssim F \lesssim 2.5\,\mathrm{pN}$, plectoneme formation and double strand denaturation can happen at the same time, leading to the intermediate slope in the extension curves, as was found in various different experimental assays (see Sec. 3.1). This intermediate regime was previously

interpreted as consisting of two distinct, spatially separated phases, one corresponding to a denatured and the other to a plectonemically coiled region of the double strand [3, 2, 63]. In contrast to this simple picture, simulations presented in this thesis show that both structural features do not occur independently, but are highly co-localised; for all simulations in the relevant parameter regime, denaturation bubbles were found to accumulate in the end-loop of the plectoneme, as shown in the reference structures in Fig. 4.1. Physically, this behaviour is plausible, because plectoneme end-loops represent highly bent regions of the double strand, where bending stress might be expected to favour bubble opening. Furthermore, denatured regions of the DNA strand possess a significantly enhanced flexibility, making them preferred sites for plectoneme nucleation. Co-localisation of bubbles and plectonemes also explains better the experimentally observed spectrum of end-to-end fluctuations as a function of σ and F, as will be shown in Sec. 4.1.3.

Reflecting this behaviour, a strong correlation is observed between the centre of denatured regions and the plectoneme position determined using the plectoneme detection algorithm introduced in Sec. 2.1.3.2. The configuration shown in Fig. 4.1 thus represents a novel conformational regime of DNA, a tip-bubble plectoneme, which arises from the coupling of denaturation and writhing in DNA.

To characterise this novel conformational state, a microscopic order parameter has to be defined. As suggested by the reference configuration shown in Fig. 4.1, a key criterion of this state is that the plectoneme position and the position of the midpoint of the largest denaturation bubble in the system coincide within a small margin M_c. Using this observation, a tip-bubble plectoneme can be defined as follows:

- A denaturation bubble with size $l_b \geq 2$bp exists in the system.
- A plectoneme is detected in the system using the algorithm described in Sec. 2.1.3.2.
- The bubble midpoint and the plectoneme position are separated by less than M_c bp

The restriction to bubbles with size $l_b \geq 2\text{bp}$ was introduced in order to discard short-lived single base-pair denaturations, which can occur in the strand due to thermal noise. For a large majority of configurations, the bubble-plectoneme distance is less than 20bp, so that chosing $M_c = 20\text{bp}$ captured almost all microscopic configurations, as shown in Fig. 4.2.

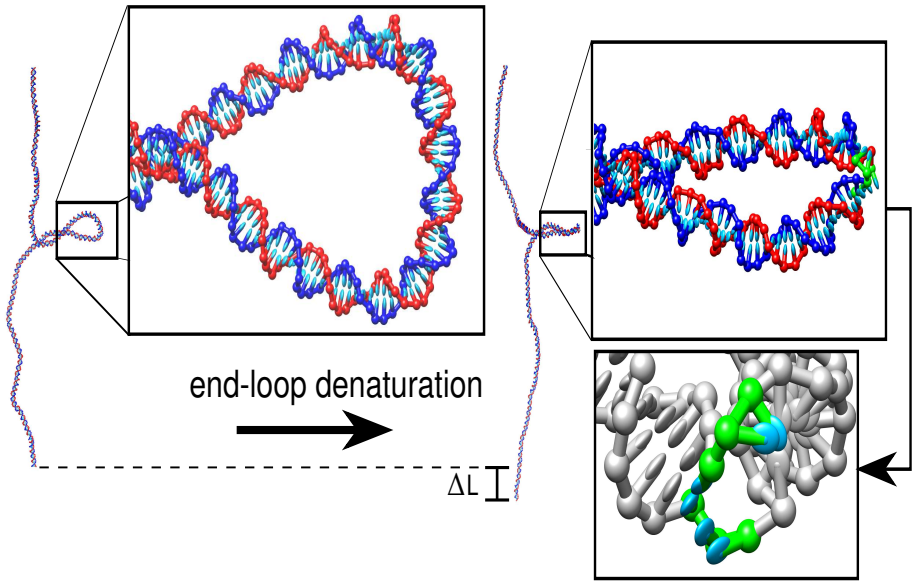


Figure 4.1: Typical microscopic formation of a tip-bubble plectoneme by an end-loop denaturation. Base-pair opening allows for a tighter winding of the end-loop, leading of a reduction ΔL in end-to-end length.

4.1.2 Populations and state diagrams

Based on the microscopic criteria for the plectoneme, bubble and tip-bubble plectoneme structures defined in Secs. 2.1.3.2, 2.1.3.1 and 4.1.1 respectively, prevalence of these states as a function of superhelical density σ and stretching force F on the system can be studied. A systematic survey of the likelihood for a given structure to be present in the system as a function of these external parameters is shown in Fig. 4.3. In all simulations, at most one plectoneme was observed, as is expected at high salt concentration and short strand length from continuum-level theories of plectoneme formation taking into account explicit electrostatics [193], and as was observed in recent single molecule experiments [1].

For $\sigma < 0$, there is a clear crossover from pure plectonemes to tip-bubble plectonemes and then to pure bubbles as a function of force. For $\sigma > 0$, a broader crossover to the tip-bubble regime occurs, which happens at significantly higher forces than for negative supercoiling. The reason for this difference is that for negative supercoiling the tip bubble is able to absorb some of the undertwist imposed on the system. This is not the case for the defect forming in the end-loop at $\sigma > 0$.

At low values of $|\sigma|$, the recognition of small plectonemes somewhat depends on the cutoff values chosen in the plectoneme detection algorithm, as described in Sec. 2.1.3.2. As shown in Fig. 4.3, the transitions in this region of the state diagram are narrow as a function of F and σ. Therefore, changes in the cutoff values have only very small effect on the values of σ for which the onset of the different regimes is detected.

Note that as long as σ is large enough to allow stable tip bubbles, the transition from plectonemes to tip bubbles appears to be at a very similar force for different σ. The reason for this is that the crossover is mainly determined by a change in the end-loop structure, rather than in the rest of the plectoneme, which grows for

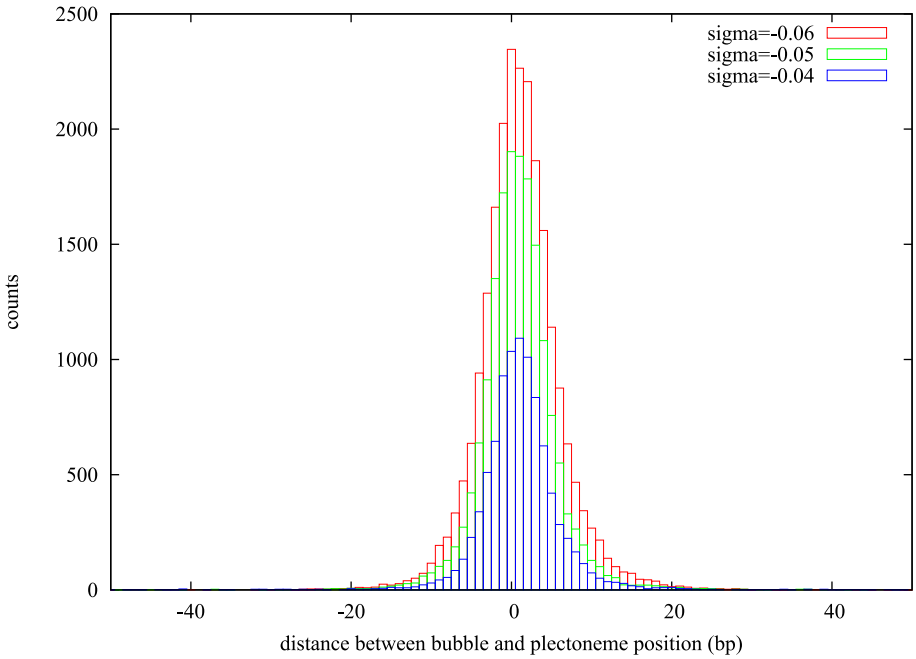


Figure 4.2: Distribution for the distance between bubbles and plectoneme midpoints for simulations of a 600-bp system, at different values of σ for $F = 1.27$pN. Strong co-localisation is observed.

increasing σ (see Fig. 3.2). Similar arguments may explain the observation that the transition from tip bubbles to extended bubbles also happens at the same force for different negative values of σ.

Combining the results shown in Fig. 4.3 allows one to construct a state diagram of DNA, which gives the dominant state in the system as a function of σ and F. To generate the state diagram in Fig. 4.4, a threshold value of 40% was used to determine the boundaries between different configurational states. Also in this case, changes in the threshold introduce only slight modifications to the precise position of the boundaries, but do not affect the overall structure of the state diagram.

In sum, tip-bubble plectonemes thus represent a separate intermediate state between extended bubbles and pure plectonemes, rather than a pure crossover regime between these two states. Further dynamical evidence for the presence of boundaries between the tip-bubble plectoneme state and the extended bubble and pure plectoneme state is given by enhanced fluctuations close to the boundaries, and will be discussed in chapter 5.

4.1.3 Free-energy profiles for bubbles and plectonemes

Bubbles in the end loop tend to shorten plectonemes due to three effects, thus causing interconversion between bubbles and plectonemes:

- Denaturation bubbles possess very small native twist, and are therefore able to absorb undertwist relative to the double-stranded state.

- Denatured single strands can twist back on themselves in a negative way, thus even leading to a negative twist contribution. As the overall linking difference ΔLk of the system is conserved according to the relation $\Delta Lk = \Delta Tw + \Delta Wr$ (see Sec. 2.1.2), this increase in ΔTw changes the system's writhe component and thus the average plectoneme size.

- A denaturation in the end loop constitutes a defect in the double helical structure, which allows the formation of a smaller, more tightly wound tip of the

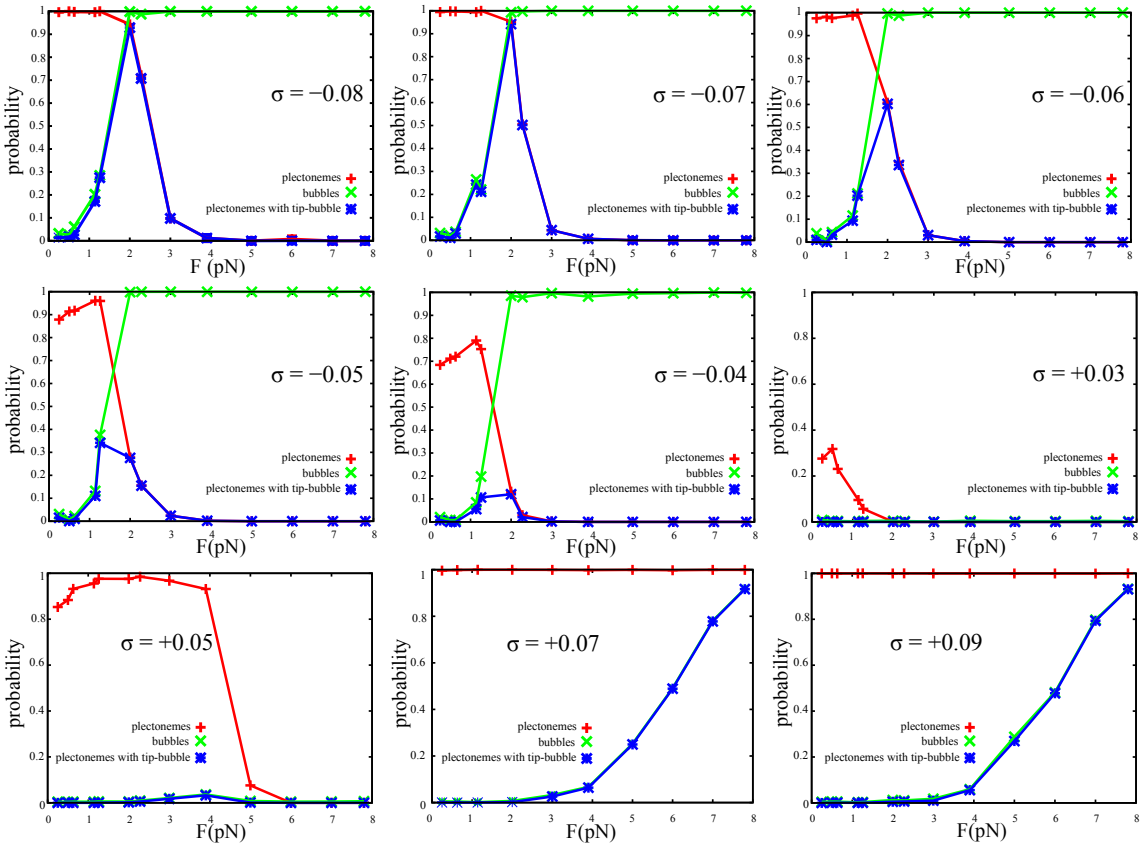


Figure 4.3: Population diagrams as a function of force for different imposed values of σ. For $\sigma < 0$, note the occurrence of a tip-bubble regime for intermediate values of F. At $\sigma > 0$, note the wide crossover to tip-bubble plectonemes for increasing F. No pure bubble regime exists for $\sigma > 0$ at the force values used here. However, other forms of DNA may occur at sufficiently larger force.

plectoneme structure, as shown in Fig. 4.9, thus causing an extension ΔL of the overall system.

Fig. 4.5 shows free-energy landscapes of the system as a function of bubble and plectoneme size for different values of F. In a 600-bp strand, due to thermal noise, very short-lived denaturations of small size can occur in a duplex at $T = 300\text{K}$. In order to separate out this contribution, bubbles with a size up of 1 or 2bp were only taken into account if they were co-localised with a plectoneme in this analysis.

Fig. 4.5 also shows that the initial formation of a 2-bp tip bubble causes a decrease in plectoneme size by roughly 100bp. Assuming that the equivalent plectoneme size is converted into extended strand length aligned with F and neglecting changes in bending energy, the free energy gain due to this extension is $\Delta G =$

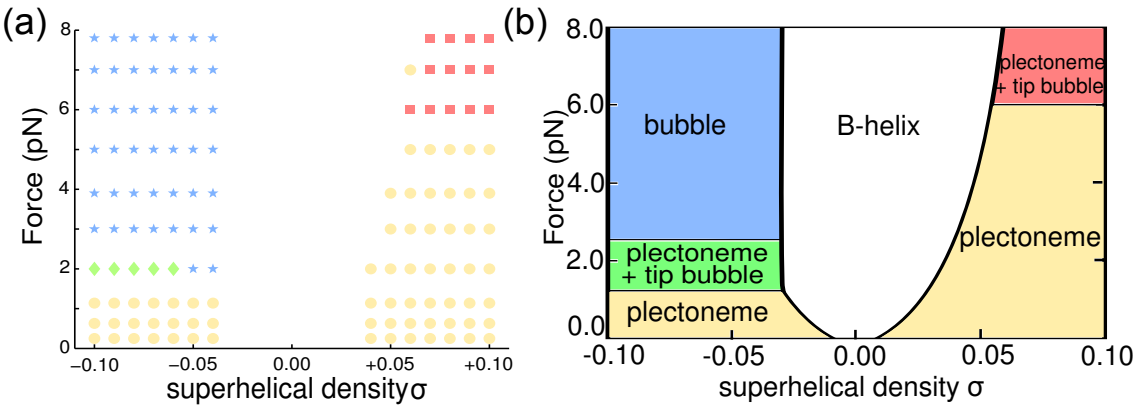


Figure 4.4: State diagram based on the populations shown in Fig. 4.3. (a) Dominant state exceeding a cutoff of 40% on a grid in the $\sigma - F$ plane. Yellow circles, blue stars, green diamonds and red squares signify pure plectonemes, extended bubbles, and tip-bubble plectonemes for negative and positive supercoiling respectively. Note that at small supercoiling, the detection of plectonemes will be somewhat cutoff-dependent as discussed in Sec. 2.1.3.2. (b) Semi-schematic interpolation of the state-diagram shown in (a).

$F\Delta L \approx 1.27\text{pN} \cdot 100 \cdot 0.34\text{nm} \approx 10 k_B T$ for $T = 300\text{K}$. This free energy gain is on the same order of magnitude as the cost of forming a small double strand denaturation in oxDNA. After denaturation of the first two base pairs and consequent shape change of the end-loop, the size of plectonemes and bubbles can be interconverted in an approximately linear fashion.

Note that this behaviour markedly differs from the one predicted by simple models that treat bubbles and plectonemes as separate structures which have been proposed in the literature [3, 63]. To describe the interconversion between these structures in a consistent way, the substantial nucleation costs of both bubble and plectoneme structures have to be considered [135, 4, 1, 145]. The free-energy landscape would then show two separate minima separated by a considerable barrier, representing bubble and plectoneme states respectively. Due to the co-localisation behaviour described in this chapter however the presence of a bubble reduces the nucleation cost of a plectoneme and vice versa, leading to the free energy landscapes shown in Fig. 4.5 which are flat along their diagonal. This feature of the free-energy landscapes for bubble and plectoneme sizes is related to the experimentally observed broad fluctuation spectrum of end-to-end distance, which will be discussed

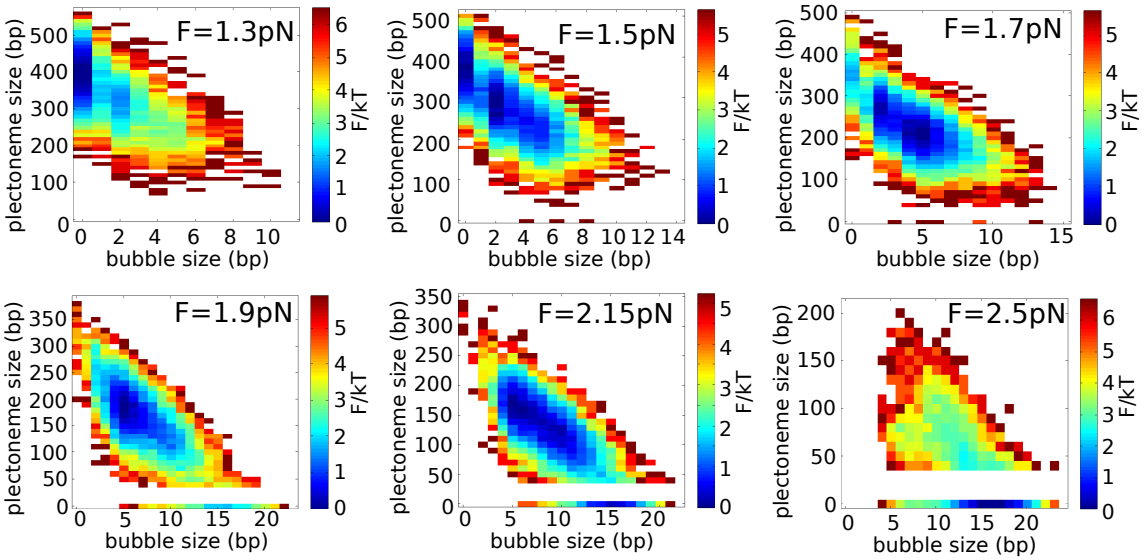


Figure 4.5: Free-energy landscapes as a function of bubble and plectoneme size at $\sigma = -0.08$ and different forces. A negative linear correlation of bubble and plectoneme size is observed after forming a kink-like tip-bubble defect of 1 to 2bp. The landscapes are obtained using the average-base parametrisation of oxDNA.

in chapter 5. Previous modelling approaches which did not take into account the co-localisation of bubbles and plectonemes could only reproduce this behaviour by neglecting the nucleation costs of bubbles and plectonemes. This strategy however contradicts known energetic features of these structures [135, 4, 1, 145], and furthermore allows formation of a large number of different independent plectoneme structures in the system, in contradiction to experimental observation [1].

Fig. 4.6 shows free energy profiles as a function of bubble size, corresponding to projections of the 2-dimensional free energy landscapes of Fig. 4.5 onto the x-axis. At $F \approx 1.5$pN, a 2-bp tip bubble denaturation starts to become favourable. This small initial denaturation is particularly stable, as it enables kinking and therefore tighter winding of the end-loop, as shown earlier in Fig. 4.1. As the force increases, it becomes more favourable to grow larger bubbles, as can be seen in the free-energy plots in Figs. 4.5 and 4.6. For large enough force, the plectoneme disappears, and the dominant states are extended bubble states. The thermodynamics of the crossover between tip-bubble plectonemes and extended bubbles is illustrated in Fig. 4.7, where the projection of Fig. 4.6 is separated into contributions due to plectonemes, and contributions due to bubble states without plectonemes. As discussed

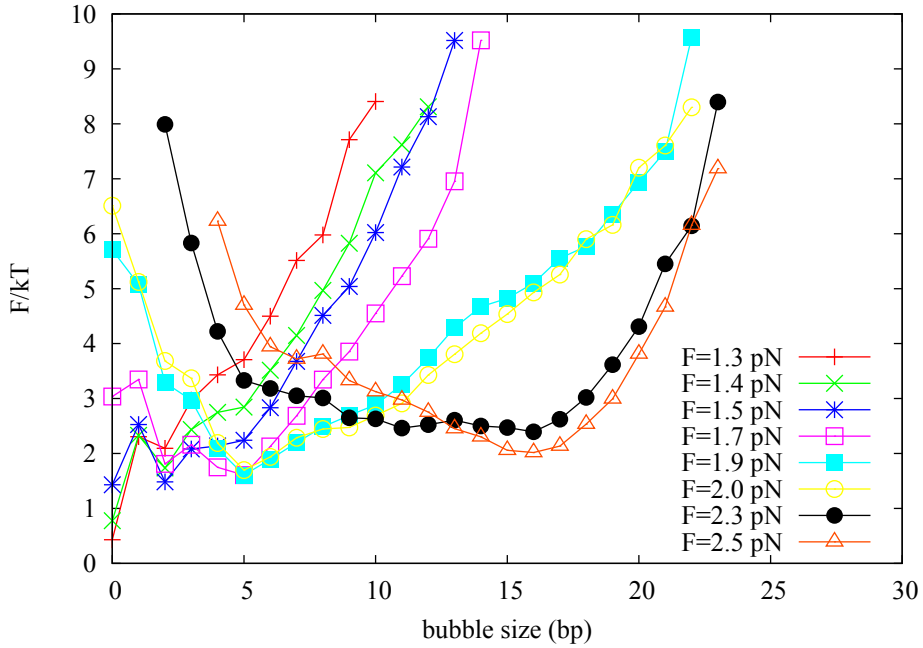


Figure 4.6: Free-energy profiles at $\sigma = -0.08$ and different values of F as a function of bubble size, corresponding to projections of the 2-dimensional free energy landscapes of Fig. 4.5 onto the x-axis. Note the stability of 2-bp bubbles at low force, due to free energy gains from initial end-loop rearrangement. For forces $F \gtrsim 2.3$pN, extended bubble states are more stable than plectonemic states, see also Fig. 4.7.

in Sec. 2.1.3.2, the classification of tip-bubble plectonemes of small size somewhat depends on the cutoff of the plectoneme detection algorithm. However, as the transition from tip-bubble plectonemes to bubbles is fairly narrow (cf. Fig. 4.4), this is not expected to significantly affect the force at which the crossover occurs. Fig. 4.7 clearly shows how, with increasing force, the system transitions from a regime with mainly tip bubbles, to a regime with mainly extended bubble states. At parameter values where extended bubbles and tip-bubble plectonemes can coexist, local kinks induced by bubbles provide a preferred site for re-nucleation of a plectoneme. Structures and a corresponding kymograph are shown in Fig. 4.9 to illustrate the crossover between bubbles and plectonemes for $\sigma = -0.09$ and $F = 2.3$pN. The average size of both bubbles and plectoneme structures as a function of F at $\sigma = -0.08$ is shown in Fig. 4.8, reflecting again the sharp boundaries between the different configurational regimes observed for the population diagrams in Fig. 4.4.

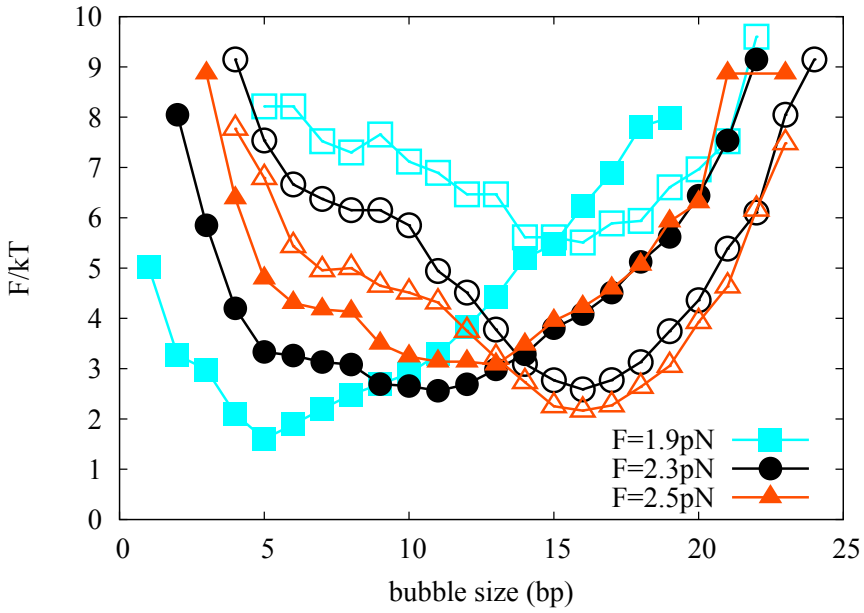


Figure 4.7: Free-energy plots that illustrate the thermodynamics of the crossover between tip bubbles and extended bubbles. The left curves (solid symbols) show the free energy of tip-bubble plectonemes at a particular size, while the right curves show the free energy of extended bubbles (open symbols) without a plectoneme. At small values of F, the tip bubble is more stable than the extended bubble. At $F \approx 2.3$pN, the extended bubble population has a free energy roughly equal to the tip-bubble plectoneme, indicating that above this force, extended bubbles become more favourable, as can be seen for the $F = 2.5$pN curves.

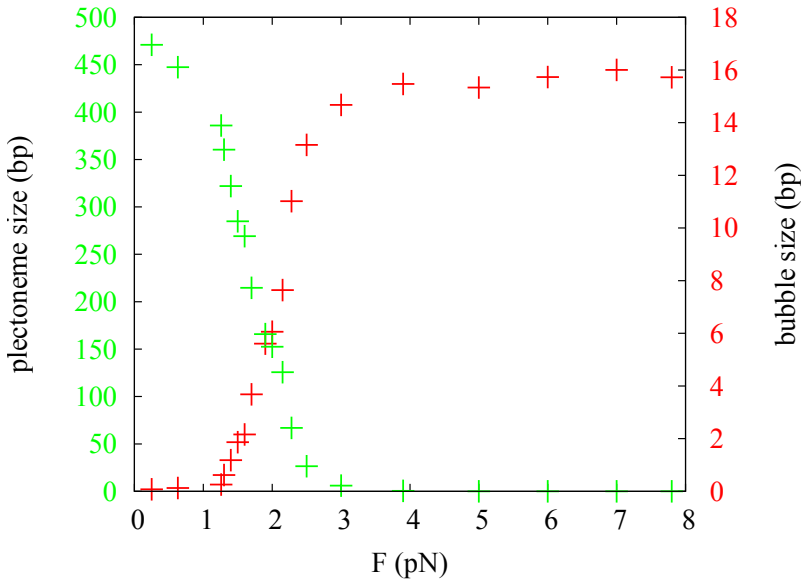


Figure 4.8: Mean sizes of plectoneme (green) and denaturation bubble (red) as a function of applied stretching force F at $\sigma = -0.08$.

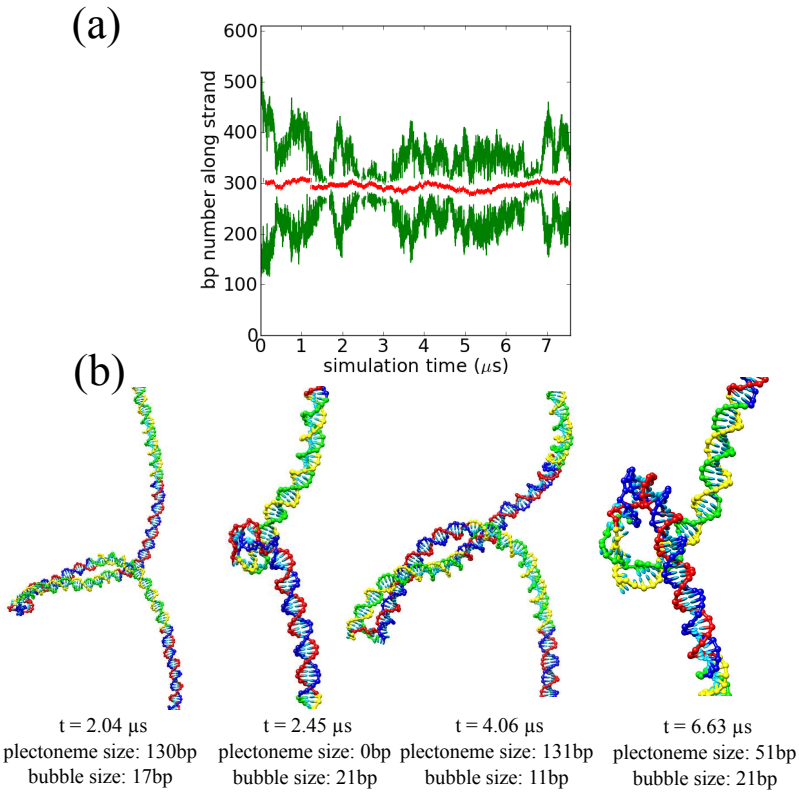


Figure 4.9: Interconversion between bubbles and plectonemes at $\sigma = -0.09$ and $F = 2.3$pN. (a) Position kymograph of plectoneme boundaries (green) and bubble centre (red). (b) Example configurations taken from trajectory shown in (a) at the times indicated. Different colouring of parts of the double strand is to facilitate comparison between structures.

4.2 Sequence effects in plectoneme localisation

4.2.1 Base-pairing energies

A particular consequence of the co-localisation of bubbles and plectonemes to form the tip-bubbble state is the coupling of sequence properties to the distribution of writhe along the double strand. Bubble formation is known to be highly sequence dependent [135, 26].

To explore the role of sequence in plectoneme positioning, simulations of the 600-bp system were performed at $\sigma = -0.06$ and $F = 1.27$pN using a sequence-dependent parametrisation of oxDNA. Two random sequences were used, a fully random sequence with a GC content of 49%, and a block-random sequence containing five 120bp stretches with a GC content alternating between 70% and 30%, and an overall GC content of 52% (the full sequences are given in Appendix C). To distinguish sequence-dependent from other effects such as the influence of boundary conditions, simulations using the sequence-independent average-base parametrisation were performed. All simulations were run for at least 60μs of simulated time, after which no significant asymmetries are discernable in the average-base plectoneme distributions (see Figs. 4.10 and 4.11). This criterion was taken as a sign of equilibration[1].

Even though the average GC content in both random sequences is similar, tip bubble prevalence is 59% for the randomised sequence and 72% for the block-random sequence, compared to 19% for the average-base model. These differences occur because bubbles can form more easily in weaker AT-rich regions. Furthermore, the presence of longer AT-rich stretches increases the overall denaturation in the strand by allowing denaturation to take place in a more localised way. The accumulation of plectonemic coils close to the denatured region will further contribute to this localisation of melting. As shown in Sec. 3.1.3, this effect is also manifest in the

[1]Note that the decay of position probability at the strand ends is due to the effects of the boundary conditions applied to the simulated system (cf. Sec. 2.4). In particular, the centre position of a plectoneme cannot be closer to the strand end than half its size as determined by the detection algorithm of Sec. 2.1.3.2.

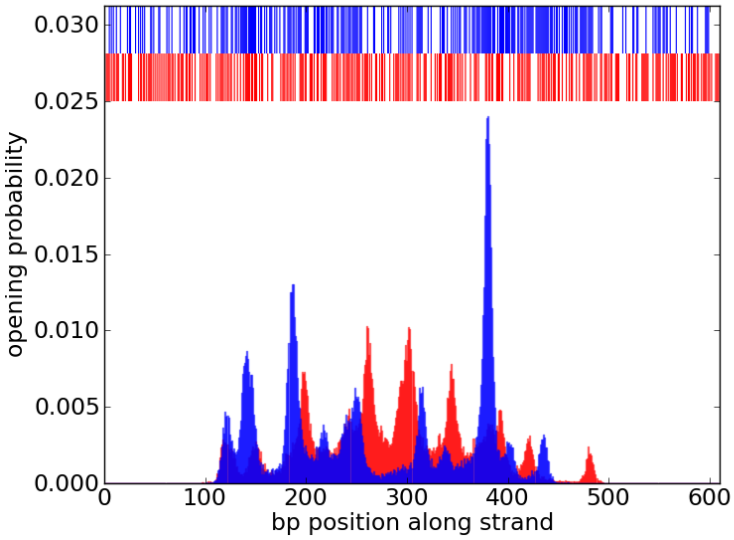


Figure 4.10: Position distribution of plectonemes at $\sigma = -0.06$ and $F = 1.27$pN. Results are shown in red for a random sequence, and in blue for a block-random sequence, as explained in the text. Plectoneme formation is suppressed near strand ends because these are clamped. For each sequence, the upper part of the figure shows coloured positions for AT basepairs and white for CG basepairs. Plectonemes in the tip-bubble regime strongly localise to AT-rich regions.

hat-curves: The random sequence deviates from the $\sigma \to -\sigma$ symmetry at lower forces than the average-base model does.

The plectoneme distribution found for the random and block-random sequences is shown in Fig. 4.10. Plectonemes can be seen to exhibit a strong tendency to accumulate in AT-rich, weak sequences, as a consequence of the smaller enthalpic cost of tip-bubble formation in these regions. In particular, the two AT-rich flanking regions of the block-random sequences show a higher plectoneme population despite being closer to the boundary walls than the central region, where the plectoneme density is suppressed by a GC-rich sequence.

To demonstrate the effect of the base sequence in a biologically relevant sequence, the plectoneme position distribution on a strand containing the promoter sequence of the human TP53 gene was obtained (see Appendix C for the full sequence used). The resulting plectoneme distribution is shown in Fig. 4.12, showing a plectoneme distribution markedly different from the average-base model also in this case.

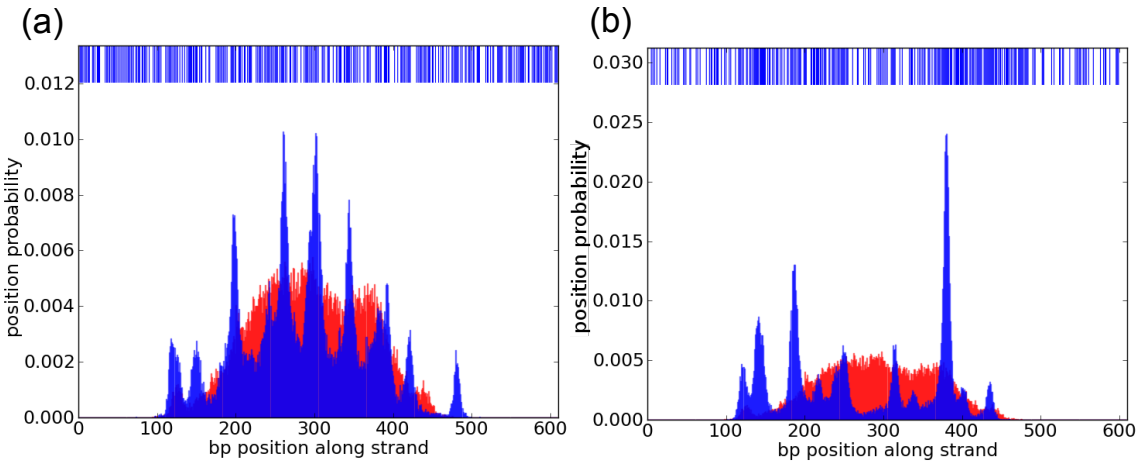


Figure 4.11: Plectoneme position distribution for the random sequences and in the average-base parametrisation at $\sigma = -0.06$ and $F = 1.27$pN. Sequences are given in the upper part of the figure, where a blue line indicates an AT base-pair. (a) Random sequence in sequence-dependent (blue) and average-base (red) parametrisation. Mean plectoneme size is 260bp for the average-base parametrisation and 195bp for the random sequence. (b) Block-random sequence in sequence-dependent (blue) and average-base (red) parametrisation. Mean plectoneme size is 260bp for the average-base parametrisation and 137bp for the block-random sequence.

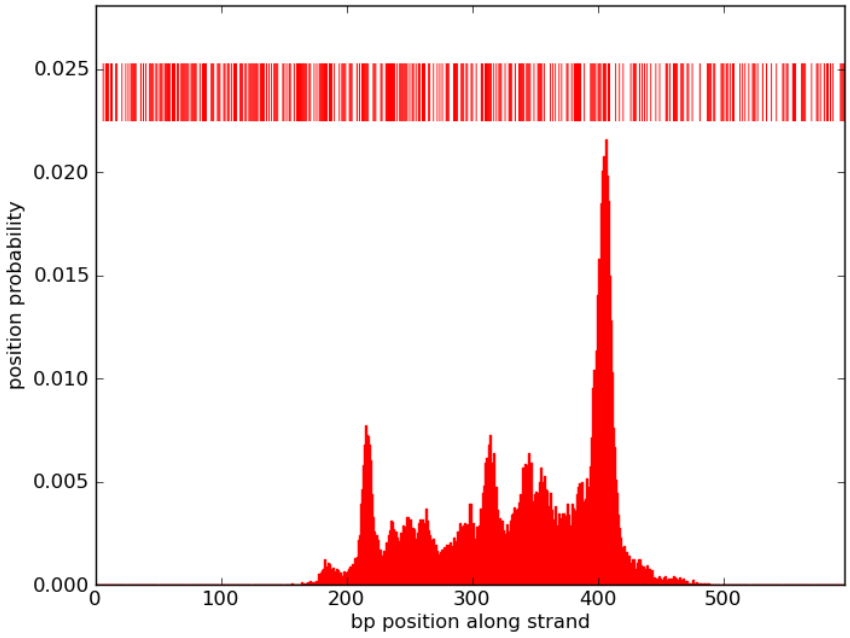


Figure 4.12: Plectoneme position distribution for a sequence containing the human TP53 gene promoter sequence.

4.2.2 Mismatches

The differences in the distribution of plectoneme positions described in Sec. 4.2.1 are due to sequence-dependent differences in the base-pairing and stacking energies, leading to a sequence-dependent bubble opening cost in a double strand consisting of otherwise fully complementary single strands.

A physically similar situation can be expected for a double strand containing mismatches, i.e. sites at which two opposing nucleotides have identities that do not allow them to form a base pair. This means that no additional free-energy cost for denaturing the double strand at the mismatch site arises, making the mismatch particularly favourable as the location of a tip-bubble. Furthermore, the high intrinsic flexibility of the unpaired region will favour formation of the plectoneme close to the mismatch.

In analogy with the sequence dependence reported in Sec. 4.2.1, Fig. 4.13 shows the plectoneme distribution for an average model with 2-bp mismatches, which have a biological relevance (see Sec. 4.4). To test the influence of boundary conditions, the mismatches were placed in the central and flanking regions of the double strand. In both cases, clear localisation of plectonemes close to the mismatch is observed.

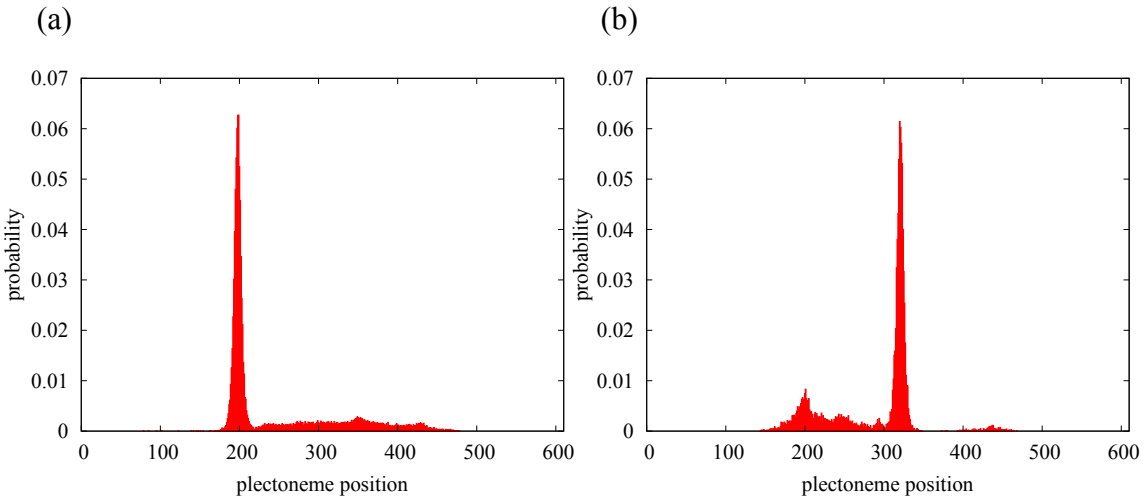


Figure 4.13: Position distribution of plectonemes for 2-bp mismatches of an otherwise fully complementary double strand in the average-base parametrisation.
(a) Mismatch on base pair number 200 and 201.
(b) Mismatch on base pair number 323 and 324.

4.2.3 Further effects of secondary structure

Besides tip-bubble denaturation, it is conceivable that other types of secondary structure may occur in the plectoneme end-loop, which might have an effect on the overall free-energy cost of positioning a plectoneme at a given location. Most of these possible secondary structures will require a particular sequence to form. For example, when the end-loop contains a palindromic sequence, cruciform formation could occur [174]. Results from continuum rod models [123] indicate that the effect of plectoneme localisation primarily depends on non-homogeneous local flexibility of the DNA double strand. Most secondary structures will locally enhance the flexibility of the double strand, thus still favouring plectoneme localisation at their position.

In this work, non-symmetric bonds were disallowed for sampling efficiency and to focus on generic denaturation effects, excluding such secondary structure effects (cf. Sec. 2.1.3.1). However, given that the denaturation tip-bubbles observed in the systems studied here typically consist of small bubbles containing only few base pairs, the effect of secondary structure is expected to be negligible for most sequences.

Even in the presence of larger bubbles, double strand denaturation might be expected to be the dominant response of the system, before forming secondary structure on a longer timescale. Therefore, the plectoneme position distributions presented in this section would remain relevant at least in this intermediate phase.

4.3 Experimental tests of plectoneme localisation

The tip-bubble plectoneme state couples sequence-dependent features of the DNA double strand to the distribution of plectoneme structures, which are central to the writhing behaviour of the strand. This constitutes a coupling between local and global strand properties, which are accessible in experiments by different techniques. In this section, experimental strategies using the methods reviewed in Sec. 1.3 are outlined to test the structural predictions based on the simulation results discussed in this section.

- Tip-bubble plectonemes lead to a stabilization of open bubbles in specific weak sequences. These could be detected by standard oligonucleotide hybridization probes in magnetic tweezers, as used before e.g. by Strick *et al.* [194]. This test does not demonstrate directly the presence of plectonemes, but allows testing for stable denaturation of a specific sequence for given σ and F, which are more abundant as a consequence of tip-bubbles.

- AFM techniques (see Sec. 1.3.1) could be used to directly image tip-bubbles, as was done in closed-circular systems using single-strand binding proteins by Jeon and co-workers [44]. To avoid artifacts in the imaging process, it may be possible to keep the plectoneme fixed by DNA bridge proteins, which have been explored experimentally before [195, 196].

- Multi-colour FRET (see Sec. 1.3.3) provides a means to measure several distances in a molecular system in parallel. Using this technique, it may be possible to detect the presence of a plectoneme stem and a bubble at the same time: Two distant sites along the double strand are brought into close contact, while two sites on opposing single strands are further apart than is compatible with a double-stranded structure, indicating denaturation [197].

- To directly probe the presence of a tip-bubble in a system held under mechanical stress, it would be necessary to combine an experiment that includes a molecular tweezer providing spatial resolution such as the assay by van Loenhout *et al.* with a molecular probe for double-strand denaturation. Such probes could consist of fluorescence-labelled single strand binding proteins or short oligonucleotides. This test would furthermore allow one to directly probe the dynamics of tip-bubble plectonemes.

- Existing single-molecule experiments have already provided indirect evidence for the tip-bubble plectoneme state, for example from the behaviour of fluctuations in end-to-end distance, as discussed in Chapter 5.

4.4 Biological implications

As shown in Sec. 4.1.2, end-loop denaturation in DNA happens in a regime of about 1-2pN of stretching force, and a superhelical density of $\sigma \lesssim -0.04$. These are parameters which are in the range of mechanical stresses experienced by DNA in living cells. The genome of most organisms is negatively supercoiled with a superhelical density $\sigma \approx -0.05$ [37] [2], which is well within the tip-bubble plectonemic regime as shown in the state diagram in Fig. 4.4. In some cases, namely genome organisation, transcription and DNA repair processes, more specific biological consequences of tip-bubbles might be imagined, and are discussed in this section.

When comparing simulation results to biological conditions, it has to be remembered that the version of oxDNA used in this work was parametrised at a monovalent salt concentration of 500mM (see Chapter 2). This value is considerably higher than in typical *in vivo* situations, where the intracellular concentration of monovalent ions is typically in the range of 150mM. A main consequence of this difference is that electrostatic interactions are less screened *in vivo* that in the model parameterisation conditions. Therefore, electrostatic repulsive interactions between the negatively charged DNA backbones are expected to increase. Physically, this is expected to lead to a decrease in denaturation free energies [131], and an increase in the plectoneme stem radius [120]. As shown in Sec. 3.1, this leads to a lowering in the force at which bubbles start to form, causing a shift in the boundaries of the tip-bubble plectoneme region.

Apart from such corrections in quantitative detail, however, the physics of tip-bubbles is set by the generic denaturation and buckling behaviour of DNA, which are well established and known to be relevant under *in vivo* conditions. Therefore, it might be expected that the basic mechanism of end-loop denaturation is indeed relevant in a biological scenario.

[2] In most bacteria, the genome is circular and intrinsically conserves supercoiling, while in eukaria, supercoiling is thought to be conserved by anchoring of chromosomes in the nuclear matrix [198].

4.4.1 Genome organisation

As demonstrated by the plectoneme distributions shown in Figs. 4.10–4.12, the geometry of tip-bubble plectonemes has the effect of guiding plectonemes to weak parts of the DNA sequence. Thus, by coupling to the position of plectonemic coils, sequence properties can influence genome organisation in a more global way than by denaturation of base pairs. The three-dimensional organisation of the genome determines the microscopic accessibility of the base sequence, and therefore has a regulatory relevance in virtually all processes involving its readout [199, 198].

For prokaryotes in particular, which lack the wrapping of DNA around histones, supercoiling is known to be a major player in the compaction of the DNA strand [200], with plectonemes being the dominant DNA conformation [201]. Electron microscopic studies have suggested that the position of plectonemic supercoils in close-circular DNA is non-random [202]. Tip-bubble plectonemes are a possible microscopic explanation of this positioning. Other effects not captured by oxDNA, such as sequence-dependent intrinsic bending [187, 202] are likely to add to that picture, and introduce further, more specific sequence-dependence into genome organisation.

4.4.2 Transcription

Transcription by RNA polymerases is a process well characterised at the bulk and single-molecule levels [203, 71, 204]. In particular, plectonemic supercoils are believed to be involved in the onset of transcription in both prokaryotic and eukaryotic organisms [201]. Structurally, a study by ten Heggeler-Brodier *et al.* using electron microscopy suggests that RNA polymerase indeed preferentially localises to the apical tips of DNA coils at transcription [204]. This behaviour is fully consistent with the findings on plectoneme localisation described in this chapter: As double-strand stability in the end-loop regions of plectonemic coils is significantly reduced, these sites might be expected to constitute favourable binding sites for RNA polymerase.

Furthermore, transcription start sites are known to contain AT-rich motifs such

as the TATA box [205, 206]. As shown in Sec. 4.2, tip-bubble plectonemes are likely to be centered around such long AT-rich stretches, increasing the probability of these sequence parts to be denatured and stably exposing them at a plectoneme apex, where they are easily accessible to binding by other molecules. Hence, the tip-bubble plectoneme regime may constitute a simple way for the organism to regulate transcription by global modification of the mechanical parameters F and σ.

4.4.3 DNA repair

In a biological setting, the DNA molecule is subject to permanent lesions by external influences such as UV irradiation or oxidative stress [3]. To handle this constant damage, both prokaryotic and eukaryotic cells have evolved complex repair mechanisms which correct damages [209]. Among the most common damages are pyrimidine dimers, which are induced photochemically [208] and cause local disruption of interstrand base pairing, enhancing the local flexibility of the double strand [210].

By using a mismatch to model these microscopic changes to DNA structure, this phenomenon can be related to the results of Sec. 4.13: Tip-bubble plectonemes are very likely to accumulate around local mismatches of the base sequence, as these parts of the strand do not incur a free-energy penalty for denaturing the double strand, and have a very high local flexibility. Even if the precise nature of the local disruption is likely to differ from this simple coarse-grained view, one might still expect that enhanced local flexibility causes plectoneme accumulation around the lesion. Local plectoneme accumulation might in turn be involved in the recruitment of other molecular species to the damaged site.

4.5 Summary

In this chapter, a novel conformational state of DNA, a plectoneme possessing a denaturation bubble at its tip, was described. It arises in the crossover regime between pure plectonemes and extended bubbles for negative supercoiling, and for high forces

[3] The frequency of DNA damage is believed to be in the range of 10^3 to 10^6 per day per cell [207, 208].

at positive supercoiling. The free-energy landscape for the crossover between the tip-bubble plectoneme state and its "neighbouring" states was calculated, showing that bubbles and plectonemes lower their mutual nucleation barrier, thus facilitating interconversion between the two motifs.

The presence of broken base pairs in the structure makes the free energy of the plectoneme sequence-dependent. Due to the coupling of denaturation and writhing behaviour, the sequence contributes to the localisation of plectonemes, and favours their positioning in weak, AT-rich stretches. Similarly, mismatches in the base sequence of the double strand provide positions where tip-bubble denaturation does not incur a denaturation enthalpy cost. Therefore, these mismatches act as preferred sites for tip-bubble plectoneme localisation.

Several experimental strategies to test for the presence of plectonemes possessing tip-bubbles are possible, some of which are similar to already established assays. Biologically, the guiding of plectonemic supercoils by tip-bubbles highlights the fact that intrinsic properties of the DNA molecule can determine its large-scale configuration far beyond double strand denaturation, even without external organising factors. Furthermore, tip-bubble plectonemes can help guide the recruitment of active molecules such as RNA polymerase to weak sequences, thus critically governing their binding kinetics, and helping them fulfill their biological function.

Chapter 5

Dynamical behaviour of plectonemes

Next to the overall mechanical behaviour of supercoiled DNA, the dynamics of plectonemes has of late become accessible in experiments. Apart from the time-resolved overall behaviour of DNA under torsion and tension, it has recently become possible to track individual plectoneme structures, and observe their displacement behaviour in real time [1].

In this chapter, the dynamics of plectoneme structures is studied. Sec. 5.1 describes fluctuations in the end-to-end distance of the strand, as well as their dependence on strand length. In Secs. 5.2 and 5.3, the continuous, diffusive and the discontinuous, "hopping" displacement modes of plectonemes are studied respectively.

5.1 Fluctuations and length dependence

5.1.1 Fluctuation spectrum

The magnitude of fluctuations in end-to-end distance is accessible in molecular tweezer experiments, and can give more detailed information about the energetics of the conformational states at given external conditions. In a series of pioneering experiments [3, 2], the fluctuation spectrum in L/L_0 was measured for the first time for systems of effective length $L_0 \approx 6\text{kbp}$ around the region for negative supercoiling, close to the crossover from plectonemes to a bubble state. In this section, analogous

results are presented for the system studied in coarse-grained simulations.

The red curve in Fig. 5.1 shows the magnitude of fluctuations in end-to-end distance of the strand for system size $L = 600\,\mathrm{bp}$ at a superhelical density $\sigma = -0.08$. At forces between 1 and 3pN, the 600-bp fluctuation spectrum exhibits enhanced fluctuations, and shows a two-peak substructure.

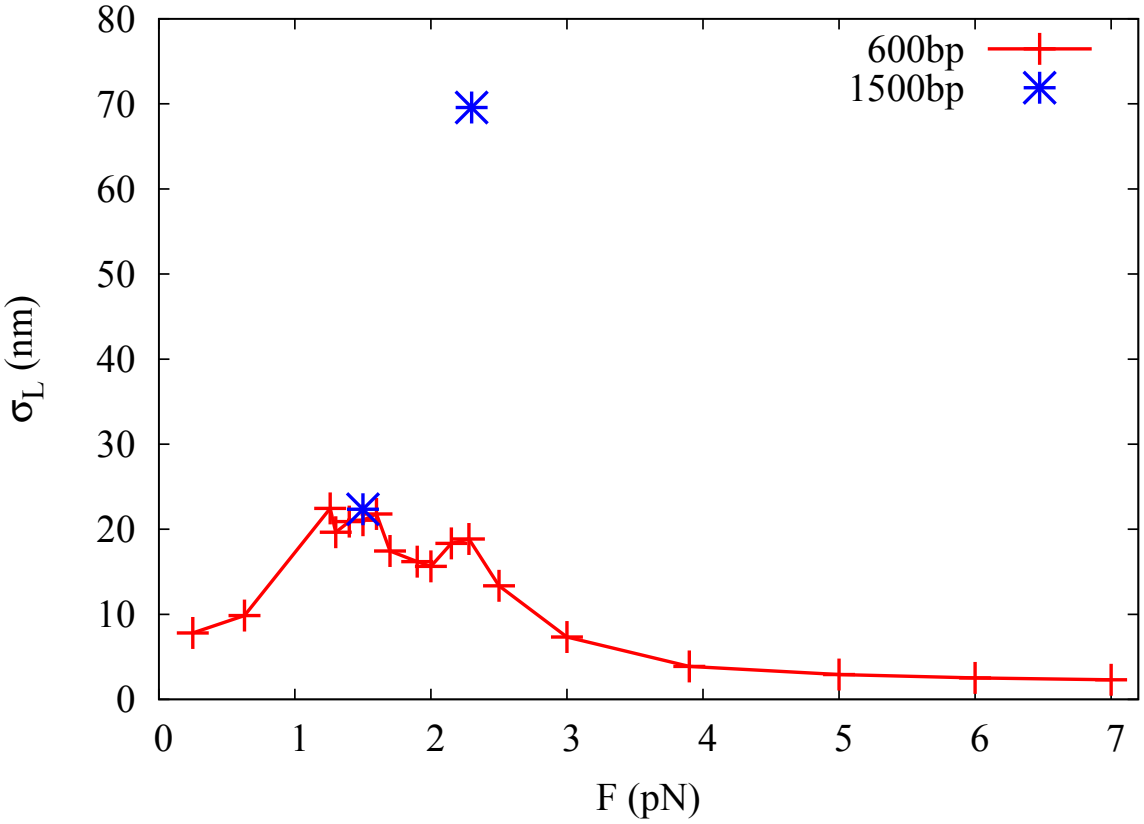


Figure 5.1: Fluctuation spectrum (standard deviations of the end-to-end distance) for $L = 600\,\mathrm{bp}$ (red points, as in inset of Fig. 1b in the main paper), together with two points obtained for a $L = 1500\,\mathrm{bp}$ system (blue points). While the maximum at low force is approximately independent of strand length, the high-force maximum grows with system size in a roughly linear fashion.

This structure of the fluctuation spectrum can be related to crossing the boundaries between the different states shown in the state diagram presented in Sec. 4.1.2: The first peak is due to fluctuations arising from the initial opening of the tip bubble, signifying the transition from the pure plectoneme to the tip-bubble plectoneme state. The second peak is due to the competition between a tip-bubble plectoneme and the extended bubble state. These two effects have a different dependence on system size, which will be discussed in Sec. 5.1.2. The overall width of the peak in

the fluctuation spectrum observed in experiments is somewhat more narrow than it is in simulations for $L = 600\text{bp}$ [2]. This may be due to two reasons:

- Experiments are expected to mainly resolve the high-force peak (cf. Sec. 5.1.2), for which the width at half-maximum is smaller than the full spectrum observed at $L = 600\text{bp}$ (see Fig. 5.1).

- Generic finite size effects may lead to a narrowing of the fluctuation peak at a transition for increasing system size, for example because of cooperative localisation effects that increase with strand length.

Apart from these differences in detail, both simulations and experiments consistently observed considerably wider peaks in the fluctuation spectrum than predicted based on simple models that only include plectonemes and bubbles, without accounting for their respective nucleation costs [3, 2, 63]. This behaviour is caused by the joint localization behaviour of bubbles and plectonemes. Due to co-localisation (see Sec. 4.1.1), pure bubbles and tip-bubble plectonemes each lower the nucleation barrier for the formation of the other, causing a broader transition between the two states, as discussed in Sec. 4.1.3. This leads to a wider fluctuation spectrum than one would observe for two states separated by a large nucleation barrier, which would generate a sharp two-state behaviour. Hence, the broad fluctuation spectra observed by Tempestini and co-workers [2], as well as independently by Meng *et al.*, are indirect evidence supporting the existence of the tip-bubble plectoneme state presented in chapter 4: Models which do not allow for co-localisation in the form of tip-bubble plectonemes can only fit the experimental data if they neglect the nucleation free-energy costs of bubbles and plectonemes. These are however well known to be substantial from both theoretical work [145, 135] and experimental observations [1].

For a similar salt concentration as used in the parametrisation of oxDNA, the experiments of Ref. [2] find that the location of the end-to-end fluctuation peak lies at $F \approx 1\text{pN}$, a force about one piconewton lower that the large-force peak seen in simulations. This observed discrepancy may be due to a number of reasons:

- Still unexplored generic finite-size effects may lower the transition force for longer strands.

- Simulations for both strand lengths reported in this section were performed for the average-base parameterisation of oxDNA. As discussed in Sec. 3.1.3, the crossover between bubble and plectoneme states takes place at slightly lower forces in the sequence-dependent model, as bubbles preferentially form in AT-rich stretches of the double strand.

- For a random sequence, the longer the strand, the higher the likelihood for finding longer AT-rich regions, which may further enhance the probability of bubble formation due to localisation of denaturation.

Taken together, these effects are expected to lower the forces at which the fluctuation peak occurs in the simulated system. Therefore, they should lead to better agreement of simulation results and experimental observations of the force at which a fluctuation maximum occurs [2].

Finally, it may be the case that oxDNA slightly overestimates the cost of forming twist-induced bubbles. It is difficult to find direct comparisons to experiment for this phenomenon that would allow for an independent check. In previous studies, good agreement with experiment was observed for force induced melting [176] without torsional constraint, and for simple duplex melting [172]. This gives confidence that the difference between simulation and experimental results are mainly accounted for by the factors listed above as opposed to generic shortcomings of the model.

5.1.2 Length dependence of fluctuation spectrum

For reasons of computational efficiency, the majority of the data presented in this thesis was taken from systems at a relatively short strand length of $L = 600\,\mathrm{bp}$, which is roughly 5-10 times shorter than common experimental systems. Previous experimental studies of the buckling transition have found a marked dependence of the buckling transition on the strand length L. In particular, the width of the buckling transition and the kinetics of buckling were found to be strongly influenced

by strand length [4]. In order to study the influence of differences in strand length and to be able to directly compare to experimental results, simulations at $L = 1500\,\mathrm{bp}$ were performed. For these systems, an implementation of oxDNA for CUDA-enabled GPUs was used to permit more efficient sampling.

To test the influence of strand length on the overall extension behaviour, force-extension curves at $\sigma = -0.08$ were calculated for $L = 600\,\mathrm{bp}$ and $L = 1500\,\mathrm{bp}$, and are shown in Fig. 5.2. Both curves collapse almost perfectly for high forces, while from $F \approx 1\,\mathrm{pN}$, the shorter system shows a slightly larger end-to-end distance, indicating the presence of finite-size effects due to interactions of the plectoneme with the system boundaries.

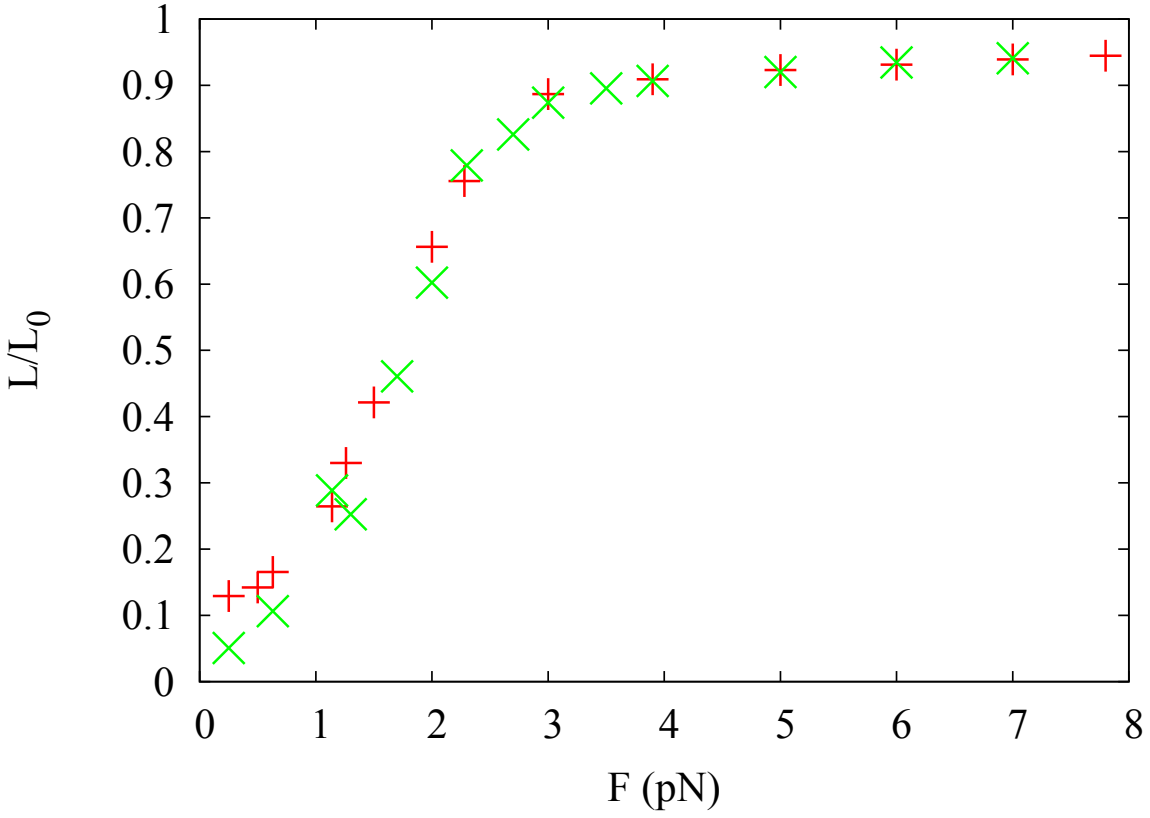


Figure 5.2: Mean end-to-end extensions for $L = 600\,\mathrm{bp}$ (red curve) and $L = 1500\,\mathrm{bp}$ (green curve). For high forces, both lengths lead to identical relative extensions. For $F \lesssim 1\,\mathrm{pN}$, finite-size effects lead to a somewhat increased end-to-end extension for the 600-bp system. Note that in the same force range, qualitatively similar deviations from the experimental torque response behaviour are found, providing another signature of finite-size effects (cf. Fig. 3.8).

The length-dependence of fluctuations in end-to-end extension gives further in-

formation on the structural properties involved at different forces. As discussed in Sec. 5.1.1, the low-force peak in the fluctuation spectrum for $L = 600\text{bp}$ shown in Fig. 5.1 corresponds to the onset of tip-bubble denaturation, while the high-force peak corresponds to the disappearance of the tip-bubble plectoneme in favour of an extended bubble state. From this consideration, the first peak of the fluctuation spectrum is expected to be approximately independent of system size, kink formation being an effect that only involves local strand rearrangement. Conversely, the high-force peak of the enhanced fluctuation region is expected to grow linearly with system size for constant σ, as it is related to the total plectoneme size, whose maximum value is set by the *absolute* linking difference $\Delta Lk = \sigma(N-1)/p$, where N is the number of base pairs in the strand, and $p \approx 10.4$ is the average number of base pairs per turn.

To test these expectations, simulations of the 1500-bp system at $\sigma = -0.08$ and $F = 1.5\text{pN}$ and $F = 2.3\text{pN}$ were performed[1]. These force values are near the expected peaks in the fluctuation spectrum for the shorter system (see Fig. 5.1).

In particular the high-force fluctuations are difficult to sample for the larger system, as it involves dissolution and re-formation of a tip-bubble plectoneme. This corresponds to a global redistribution of supercoiling, a transition which is experimentally known to be a slow process [4]. Four independent simulations were run, with a total simulation time of approximately $128\,\mu\text{s}$ at $F = 1.5\text{pN}$ and ten independent simulations of a total simulation time of approximately $443\,\mu\text{s}$ at $F = 2.3\text{pN}$. The full end-to-end distributions obtained for these state points at $L = 600\text{bp}$ and $L = 1500\text{bp}$ are shown in Figs. 5.3 and 5.4. At $F = 2.3\text{pN}$, clear bimodal behaviour is observed in both systems, where a high-extension population represents the extended bubble state, and a low-extension population the tip-bubble plectoneme state. Note that a similar distribution has been reported from experiments in Ref. [4] for positive supercoiling, where the low-extension state is the unwrithed

[1]For reasons of computational efficiency, the full fluctuation spectrum could only be obtained for the 600-bp system.

structure without bubbles and the high-extension state is expected to be plectonemes without tip-bubbles.

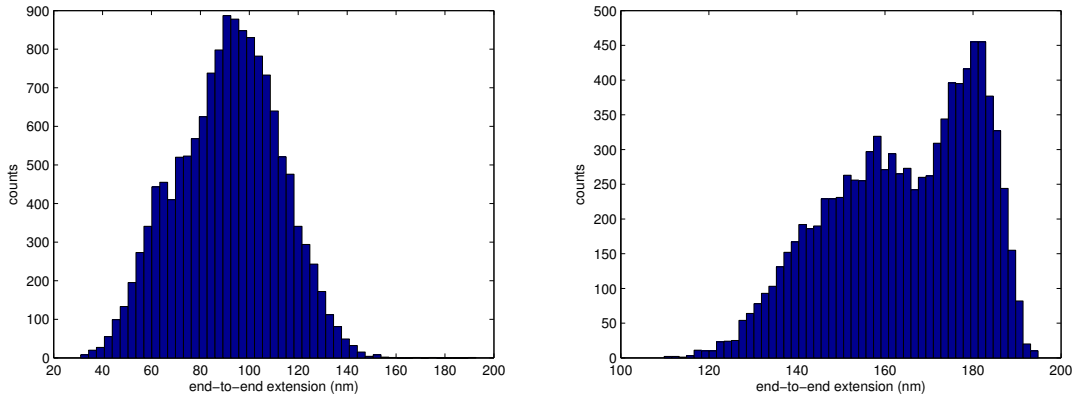


Figure 5.3: Extension distributions for $L = 600\text{bp}$, at superhelical density $\sigma = -0.08$, with $F = 1.5\text{pN}$ (left) and $F = 2.3\text{pN}$ (right). Note the weak bimodality in the right histogram, due to interconversion between bubble and tip-bubble plectoneme states.

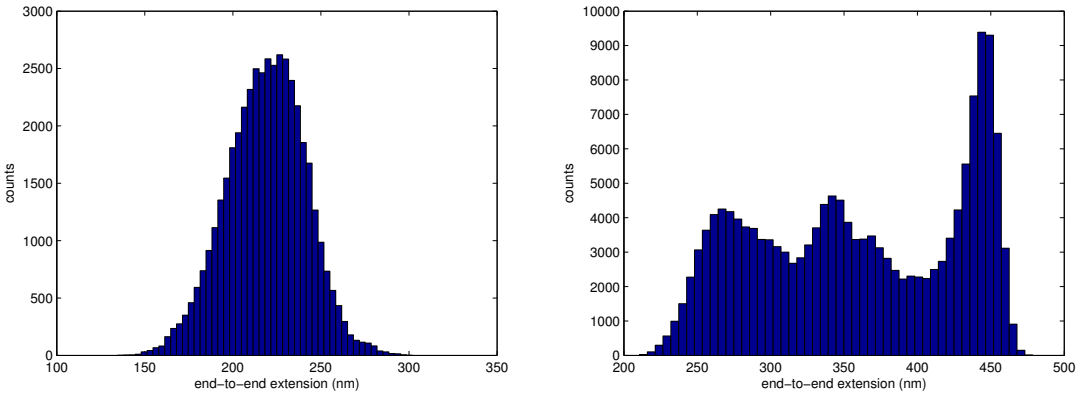


Figure 5.4: Extension distributions for $L = 1500\text{bp}$, at superhelical density $\sigma = -0.08$, with $F = 1.5\text{pN}$ (left) and $F = 2.3\text{pN}$ (right). The results are qualitatively similar to those for the $L = 600\text{bp}$ system shown in Fig. 5.3.

In the case shown in Figs. 5.3 and 5.4, the low-extension population consists of plectonemes with tip bubbles and is much broader because plectonemes with tip bubbles of different sizes have comparable stabilities relative to the extended state. Apart from this competition between high- and low-extension states, further substructure is seen in the low-extension distribution for $L = 1500\text{bp}$ (cf. Fig. 5.4). However, one has to be cautious in the interpretation of this sub-structure, as the

sampling of large supercoiled structures converges slowly, and it was therefore not possible to verify that this feature is robust due to computation time constraints on the GPU processors. At any rate, the influence of this additional sub-structure on the standard deviation calculated for the distribution is expected to be relatively small.

Fluctuation results for both 600bp and 1500bp are plotted together in Fig. 5.1, where the results for $L = 1500\text{bp}$ are represented by blue stars. The results corroborate the expectation that the first peak is largely independent of system size, while the second peak grows much more strongly, with an amplitude increase consistent with the expectation that it would grow linearly with system size.

In Fig. 5.5, typical tip-bubble plectoneme configurations at $L = 1500\text{bp}$ observed for $F = 2.3\text{pN}$ and $\sigma = -0.08$ are shown. Note that in agreement with experimental results [4], these structures are very long-lived, and only show very little diffusion during simulation time. This is because formation and melting of these structures requires a global redistribution of writhe, while diffusion can only take place by reptation of the strand through the entire plectoneme and a concerted motion of the tip bubble. A kymograph for a tip-bubble plectoneme in the $L = 1500\text{bp}$ system at $F = 2.3\text{pN}$ and $\sigma = -0.08$ is shown in Fig. 5.6, demonstrating that these structures are stable on a μs timescale.

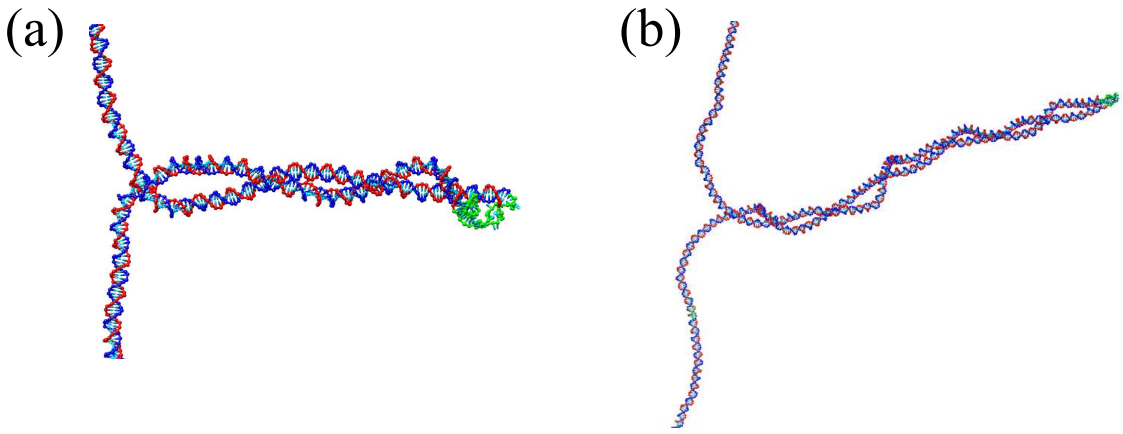


Figure 5.5: (a) Tip-bubble plectoneme structure at $L = 1500$bp, for $\sigma = -0.08$ and $F = 2.3$pN, with plectoneme size 211bp and bubble size 33bp. (b) Tip-bubble plectoneme structure at $L = 1500$bp, for $\sigma = -0.08$ and $F = 2.3$pN, with plectoneme size 576bp and bubble size 10bp. Note that both structures are obtained at the same values of σ and F.

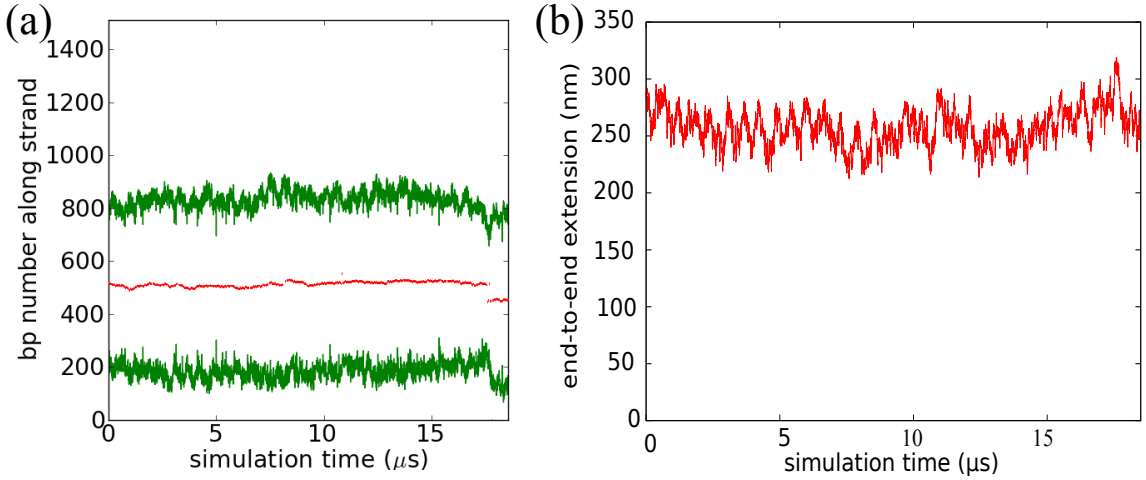


Figure 5.6: (a) Position kymograph of a tip-bubble plectoneme at $L = 1500$bp, for $\sigma = -0.08$ and $F = 2.3$pN, showing stability of the structure on a μs time scale. (b) End-to-end extension kymograph for the run shown in (a).

5.2 Diffusion

Recently, the dynamics of plectonemes has been directly probed in pioneering experiments by VAN LOENHOUT *et al.* [1]. In this section, the diffusive dynamics of individual plectonemes is studied in the simulated system.

Plectonemes can displace along the supercoiled double strand via a continuous, reptational movement, as illustrated in Fig. 5.7. Depending on the applied level of σ and F, plectoneme movement can be interrupted by opening of a tip bubble, as shown in Fig. 5.8. As displacement of a tip-bubble plectoneme requires the coupled motion of the plectoneme and the writhed bubble at its tip, it might be expected to be much slower than displacement without a tip bubble.

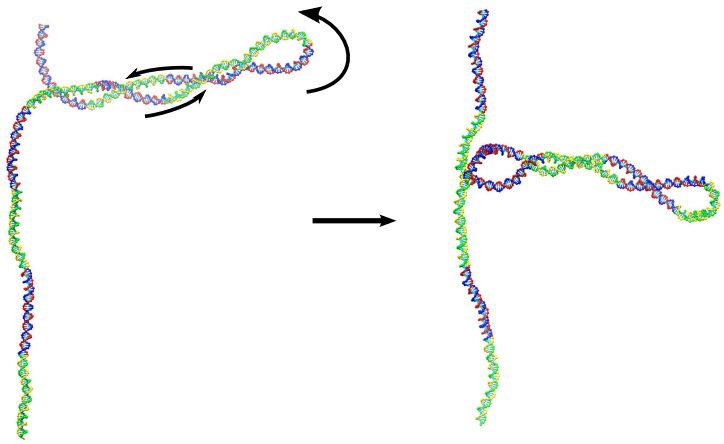


Figure 5.7: Typical displacement of a plectoneme structure in a $L = 600\,\mathrm{bp}$ system by reptation of the double strand. Stretches of 50 bp are shown in different coulour to visualise strand displacement.

Fig. 5.9 shows the mean-square displacement of plectoneme structures with and without tip bubbles in a 600-bp system, using the random sequence given in Appendix C for $\sigma = -0.04$, $\sigma = -0.05$ and $\sigma = -0.06$, and a stretching force $F = 1.27\,\mathrm{pN}$. The data was extracted from sections of trajectories of a total simulated time of $140\,\mu\mathrm{s}$ for each value of σ. On a μs time scale, the mean-square displacements (MSD) $\langle (d(t))^2 \rangle$ of the positions of plectonemes without tip bubbles show an approximately linear behaviour in time, indicating diffusive motion of these structures. In contrast, tip-bubble plectonemes do not show any signature of nonzero

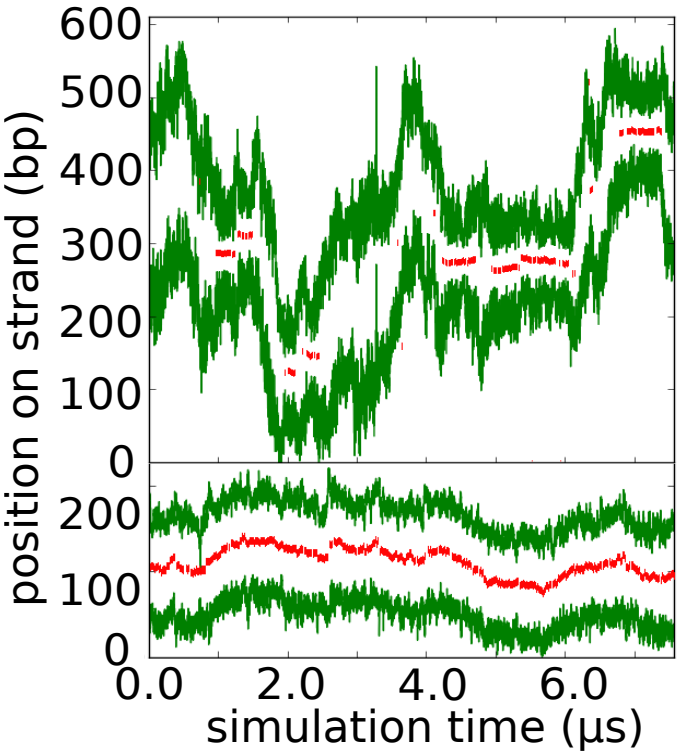


Figure 5.8: Plectoneme kinetics depicted by kymographs of the plectoneme boundaries (green lines). Red denotes the centre of the tip-bubble, which pins the plectoneme and slows down diffusion. The upper panel shows a simulation at $\sigma = -0.05$ and $F = 1.27$pN, where reversible pinning occurs. The lower panel shows a simulation for a fully pinned state at positive supercoiling, $\sigma = +0.08$, $F = 7.9$pN, exhibiting much slower effective diffusion due to the presence of a tip-bubble.

MSD on that time scale.

The diffusion coefficient D and the MSD are related by $\langle (d(t))^2 \rangle = 2Dt$, allowing determination of D from the slope of linear fits to the data shown in Fig. 5.9. The determined diffusion coefficient for unpinned structures decreases as σ becomes more negative because of the larger plectonemes that form at larger $|\sigma|$ (cf. Sec. 5.1.2). Numerical values of D determined from the fitting procedure are given in Table 5.1.

σ	D (kbp^{2}s^{-1})	mean plectoneme size (bp)
-0.04	1.2×10^4	161
-0.05	5.1×10^3	218
-0.06	3.9×10^3	272

Table 5.1: Unpinned diffusion coefficients obtained for different superhelical densities at $F = 1.27$pN from the linear fits shown in Fig. 5.9.

In the experimental setup of Ref. [1], a positive $\sigma > 0$ was chosen such that the

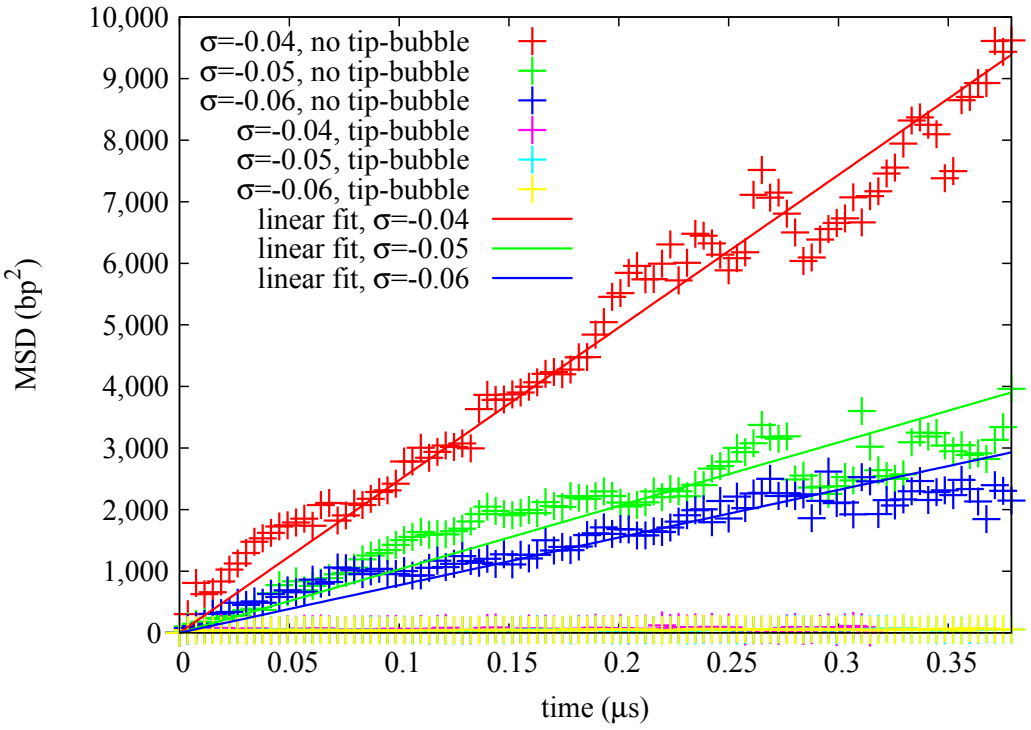


Figure 5.9: Mean-square displacement of plectonemes with and without tip bubble at different superhelical densities and $F = 1.27\text{pN}$. While tip-bubble plectonemes are pinned on a μs timescale, plectonemes without tip bubble exhibit significant diffusion.

plectoneme structure contained approximately 25% of the strand contour length L_0 at a given value of F. An analogous condition was set in simulations, and plectoneme diffusion coefficients were obtained for a range of stretching forces $1\text{pN} \leq F \leq 8\text{pN}$ as outlined above. From simple scaling arguments from polymer knot diffusion due to De Gennes [211], a marked exponential decrease of D is expected for increasing F due to the tightening of the plectoneme structure. Moreover, from the state diagram presented in Sec. 4.1.2, tip-bubble denaturations are expected to occur in diffusing plectonemes also for $\sigma > 0$ at high enough forces $F \gtrsim 5\text{pN}$, leading to a further decrease of diffusion. To separate these two effects, both the effective diffusion coefficient D_{eff} insensitive to the denaturation state of the tip bubble, and the unpinned diffusion coefficient D_{unp} for plectonemes without tip bubbles were calculated. The results are shown in Fig. 5.10, exhibiting a marked decrease of diffusion as a function of applied stretching force F, combined with a further, steeper decrease from $F \approx 5\text{pN}$, where tip-bubble states start to be populated (cf.

Sec. 4.1.2). In order to establish that this additional slowdown is due to tip-bubble formation, the relative slowdown $D_{\mathrm{eff}}/D_{\mathrm{unp}}$ was determined for each force point, and compared to the probability of tip-bubble formation. Fig. 5.11 shows that both quantities indeed behave similarly, suggesting that the effective diffusion coefficient D_{eff} is decreases like

$$D_{\mathrm{eff}} = D_{\mathrm{unp}}(1 - p_{\mathrm{TB}}) + D_{\mathrm{pinned}}\, p_{\mathrm{TB}}, \tag{5.1}$$

where p_{TB} is the probability for a tip bubble to be present, and D_{pinned} is the very small diffusion coefficient of tip-bubble plectonemes. Setting $D_{\mathrm{pinned}} = 0$, Eq. 5.1 directly leads to $D_{\mathrm{eff}}/D_{\mathrm{TB}} = 1 - p_{\mathrm{unp}}$, which is consistent with the data shown in Fig. 5.11.

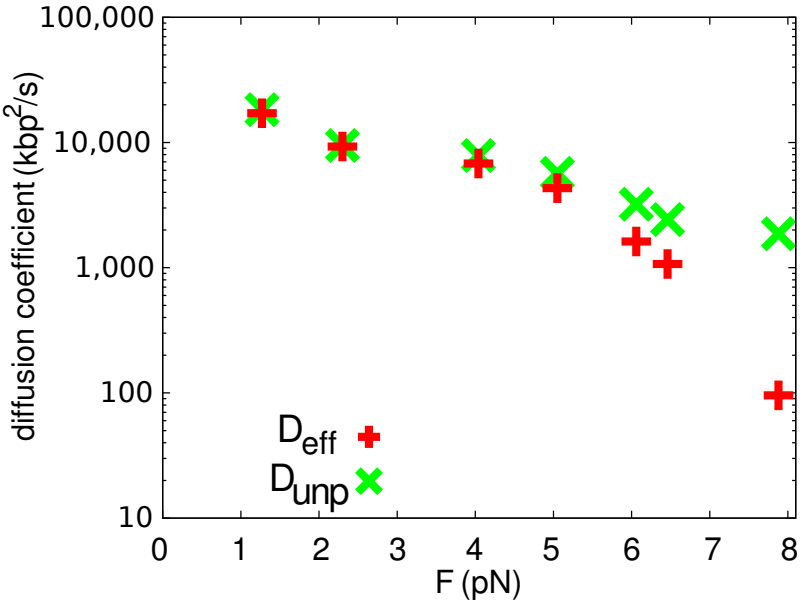


Figure 5.10: Diffusion coefficients D_{unp} and D_{eff} for plectonemes without tip bubbles and for plectonemes independent of tip-bubble presence respectively. From $F \gtrsim 5\mathrm{pN}$, D_{eff} decreases more steeply than for smaller stretching forces. As shown in Fig. 5.11, this is due to tip-bubble denaturations.

The absolute values of the diffusion coefficients obtained in this section are much higher than those measured in Ref. [1]. There are several reasons for this difference.

Firstly, at a strand length of 600bp used in simulations, even the highest superhelical density used here to study diffusive motion, $|\sigma| = 0.06$, corresponds to a linking difference of $\Delta Lk \approx 3.5$. Thus, plectonemes of at most 3-4 double strand

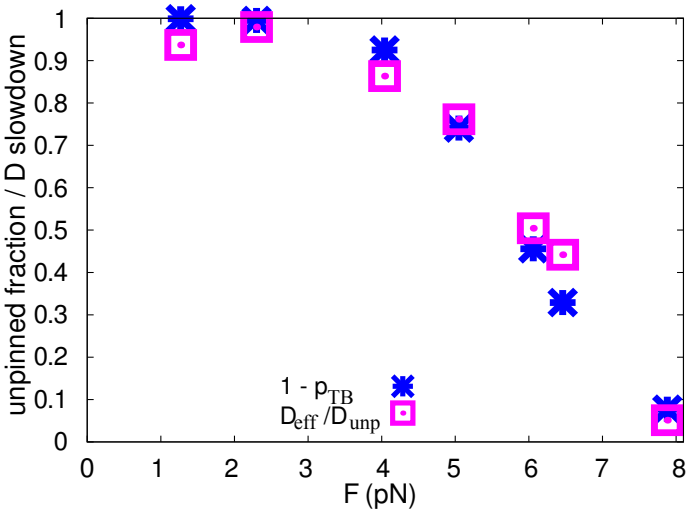


Figure 5.11: Tip-denaturation probability and relative slowdown of the determined effective diffusion coefficient D_{eff} as a function of stretching force, showing that diffusive slowdown is due to tip-bubble formation.

self-crossings are expected to be present. In practice, plectoneme structures are somewhat shorter, due to positive values of $\langle Tw \rangle$. In contrast to this, linking differences $\Delta Lk > 40$ were imposed in Ref. [1], leading to much larger plectoneme structures which are expected to diffuse far more slowly than the simulated structures [211].

Secondly, a high effective monomer diffusion coefficient $D_{\text{sim}} = 6 \times 10^{-7} \text{m}^2\text{s}^{-1}$ was chosen in simulations for this work. This is a common choice in coarse-grained models, and is done in part to increase the efficiency of sampling slow processes. From measurements of diffusive motion of DNA single strands, these effects are estimated to speed up the simulations by up to two orders of magnitude in simulated time [175].

Finally, underlying free-energy landscapes tend to be smoothed out in coarse-grained models, leading to accelerated motion of the simulated diffusion over barriers compared to the experimental system (see e.g. Ref. [212]). The time scale used in the integrator employed by oxDNA (cf. Sec. 2.3 and Appendix A) is set by the mass, length and energy scales used, which determine the frequencies of intra-molecular vibration modes. Coarse-graining may affect differently processes such as the overall bending modes of the double strand, which are important for strand reptation, mak-

ing it difficult to define a homogeneous time scale for all these processes. Therefore, it is generally safer to compare the relative time scale of two processes rather than their absolute duration.

Although a reduction in diffusion is seen in Ref. [1] with increasing applied force F, the forces probed are lower than those at which tip bubbles are seen for positive supercoiling in oxDNA. There are several possible explanations for this slowdown.

Increasing the force is expected to lead to a slowdown of diffusion via reptation, due to tightening of the plectoneme structure [211]. This effect is illustrated in Figs. 5.10 and 5.11 at $F \lesssim 5\text{pN}$. Additionally, sequence-dependent mechanical properties of the strand beyond denaturation may come into play, such as intrinsic curvature of certain double-stranded sequences.

Hence, it may be the case that the data in the specific experimental system of Ref. [1] can be explained without the formation of plectoneme tip bubbles. In that case, plectoneme tip bubbles would be expected to cause an additional subsequent slowdown of diffusion at higher forces. To directly demonstrate the slowdown of plectoneme diffusion by tip bubbles in experiment, it would be necessary to resolve the position of double-strand denaturations and plectonemes at the same time. This could be done for example by using fluorescent probes of single stranded regions.

5.3 Hopping

Next to the continuous, diffusive displacement described in Sec. 5.2, the single-molecule experiments of Ref. [1] suggested the presence of an abrupt, long-range "hopping" displacement mode of plectoneme structures. Also in simulations, long-range displacement of plectonemes by "hopping" was observed, albeit that hopping events occurred fairly rarely. In 21 independent simulations of a 600-bp system at $F = 1.27\text{pN}$ and $\sigma = -0.04$ run for a total simulation time of $148\,\mu\text{s}$, 4 hopping events which involved rapid displacement of the plectoneme center position over a distance of more than 100 bp occurred. In all cases, the plectoneme in the initial

location fully unformed, and then quickly re-nucleated at a distant site (see Fig. 5.12 for a detailed plectoneme position kymograph and microscopic strand structures).

Coexistence of two plectonemes was not observed in equilibrium, as expected at the high-salt conditions used in oxDNA [193, 1]. Due to the absence of equilibrium multi-plectoneme states in the 600-bp system (cf. Sec. 4.1.2), no "writhe exchange" between two plectoneme structures is observed. Hopping only occurs close to the critical buckling superhelical density σ_b, where the system has a non-zero probability to fully dissolve a plectoneme and return to the extended state.

However, for DNA strands of length 1500bp (see Sec. 5.1), long-range writhe exchange between two simultaneously present plectonemes could be observed in an out-of-equilibrium situation. Simulations were started from a linear, homogeneously undertwisted double strand at $\sigma = -0.08$ and $F = 2.3$pN. A typical example of such a run is shown in Fig. 5.13, where the kymograph initially shows the simultaneous presence of two plectonemes. Coexistence is followed by the disappearance of the smaller plectoneme, mediated by long-range transport of its writhe to the larger plectoneme, which then stays stable for a long time.

Van Loenhout *et al.* [1] observed a similar long-range displacement behaviour over distances of up to 15kbp at $F = 0.8$pN and $\sigma \approx +0.04$. Some hopping events observed in the experiments of Ref. [1] showed immediate dissolution of the initial plectoneme, and re-nucleation at a distant site, which is reminiscent of the equilibrium mechanism described above. However, the experiments are performed at ambient conditions that permit stable coexistence of multiple plectonemes [193, 1]. Hence the writhe-exchange mechanism reminiscent of the one observed in non-equilibrium situations in simulations is likely to be present as well.

5.4 Summary

In this chapter, the dynamical characteristics of plectoneme structures were observed. In Sec. 5.1, it was shown that the fluctuation behaviour of the end-to-end extension of a supercoiled DNA strand contains non-trivial information on the un-

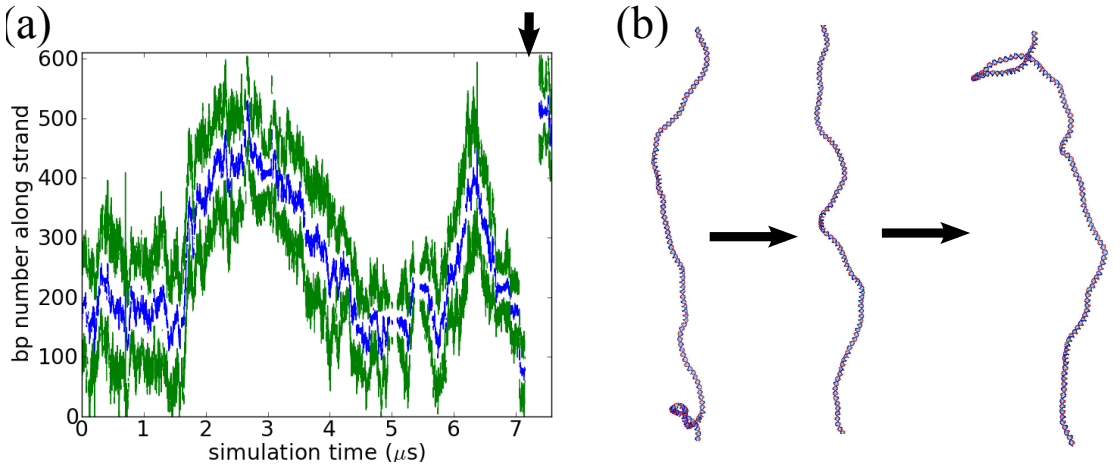


Figure 5.12: Hopping event observed in a simulation at $\sigma = -0.04$ and $F = 1.27$pN: (a) Kymograph of plectoneme boundaries (green lines) and centre position (blue line) showing a hopping event at the time marked by a black arrow. (b) Structure of the DNA strand close to the hopping event at $t = 7.09\mu$s (left), $t = 7.34\mu$s (middle) and $t = 7.39\mu$s (right).

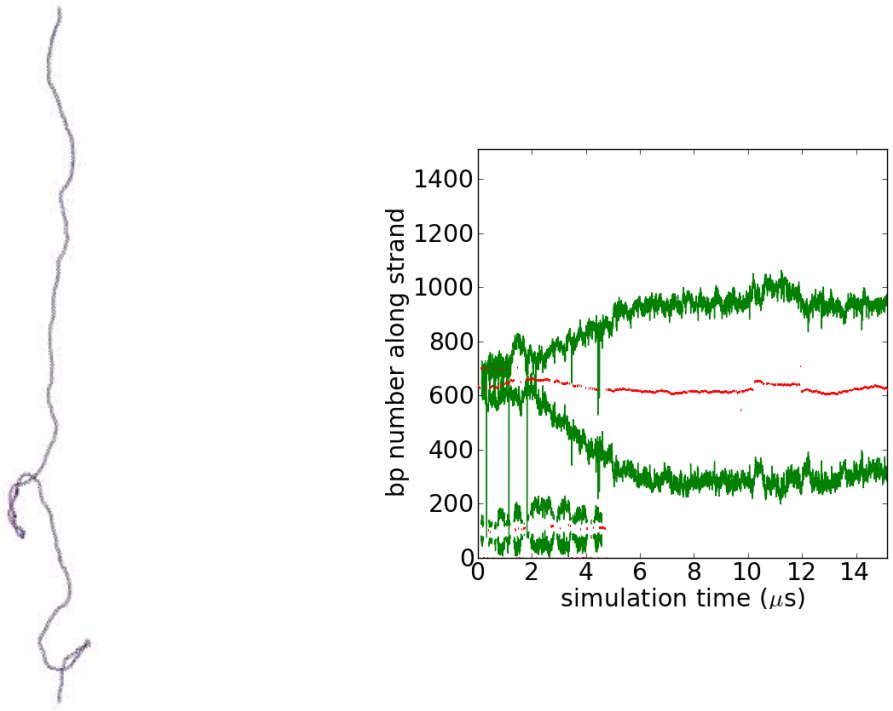


Figure 5.13: Coexisting plectonemes in an out-of-equilibrium run started from a homogeneously underwound linear double strand of length 1500bp at $\sigma = -0.08$ and $F = 2.0$pN, as shown in the Supplementary Movie. Left: Molecule configuration showing two coexisting plectonemes. Right: Kymograph of two plectoneme sites. Ultimately, the smaller plectoneme dissolves and a single, large plectoneme is formed.

derlying competition between microscopic structures. In particular, available experimental results on end-to-end fluctuations can only be explained by simple models that treat bubbles and plectonemes as spatially separate populations if the nucleation costs of these structures are ignored. The presence of the tip-bubble plectoneme state, in contrast, can explain the observed behaviour without this unphysical assumption. The diffusive displacement behaviour of individual plectonemes was investigated in Sec. 5.2, and it was shown that tip bubbles lead to an arrest of plectoneme motion which causes a stronger slowdown at larger stretching forces than expected from models that ignore the ability of the double strand to denature. Finally, Sec. 5.3 described discontinuous "hopping" behaviour of plectonemes, which was observed in simulations, in agreement with the results of recent single molecule experiments.

Chapter 6

Bending and denaturation in DNA minicircles

DNA minicircles are a popular model system for the study of molecular properties of DNA for several reasons.

Firstly, most bacteria possess circular genomes, and in higher cells, DNA is wound around histones in small solenoidal loops [26]. Furthermore, circular plasmids are a common in bacterial conjugation, and are used in biotechnology as cloning vectors.

Secondly, following initial work by Cloutier and Widom [213, 214] suggesting unexpected cyclisation behaviour, short minicircles have genererated considerable interest as an experimental system for the study of short-range DNA elasticity [215, 216, 217, 218, 219, 220, 221, 82, 168, 222, 109, 110]. From the perspective of modelling, minicircles of a size of about 100bp are interesting because they are accessible by atomistic as well as coarse-grained and continuum models, and thus provide an opportunity to study the same system using different theoretical descriptions. Comparing the results from these different approaches and testing how well they reproduce experimental data can shed light on the strengths and limits of the different models, and clarify some of the relevant physics of DNA.

In this chapter, minicircles of a length of 102-108bp are studied using oxDNA. It is found that the topology of the system influences its denaturation behaviour, and introduces long-range correlations of the denaturation profile of the minicircle, a mechanism that is supported by recent experimental results.

Finally, the predictions made in the framework of oxDNA are related to the predictions made by all-atom simulations and a discrete model of local denaturation.

6.1 System properties

Closed minicircles of length 102, 106 and 108bp were studied at $T = 300$K each with a set linking difference of $Lk = 9$. Assuming the equilibrium pitch of $p_0 = 10.36$bp/turn of oxDNA at this temperature as determined in Sec. 3.2, these lengths lead to a superhelical density σ of $\sigma_{102} = -0.09$, $\sigma_{106} = -0.12$ and $\sigma_{108} = -0.14$ respectively. Simulations were started from perfectly circular states, which however relaxed to an equilibrated structure very quickly (see Table 6.2 for decorrelation times of writhe Wr and the local twist angle θ). The total simulated time of production runs was at least 70μs.

Three different base sequences are investigated. The DESIGN sequence is an artificial sequence that possesses long AT and GC-rich stretches and is used to illustrate generic denaturation behaviour. The RANDOM sequence was used before in experiments by Du *et al.* [217] to study bending-induced deformations of DNA. Finally, a biological sequence from the far upstream element (FUSE) of the human *MYC* gene [223] was integrated into the RANDOM sequence. The explicit sequences used in simulations are given in Appendix C.

To check the generic behaviour of the system and as an independent test for equilibration, every system length was also run in the average-model parametrisation of oxDNA. Furthermore, for all strand lengths a relaxed control minicircle with $Lk = 10$ was simulated, which corresponds to small superhelical densities of $\sigma_{102} = 0.02$, $\sigma_{106} = -0.02$ and $\sigma_{108} = -0.04$ for strand lengths 102,106 and 108bp respectively. In these near-relaxed minicircles, no significant stable denaturation was observed.

6.2 Denaturation properties

Denaturation profiles were determined for all circle lengths and DNA sequences. For each system, 5 independent production runs with a simulated time of at least

70μs each were performed, and the results combined to determine the equilibrium denaturation pattern. In this section, only the results for $L = 102$bp are presented. Corresponding results for the other lengths are given in Appendix B.

6.2.1 Opening distribution and dynamics

Kymographs of double-strand denaturation over a time scale of approximately 80μs are shown for all three sequences in Fig. 6.1. While the dynamics of individual base pair opening and closing happens on a very short time scale, larger changes to the denaturation pattern happen more rarely. For example, this behaviour can be seen in Fig. 6.1(a), where the main focus of denaturation shifts between the first two AT-rich stretches in the strand. This behaviour overlaying the faster dynamics of individual base-pair breaking and formation is associated with global rearrangements of the minicircle configuration.

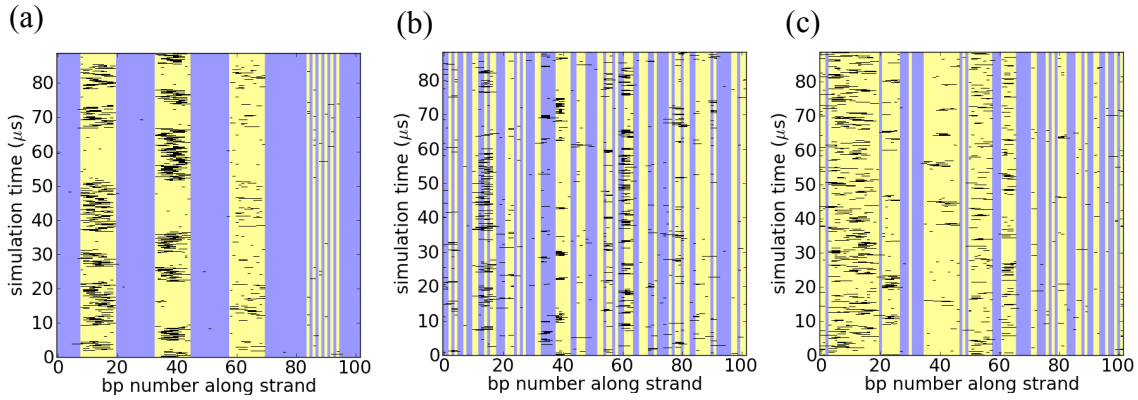


Figure 6.1: Denaturation kymographs for minicircles for length $L = 102$bp. AT base pair positions are coloured in yellow, and GC base pair positions in blue. Note the slow change in denaturation patterns, particular in (a), associated with global rearrangement of the main denaturation site in the minicircle. (a) DESIGN sequence (b) FUSE sequence (c) RANDOM sequence.

Denaturation probabilities for the three different sequences studied here are shown in Figs. 6.2 - 6.4. Unsurprisingly, the overall base pair identity and length of weak sequence stretches are the main determinant of the denaturation distribution. However, effects that go beyond the local base identity are seen to play a role in determining the denaturation pattern.

Firstly, not only the local base identities, but also the sequence of bases along a strand matter for the denaturation probability. This effect is most clearly seen in the denaturation pattern of the DESIGN sequence (Fig. 6.2), where the denaturation probability of the third AT-rich stretch (base pairs 59-70) is considerably reduced compared to the two other AT-rich stretches of equal size in the sequence. The difference between both parts of the sequence is that the first weak stretch (base pairs 9 to 20) consists of a $(TA)_6$ sequence, while the the second stretch has the sequence $(AT)_6$ and the third stretch consists of the sequence $(AA)_6$. From the sequence-dependence of physical interactions in DNA as described by the SantaLucia model [131], on which the parametrisation of oxDNA is based [173], these stretches are expected to show different denaturation behaviour, although they have precisely the same content of A and T bases: The stacking interaction between two consecutive nucleotides depends on identities and order of both bases. For the AA/TT step, the stacking energy is higher than for the AT or TA steps, which explains the differences

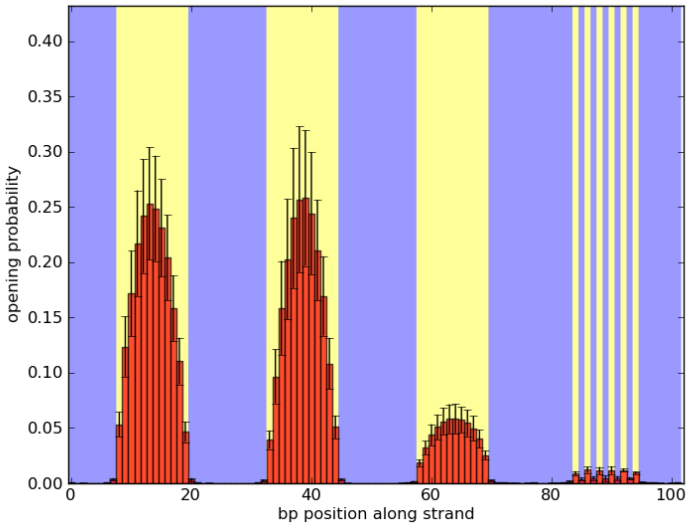


Figure 6.2: Denaturation profile for a minicircle of length $L = 102$bp posessing the DESIGN sequence. AT base pair positions are coloured in yellow, and GC base pair positions in blue. Error bars indicate standard deviation between 5 independent simulations. Note that the third AT-rich stretch (possessing a $(AA)_6$ sequence) has a lower overall denaturation probability than the first two AT-rich stretches (possessing $(TA)_6$ and $(AT)_6$ sequences respectively; see Appendix C).

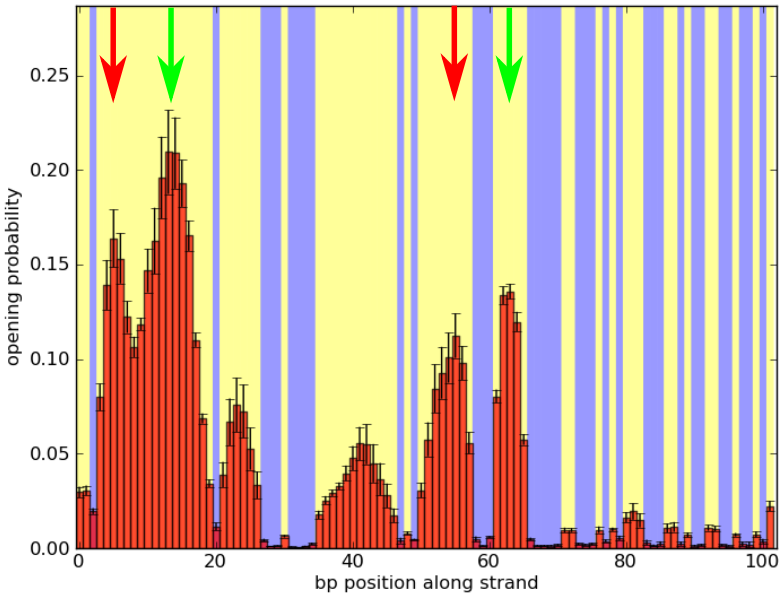


Figure 6.3: Denaturation profile for a minicircles of length $L = 102$bp posessing the FUSE sequence. AT base pair positions are coloured in yellow, and GC base pair positions in blue. Error bars indicate standard deviation between 5 independent simulations. Arrows indicate peaks separated by approximately half a circle length that are augmented by cooperative denaturation (see Fig. 6.6). Sites of correlated opening are marked by arrows of the same colour.

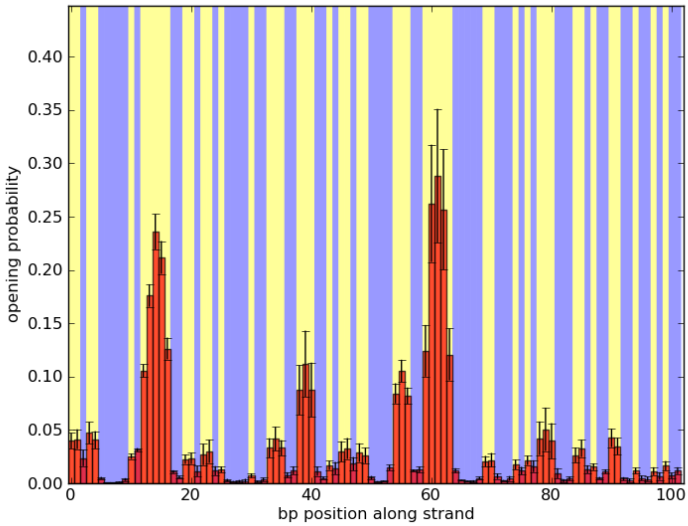


Figure 6.4: Denaturation profile for a minicircles of length $L = 102$bp posessing the RANDOM sequence. AT base pair positions are coloured in yellow, and GC base pair positions in blue. Error bars indicate standard deviation between 5 independent simulations.

in denaturation between the different AT-rich stretches observed in Fig. 6.2.

Secondly, substructure is observed within the melting profiles of AT-rich regions, such as the two-peak structure in denaturation probability within the AT-rich stretch between bp positions 3 and 20 of the FUSE sequence (Fig. 6.3). These positions are associated with enhanced denaturations at sites separated by a distance of approximately half the minicircle length.

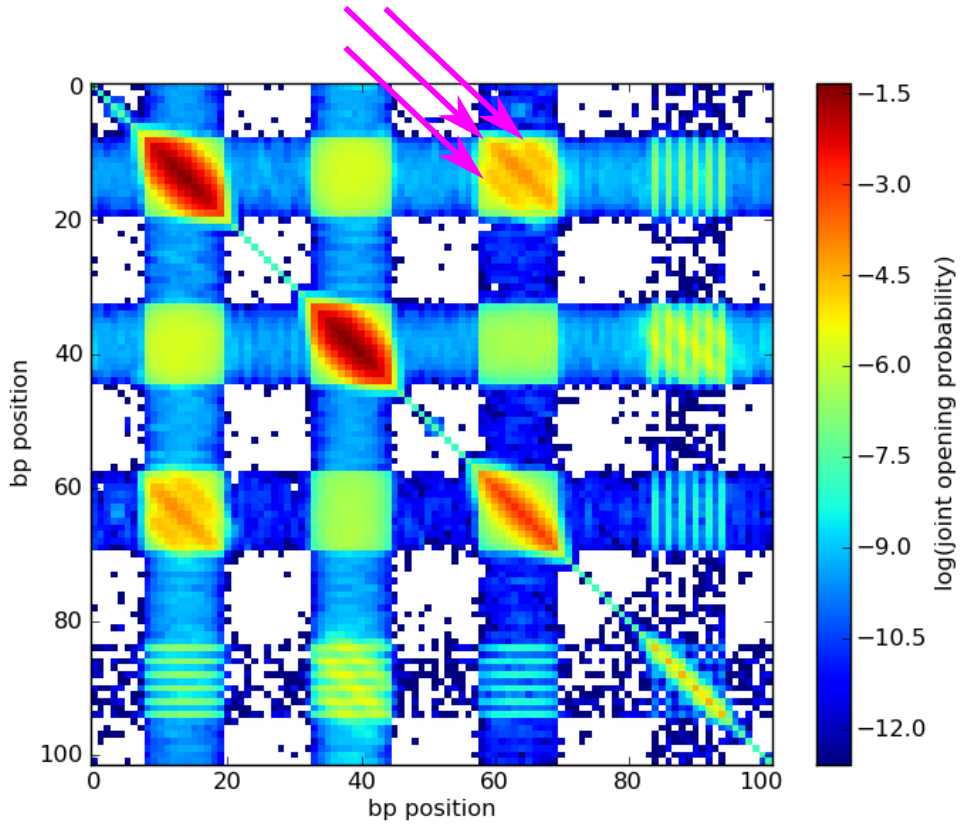


Figure 6.5: Joint opening probabilities for the DESIGN sequence of length 102bp. States never observed in simulations are coloured white. High correlations along the diagonal are unsurprising, as the close-by presence of a denatured bp increases the probability of a given basepair to be denatured as well. High values in off-diagonal elements indicate long-range cooperative bp opening. The most significant long-range correlations are observed for AT rich stretches with a distance of approximately half the minicircle size. Purple arrows indicate substructure in the denaturation pattern, which arises because denaturation of parallel base pairs is preferred (see Fig. 6.8). This leads to two additional maxima at a distance of 5bp from the off-diagonal, corresponding to half a helical turn.

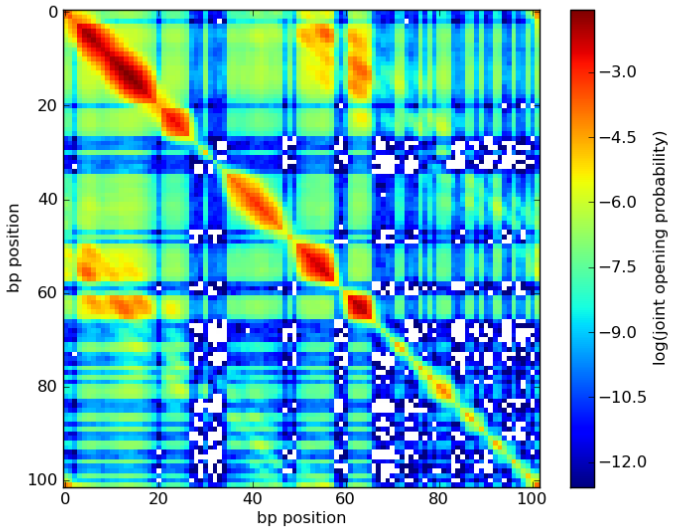


Figure 6.6: Joint opening probabilities for the FUSE sequence of length 102bp. States never observed in simulations are coloured white. The most significant denaturation correlation is observed for AT-rich stretches separated by about half the minicircle length. The geometrical substructure is similar to the one indicated in Fig. 6.5.

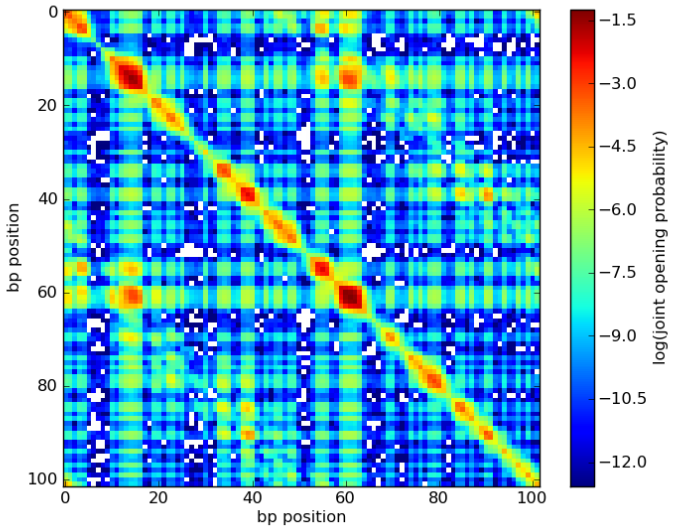


Figure 6.7: Joint opening probabilities for the RANDOM sequence of length 102bp. States never observed in simulations are coloured white. Also for this sequence, the dominant long-range correlation is for AT-rich stretches separated by about half the minicircle size. Absence of other long AT-rich stretches leads to a concentration of joint denaturation on the longest AT-rich regions around basepairs 10 and 60.

To further quantify these long-range effects in the denaturation profiles, the joint opening probability of every possible base pair in the system was calculated (Figs. 6.5- 6.7). Strong correlations are typically observed for AT-rich regions with a distance corresponding to approximately half the minicircle size. The structural basis of these long-range correlations along the system is cooperative kinking on two opposite ends of the minicircle, as shown Fig. 6.8: In a typical "droplet" shape configuration with only one denaturation bubble, the highest local curvature of the double strand is observed at an angle approximately 180° away from the bubble. The overall bending energy in the system can then be reduced by opening of a second denaturation bubble in this region. These structures are particularly stable when they exhibit a high AT content, corresponding to a smaller enthalpic cost of the denatured regions.

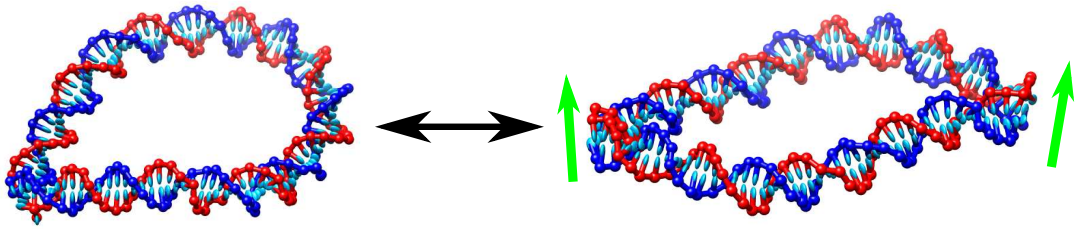


Figure 6.8: Schematics of long-range cooperative bubble opening in a short minicircle. If one bubble is opened, the minicircle is in a "droplet" shape. Bending is then strongest on the opposite end of the minicircle, where long-range cooperative denaturations are most likely to form. Green arrows indicate the approximate base pair orientation at the denatured sites. As explained in the main text and indicated in Fig. 6.5, denaturation is favoured if the two vectors are parallel or antiparallel, leading to the substructure in long-range correlations marked by purple arrows in Fig. 6.5.

While denaturation on opposite regions of the minicircle is seen to be dominant from Figs. 6.5-6.7, additional, more subtle effects may also influence the denaturation behaviour:

- Commensurability effects stemming from the geometry of the helix may play a role in the determination of the site where the bubbles form. For example, cooperative breaking of base pairs is favoured at sites which have approximately

parallel base pair vectors (see Fig. 6.8), and allow the denaturation bubbles to form without writhing the structure. The substructure observed close to the off-diagonal denaturation correlations of the AT-rich regions is an indication for this behaviour. This effect can be seen most clearly for the DESIGN sequence (Fig. 6.5), but is also present in the FUSE and RANDOM sequences. Representing a spatial correlation along the double strand, these effects are expected to vanish for strands considerably longer than one persistence length.

- Correlations between a higher number of denaturations along the minicircle may occur. For example, three small denaturation bubbles may form along the strand in approximately equal distances, which act as hinges between almost unbent double-stranded regions in between. However, this kind of denaturation pattern is expected to be suppressed for small circles because of the considerable nucleation costs of denaturation bubbles [135].

6.2.2 Topological properties

As closed-circular systems, minicircles offer the possibility to directly test the influence of imposed topology on molecular conformations, without the mathematical subtleties of defining topological quantities for open strands [164]. Of particular interest is the partitioning of the linking number into twist and writhe according to the Fuller-White-Călugăreanu relation (see Sec. 2.1.2), which can indicate if superhelical stress is primarily released trough local untwisting or though global buckling of the system [224].

To quantify the topology of the minicircles studied in this chapter, discretized versions of the formulae for twist (Tw) and writhe (Wr) were used, as introduced in Sec. 2.1.2.2 and detailed in Appendix B. The resulting values for Tw and Wr are listed for the different systems in Table 6.1. In all cases, the sum of both quantities differs from the exact value $Lk = 9$ given by the Fuller-White-Călugăreanu relation by less than 1%, indicating that the approximation schemes used are consistent.

sequence (length)	$\langle Tw \rangle$	$\sigma(Tw)$	$\langle Wr \rangle$	$\sigma(Wr)$	$\langle Tw + Wr \rangle$	$\sigma(Tw + Wr)$
DESIGN(102)	9.1300	0.1404	-0.1287	0.1493	9.0014	0.1785
DESIGN(106)	9.2744	0.2205	-0.2346	0.2138	9.0398	0.2353
DESIGN(108)	9.3885	0.2672	-0.3395	0.2421	9.0490	0.2697
FUSE(102)	9.1352	0.1362	-0.1517	0.1519	8.9835	0.1773
FUSE(106)	9.2886	0.2136	-0.2929	0.2266	8.9957	0.2417
FUSE(108)	9.4048	0.2606	-0.3736	0.2504	9.0313	0.2685
RANDOM(102)	9.1489	0.1247	-0.1823	0.1424	8.9666	0.1593
RANDOM(106)	9.3478	0.2074	-0.4089	0.1979	8.9389	0.2411
RANDOM(108)	9.4855	0.2561	-0.4779	0.2161	9.0076	0.2814

Table 6.1: Thermal averages $\langle \cdot \rangle$ and standard deviations σ of Tw, Wr and $Tw+Wr$. In all cases, the value of the thermal average $\langle Tw + Wr \rangle$ deviates from the exact value $Lk = 9$ given by the FULLER-WHITE-CĂLUGĂREANU relation by less than 1%.

For all strand lengths, the thermal average of Wr is found to be negative and relatively small, with $|\langle Wr \rangle| < 0.5$. This gives an indication that at the short strand length of the minicircles studied in this chapter, the undertwist imposed on the system is mainly released through twisting, in particular through localisation of the undertwist into denatured regions rather then in writhing, which would incur a high bending energy. Indeed, denaturations in the minicircle are observed for well over 99% of the simulation time for all systems studied in this chapter.

In the determination of writhe, difficulties can arise due to the high curvature of the strand axis in denatured regions, which can lead to a blowup of individual terms when calculating Wr according to Eq. 2.11. To handle these rare cases, a cutoff was imposed on the magnitude of individual terms in Eq. 2.11. More information on this subtlety is given in Appendix B.

To quantify the dynamics of twist and writhe interchange, the autocorrelation in the time series of Wr and the local twist angles θ_i was determined for all systems studied here. While the autocorrelation of Wr gives a characteristic time scale of global strand rearrangement, the autocorrelation of θ_i reflects the time scale on which local twist transfer takes place[1]. In general, the opening probability of a particular base pair will influence the dynamics by which twist is absorbed at a particular site.

[1]Note that Wr and θ are not related by a FULLER-WHITE-CĂLUGĂREANU-type relation, as θ_i is a local quantity. Twist can be locally redistributed without changing the overall value of Tw.

Hence, sequence effects will influence the autocorrelation time of θ_i at a given time. In order to obtain a characteristic of the entire system, the autocorrelations θ_i were averaged over all base pairs, yielding the average autocorrelation

$$\mathrm{ACF}_\theta(\tau) = \overline{\langle \theta_i(t)\theta_i(t+\tau)\rangle}, \tag{6.1}$$

where the angular brackets indicate thermal averaging and the overline site averaging. Similarly, the writhe autocorrelation is given by

$$\mathrm{ACF}_{Wr}(\tau) = \langle Wr(t)Wr(t+\tau)\rangle. \tag{6.2}$$

The autocorrelation functions for the system of size 102bp carrying the RANDOM sequence is shown in Fig. 6.9. Writhe is seen to decorrelate faster than the average local twist angle. This behaviour might be expected, as the dynamics of local twist is set by the motion of denaturation sites (see Fig. 6.1(c)), which takes place on a longer time scale than the change of writhe, at least for the minicircle systems studied here, which have a relatively small writhe component $|\langle Wr\rangle| < 0.5$.

The decorrelation times for the systems studied in this chapter were determined by fitting single exponential functions to the autocorrelations as $\mathrm{ACF}_{Wr}(\tau) \sim exp\left(-\frac{\tau}{T_{Wr}}\right)$ and $\mathrm{ACF}_\theta(\tau) \sim exp\left(-\frac{\tau}{T_\theta}\right)$. The results of this fitting procedure are given in Table 6.2. While the autocorrelation functions in the RANDOM and FUSE sequence systems all showed approximately exponential decay with τ, fitting was more difficult for the DESIGN sequences, where the twist autocorrelation function $\mathrm{ACF}_\theta(\tau)$ showed heavy tails after a faster initial decrease, yielding long decorrelation times T_θ. This reflects the relatively slow local twist dynamics visible in the kymograph of the DESIGN sequence shown in Fig. 6.1(a). For comparability of the different cases, the same fitting procedure was also applied for these system.

When interpreting the absolute duration of a simulated process, the caveats concerning the interpretation of simulated time in coarse-grained simulations discussed in Sec. 5.2 have to be kept in mind. However, comparing relative time scales might be expected to be less sensitive to these issues. In all cases, $\tau_\theta/\tau_{\mathrm{Wr}} > 1$, showing that writhe fluctuations decorrelate about two to ten times faster than local twist values.

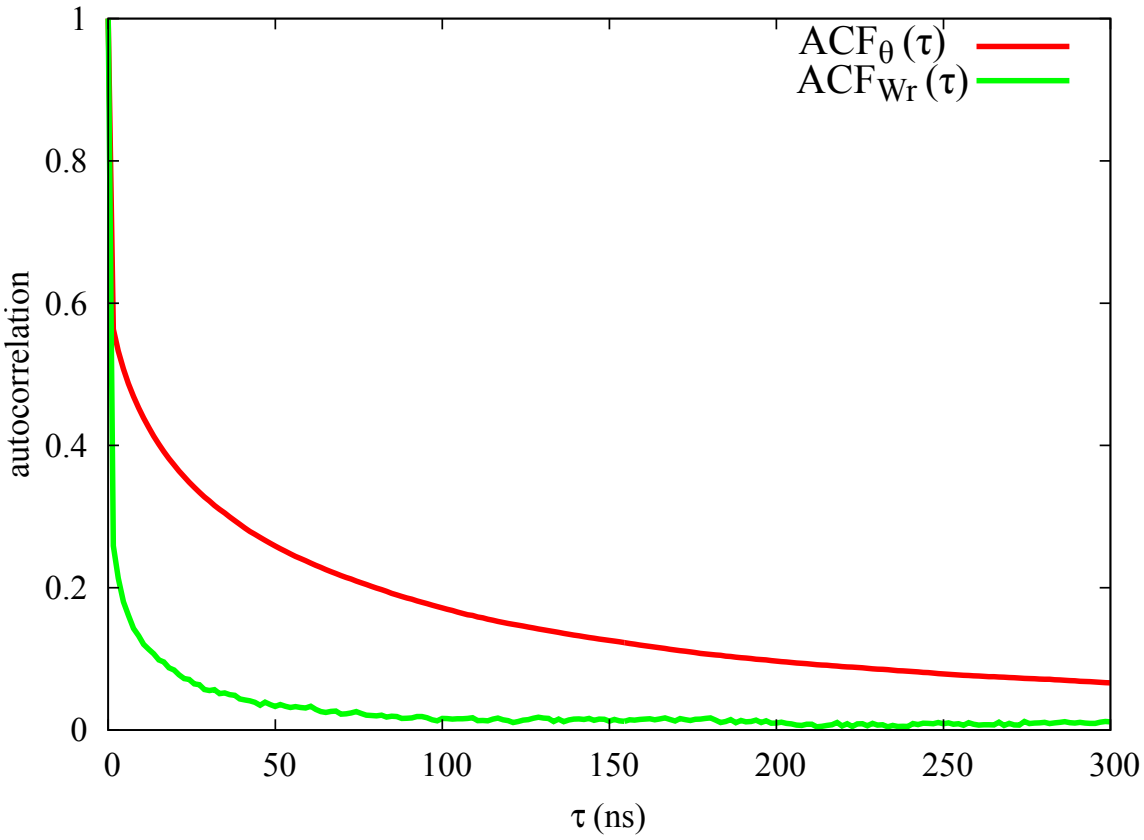


Figure 6.9: Autocorrelation of writhe (green curve) and local twist angle θ (red curve) in the 102bp system carrying the RANDOM sequence. Local twist decorrelates more slowly than writhe, as might be expected due to the relatively slow motion of denaturation bubbles in the system (see Fig. 6.1(c)).

sequence (length)	T_θ(ns)	T_{Wr}(ns)	T_θ/T_{Wr}
DESIGN(102)	320	29	11.0
DESIGN(106)	590	81	7.3
DESIGN(108)	668	65	10.3
FUSE(102)	109	19	5.7
FUSE(106)	195	74	2.6
FUSE(108)	230	82	2.8
RANDOM(102)	151	27	5.6
RANDOM(106)	274	80	3.4
RANDOM(108)	223	51	4.4

Table 6.2: Decorrelation times of the θ and Wr time series. Writhe is seen to decorrelate about one order of magnitude faster than the local twist angle θ.

6.3 Comparison to experiment and other modelling approaches

The cooperative denaturation mechanism proposed in this chapter is supported by experimental findings by Lionberger*et al.* [219]. In this work, the authors observed

cooperative kinking at opposite ends of minicircles of lengths of about 100bp using cryo-EM imaging methods. They also performed BROWNIAN dynamics simulations of a discretisation of the worm-like chain model which represents DNA as a sequence of sections of length 1nm [225]. While the cooperative kinking behaviour observed in experiments could be observed, no conclusions could be drawn as to the microscopic structure underlying these kinks, as the model does not include a representation of double strand denaturation. The results presented in this chapter are compatible with these findings. Additionally, results from coarse-grained simulations suggest a role of sequence in determining the location of the kink defects.

In addition to the simulations using oxDNA, the minicircle systems presented in this chapter were also studied using atomistic simulations using the AMBER ff99parmbsc0 forcefield [226] by a collaborating group, as well as the local double strand denaturation model developed by BENHAM and co-workers [135]. The latter model treats double-strand denaturation upon supercoiling as local two-state transition, which has a free-energy cost that depends on the identity of a given base pair. While the different intrinsic twist and torsional stiffnesses of double-stranded and denatured regions are taken into account, the model does not give any information on the spatial arrangement of the molecule. Results for system denaturation from all three approaches are shown in Fig. 6.10 for the 100bp system.

Both coarse-grained simulations and the BENHAM model can be used to predict the denaturation profiles of the minicircle, albeit at very different computational expense. Predictions using the BENHAM model are very fast for short strands, and can access the overall denaturation behaviour of DNA of kilobase length and above. In comparison to this, calculations of minicircle melting profiles using oxDNA took several weeks on a current CPU to equilibrate, making it difficult to obtain such predictions for longer systems. Both models yield the unsurprising overall prediction that AT-rich regions are more likely to denature than GC-rich regions, and give fair agreement on the overall position of the denaturation peaks. As shown in Sec. 6.2, oxDNA predicts additional effects in the local denaturation probability due to sequence ordering, overall strand topology, and the geometry of the double

helix. This is reflected in the differences between both model predictions in Fig. 6.10. OxDNA was parametrised to reproduce sequence dependent thermodynamics, but more subtle mechanical sequence effects such as intrinsic bending of A-tracts [187] are not explicitly taken into account in this coarse-graining approach. However, one might expect that these effects are weaker than the generic thermodynamics effects of melting for the large majority of sequences, as suggested by the empirical success of the SantaLucia model underlying oxDNA's sequence-dependent parametrisation. Such behaviour may however by studied in atomistic simulations.

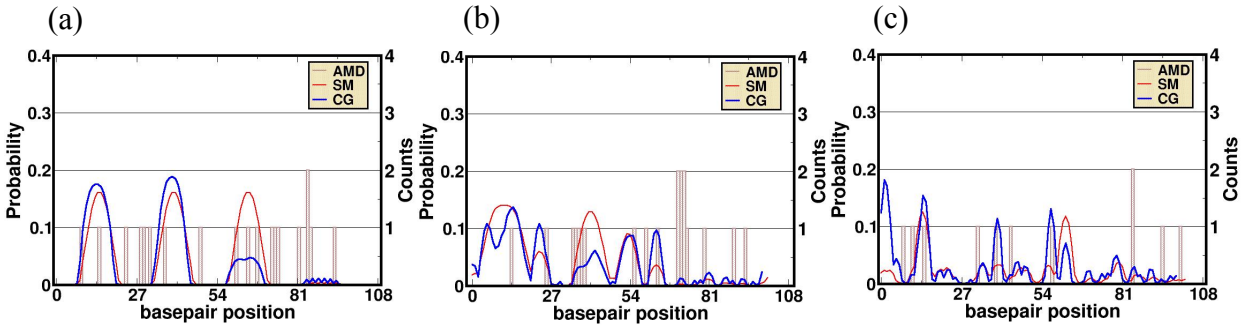


Figure 6.10: Comparison of denaturation predictions of the coarse-grained model oxDNA (CG) to predictions by the statistical mechanical local denturation model by Benham and co-workers (SM) and denaturation events in atomistic simulations (AMD) for 100-bp minicircles. Due to a difference in the equilibrium helical pitch in the AMBER and oxDNA force fields, this length corresponds to $\sigma = -0.05$ in AMBER and $\sigma = -0.06$ in oxDNA. (a) DESIGN sequence, (b) FUSE sequence, (C) RANDOM sequence. Plots provided by T. Sutthibutpong.

All-atom simulations are currently limited from accessing a timescale beyond the microsecond regime due to their computational expense. In fact, the denaturation events observed in atomistic simulations on the time scale of 100 ns shown in Fig. 6.10 do not show a strong correlation with the denaturation preditions of oxDNA and the Benham model, suggesting that they do not reflect the equilibrium denaturation distribution of the strand. While not allowing one to study the equilibrium properties of denaturations, molecule structures obtained in atomistic simulations do support the picture of correlated denaturation on opposite positions ends of the DNA minicircle. Furthermore, these simulations can access sequence dependent mechanical properties of DNA. In this context, an analysis of bending wedge angles

in the atomistic system indicates that pyrimidine-purine base pair steps possess the highest bending flexibility. In particular, the AT basepair step was found to be the most flexible in the system. This is consistent with the parametrisation of oxDNA, in which this base pair step is assigned the lowest stacking energy. This agreement may thus support the different melting behaviour observed for $(AT)_6$ and $(AA)_6$ regions in Sec. 6.2 using oxDNA. The Benham model shows identical denaturation behaviour for both regions, as it quantifies the denaturation cost of a base pair by a single free-energy cost depending on the base-pair identity, and does not separately account for stacking and hydrogen bonding.

6.4 Summary

In this chapter, the coupled denaturation and writhing behaviour of DNA minicircles was studied. The calculated denaturation profiles suggest that not only the relative base content, but also the sequence arrangement can influence the local melting behaviour of DNA. Furthermore, it was found that the topology of the minicircle can introduce non-trivial long-range correlations in the denaturation profile of the double strand. In particular, correlated base pair opening of sites which are approximately opposite of each other is favoured. Sequence properties determine the preferred axis on which the two long-range correlated defects lie. Data from recent experimental cryo-EM experiments support the observed behaviour.

Comparison with atomistic simulations and a local model of double strand denaturation underlines that the coarse-grained model reflects a tradeoff between microscopic accuracy and computational expense. While it is able to probe the equilibrium denaturation behaviour of minicircles due to its efficiency, it retains enough microscopic detail to be able to predict nontrivial effects due to the molecular geometry. Effects stemming from the details of microscopic structure may introduce further subtleties into these results. Therefore, the results may not agree with the behaviour in a full quantitative sense. Despite this caveat, the coarse-grained model gives insight into the basic physical mechanisms of the interplay between sequence

and mechanical constraints imposed on the minicircle, which is supported by experimental findings. For higher microscopic resolution, or larger system sizes, this view can be complemented by atomistic and local denaturation models.

Chapter 7

Conclusions

7.1 Summary

In this thesis, the behaviour of DNA and RNA under mechanical stresses was investigated using coarse-grained models. Particular emphasis has been placed on the interaction of sequence properties with the large-scale conformational response of nucleic acid polymers. This interdependency is known to be particularly relevant in the realm of biological applications of nucleic acid mechanics, where the dynamical interplay between sequence and structure lies at the center of many mechanisms that read and process genetic information. However, neither fully microscopic models, nor continuum models that lack a faithful representation of the chiral double-helical structure of nucleic acids can address this important regime due to their respective intrinsic limitations, making this interplay a natural field of application for nucleotide-level coarse-grained approaches. Using this modelling philosophy, insight may be gained into the basic physical mechanisms involved in the mesoscale structure and dynamics of nucleic acids, which can then be further refined and complemented using more microscopically detailed models.

Studying mechanically stressed nucleic acids with oxDNA and oxRNA has led to the following results:

- Detailed simulations of extension "hat curves" of DNA and RNA as a function of twist and tension were performed. To the author's knowledge, this work reports the first such calculation that derives the molecular response from

a nucleotide-level model, and does not pre-suppose a particular form of the global elastic response of a molecular conformation.

- Extensive comparison to data from molecular tweezer experiments yielded quantitative agreement with the theoretical behaviour, including more subtle effects such as the influence of sequence on post-buckling slopes and the torque overshoot upon buckling

- A new conformational state of DNA was proposed, consisting of a plectonemic structure with a partially denatured end-loop. This state was predicted to occur for positive supercoiling at high stretching forces, as well as in the biologically relevant regime of intermediate stretching forces and negative supercoiling. Transitions between this novel state and other known molecular conformations were studied in detail.

- The sequence-dependent properties of DNA following from the presence of the tip-bubble plectoneme state were studied. It was predicted that due to this state, plectonemes tend to accumulate in weak, AT-rich stretches of the DNA double strand. Several extensions to existing single-molecule assays were proposed to test this prediction.

- It was shown that taking into account the co-localisation of denaturations and plectoneme structures accounts for the end-to-end fluctuation behaviour observed in recent single-molecule experiments more consistently than simple models that ignore the nucleation costs of bubbles and plectoneme structures.

- A slowdown of the diffusion of plectonemes was found as a consequence of denaturations in the plectoneme end-loop, providing a microscopic mechanism for the slow diffusion of these structures observed in recent single-molecule experiments.

- Hopping dynamics of plectoneme structures was observed that is reminiscent of the one reported in single-molecule experiments.

- Biological implications of the tip-bubble plectoneme state were explored. In particular, possible consequences of the tip-bubble plectoneme for genome organization, transcription regulation and DNA mismatch repair processes were considered.

- Finally, the denaturation behaviour of short minicircles of a length of about 100bp was studied. It was found that the closed topology of the system can lead to cooperative denaturations in AT-rich sites separated by about half a minicircle length, an effect supported by molecular imaging results using cryo-EM. Comparison to atomistic simulations and a local model of strand denaturation showed that the coarse-grained model is particularly suited to explore these topological effects.

In sum, the described phenomena show a complex interplay between sequence properties and the polymeric features of nucleic acids. The base sequence does not only determine the local interaction strength between the single strands, but can influence, through coupling to global conformational motifs such as plectonemes, the three-dimensional structure of the molecule over considerable distances. In that sense, imposing mechanical, and in particular torsional stress on the double strand uncovers a whole new level of information that is not directly encoded in the base sequence, but in the mechanical properties of the polymer. Ultimately, it is the interplay of the base sequence with the physical context of the heteropolymer on which it is stored that leads to the emergence of the three-dimensional molecular structures which feature in biochemical reaction mechanisms. Therefore, it seems unlikely that this physical context can be ignored when trying to grasp the full role of nucleic acid molecules in biological systems.

7.2 Further work

The work presented in this thesis suggests several avenues for further investigation.

- While the basic physics of the phenomena studied in this thesis is expected to hold over a range of solvent salt concentrations, many more detailed aspects such as the precise location of state boundaries and crossover forces are expected to be dependent on the ionic strength of the buffer, and consequently on the strength of long-range electrostatic interactions in the molecule. Therefore, it would be interesting to see the consequences of including explicit electrostatics in the coarse-grained models of DNA and RNA. The by now experimentally well characterized dependence of DNA extension behaviour on buffer strength may help test such augmented models.

- Although this work is focused on the effects of simple double-strand denaturation in twisted and stretched DNA, the same methodology may be easily extended to strands containing more complex secondary structures, such as three-way junctions or cruciforms. Simulations of the mechanical response of these structures may suggest signatures of these structures that can be detected in single-molecule experiments.

- Consequences of the tip-bubble plectoneme state on the plectoneme distribution could be determined for a larger range of biological sequences. It would be interesting to find out if certain sequence motifs with known functions, such as regulatory regions, represent preferred locations for plectoneme accumulation.

- To further investigate the biological implications of combined writhing and denaturation, combining the coarse-graining approach followed in the models used in this thesis with a coarse-grained representation of proteins would be an important next step. This might provide more detailed insights in the influence of the comformations studied in this thesis on specific biochemical mechanism, such as transcription or topoenzyme action.

Bibliography

[1] M. T. J. van Loenhout, M. V. de Grunt, and C. Dekker. Dynamics of DNA supercoils. *Science*, 338(6103):94–97, 2012.

[2] A. Tempestini, V. Cassina, D. Brogioli, R. Ziano, S. Erba, R. Giovannoni, M. G. Cerrito, D. Salerno, and F. Mantegazza. Magnetic tweezers measurements of the nanomechanical stability of DNA against denaturation at various conditions of pH and ionic strength. *Nucleic Acids Res.*, 41(3):2009–2019, 2013.

[3] D. Salerno, A. Tempestini, I. Mai, D. Brogioli, R. Ziano, V. Cassina, and F. Mantegazza. Single-molecule study of the DNA denaturation phase transition in the force-torsion space. *Phys. Rev. Lett.*, 109(11):118303, 2012.

[4] H. Brutzer, N. Luzzietti, D. Klaue, and R. Seidel. Energetics at the DNA supercoiling transition. *Biophys. J.*, 98(7):1267–1276, 2010.

[5] X. J. A. Janssen, J. Lipfert, T. Jager, R. Daudey, J. Beekman, and N. H. Dekker. Electromagnetic torque tweezers: a versatile approach for measurement of single-molecule twist and torque. *Nano Lett.*, 12(7):3634–3639, 2012.

[6] F. Mosconi, J. F. Allemand, D. Bensimon, and V. Croquette. Measurement of the torque on a single stretched and twisted DNA using magnetic tweezers. *Phys. Rev. Lett.*, 102(7):078301, 2009.

[7] F. Miescher. *Die Histochemischen und Physiologischen Arbeiten von Friedrich Miescher.*, chapter Letter to Wilhelm His, February 26th, 1869, pages 33–38. Vogel, Leipzig, 1869.

[8] A. Kossel. Ueber die chemische Zusammensetzung der Zelle. *Arch. Physiol.*, 181:181–186, 1891.

[9] P. Levene. The structure of yeast nucleic acid. IV. ammonia hydrolysis. *J. Bio. Chem.*, 40:415–424, 1919.

[10] E. Chargaff, R. Lipshitz, and C. Green. Composition of the deoxypentose nucleic acids of four genera of sea-urchin. *J. Biol. Chem.*, 1:155–160, 1952.

[11] W. T. Astbury and F.O. Bell. X-ray studies of thymonucleic acid. *Nature*, 141:747–748, 1938.

[12] R. E. Franklin and R. G. Gosling. Molecular configuration in sodium thymonucleate. *Nature*, 171(4356):740–741, 1953.

[13] J. D. Watson and F. H. Crick. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171(4356):737–738, 1953.

[14] Wikipedia Commons. DNA structure. http://commons.wikimedia.org, 2014.

[15] A. Ghosh and M. Bansal. A glossary of DNA structures from A to Z. *Acta Crystallogr. D*, 59:620–626, 2003.

[16] K. Hoogsteen. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta crystallographica*, 16(10):907–916, 1963.

[17] S. Neidle. *Principles of Nucleic Acid Structure.* Academic Press, 2007.

[18] S. C. Ha, K. Lowenhaupt, A. Rich, Y.-G. Kim, and K. K. Kim. Crystal structure of a junction between B-DNA and Z-DNA reveals two extruded bases. *Nature*, 437(7062):1183–1186, 2005.

[19] D. Elliott and M. Ladomery. *Molecular Biology of RNA.* Oxford University Press, 2011.

[20] M. Egli, N. Usman, S. G. Zhang, and A. Rich. Crystal structure of an Okazaki fragment at 2-A resolution. *Proc Natl Acad Sci U S A*, 89(2):534–538, Jan 1992.

[21] O. Romainczyk, B. Endeward, T. F. Prisner, and J. W. Engels. The RNA-DNA hybrid structure determined by EPR, CD and RNase H1. *Mol. Biosyst.*, 7(4):1050–1052, Apr 2011.

[22] Hippocrates of Kos. *The Hippocratic Treatises "On Generation" "On the Nature of the Child" "Deseases IV"*, chapter On Generation, pages 1–5. De Gruyter, 1981.

[23] G. Mendel. Versuche über Pflanzen-Hybriden. *Verhandlungen des Naturforschenden Vereines in Brünn*, 4:3–47, 1866.

[24] O. T. Avery, C. M. Macleod, and M. McCarty. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type III. *J. Exp. Med.*, 79(2):137–158, 1944.

[25] E. Schrödinger. *What is Life?* Cambridge University Press, 2012.

[26] B. Alberts, A. Johnson, P. Walter, J. Lewis, M. Raff, K. Roberts, and N. Orme. *Molecular Biology of the Cell.* Taylor & Francis, 5th edition, 2007.

[27] J. S. Mattick and I. V. Makunin. Non-coding RNA. *Hum. Mol. Genet.*, 15 Spec No 1:R17–R29, 2006.

[28] F. H. Crick. Central dogma of molecular biology. *Nature*, 227(5258):561–563, 1970.

[29] E. Pennisi. Genomics. ENCODE project writes eulogy for junk DNA. *Science*, 337(6099):1159, 1161, 2012.

[30] A. Jeltsch. Beyond watson and crick: DNA methylation and molecular enzymology of DNA methyltransferases. *Chembiochem*, 3(4):274–293, 2002.

[31] B. McClintock. The significance of responses of the genome to challenge. *Science*, 226(4676):792–801, 1984.

[32] D. Noble. Physiology is rocking the foundations of evolutionary biology. *Exp. Physiol.*, 98(8):1235–1243, 2013.

[33] J. D. Watson and F. H. Crick. Genetical implications of the structure of deoxyribonucleic acid. *Nature*, 171(4361):964–967, 1953.

[34] F. L. Holmes. The DNA replication problem, 1953-1958. *Trends Biochem. Sci.*, 23(3):117–120, 1998.

[35] M. Delbrück. On the replication of desoxiribonucleic acid (DNA). *Proc. Natl. Acad. Sci. U S A*, 40(9):783–788, 1954.

[36] R. Weil and J. Vinograd. The cyclic helix and cyclic coil forms of polyoma viral DNA. *Proc. Natl. Acad. Sci. U S A*, 50:730–738, 1963.

[37] R. R. Sinden. *DNA structure and function.* Academic Press Inc., 1994.

[38] A. D. Bates and A. Maxwell. *DNA topology.* Oxford University Press, 2005.

[39] A. Griswold. Genome packaging in prokaryotes: the circular chromosome of E. coli. *Nature Education*, 1:57, 2008.

[40] A. N. Kapanidis and T. R. Strick. Biology, one molecule at a time. *Trends Biochem. Sci.*, 34:234–243, 2009.

[41] G. Binnig and H. Rohrer. Scanning Tunneling Microscopy. *Trends in Physics*, 1:38–46, 1984.

[42] Y. L. Lyubchenko, L. S. Shlyakhtenko, and T. Ando. Imaging of nucleic acids with atomic force microscopy. *Methods*, 54(2):274–283, 2011.

[43] J. Adamcik, J.-H. Jeon, K. J. Karczewski, R. Metzler, and G. Dietler. Quantifying supercoiling-induced denaturation bubbles in DNA. *Soft Matter*, 8:8651–8658, 2012.

[44] J.-H. Jeon, J. Adamcik, G. Dietler, and R. Metzler. Supercoiling induces denaturation bubbles in circular DNA. *Phys. Rev. Lett.*, 105(20):208101, 2010.

[45] A. Y. Lushnikov, V. N. Potaman, and Y. L. Lyubchenko. Site-specific labeling of supercoiled DNA. *Nucleic Acids Res.*, 34(16):e111, 2006.

[46] B. D. Sattin and M. C. Goh. Direct observation of the assembly of RecA/DNA complexes by atomic force microscopy. *Biophys. J.*, 87(5):3430–3436, 2004.

[47] Y. Zhang, X. Hu, J. Sun, Y. Shen, J. Hu, X. Xu, and Z. Shao. High-resolution imaging and nano-manipulation of biological structures on surface. *Microsc. Res. Tech.*, 74(7):614–626, 2011.

[48] C. Bustamante, Z. Bryant, and S. B. Smith. Ten years of tension: single-molecule DNA mechanics. *Nature*, 421(6921):423–427, 2003.

[49] I. De Vlaminck and C. Dekker. Recent advances in magnetic tweezers. *Annu. Rev. Biophys.*, 41:453–472, 2012.

[50] Z. Bryant, F. C. Oberstrass, and A. Basu. Recent developments in single-molecule DNA mechanics. *Curr. Opin. Struct. Biol.*, 22(3):304–312, 2012.

[51] R. A. Forties and M. D. Wang. Discovering the Power of Single Molecules. *Cell*, 157(1):4–7, 2014.

[52] M. D. Wang, H. Yin, R. Landick, J. Gelles, and S. M. Block. Stretching DNA with optical tweezers. *Biophys. J.*, 72(3):1335–1346, 1997.

[53] S. B. Smith, Y. Cui, and C. Bustamante. Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules. *Science*, 271(5250):795–799, 1996.

[54] A. Mossa, M. Manosas, N. Forns, J. M. Huguet, and F. Ritort. Dynamic force spectroscopy of DNA hairpins: I. Force kinetics and free energy landscapes. *J. Stat. Mech.*, page P02060, 2009.

[55] M. Manosas, D. Collin, and F. Ritort. Force-dependent fragility in RNA hairpins. *Phys. Rev. Lett.*, 96(21):218301, 2006.

[56] J.-D. Wen, M. Manosas, P. T. X. Li, S. B. Smith, C. Bustamante, F. Ritort, and I. Tinoco. Force unfolding kinetics of RNA using optical tweezers. I. Effects of experimental variables on measured results. *Biophys. J.*, 92(9):2996–3009, 2007.

[57] A. La Porta and M. D. Wang. Optical torque wrench: angular trapping, rotation, and torque detection of quartz microparticles. *Phys. Rev. Lett.*, 92(19):190801, 2004.

[58] S. Forth, C. Deufel, M. Y. Sheinin, B. Daniels, J. P. Sethna, and M. D. Wang. Abrupt buckling transition observed during the plectoneme formation of individual DNA molecules. *Phys. Rev. Lett.*, 100(14):148301, 2008.

[59] T. R. Strick, M.-N. Dessinges, G. Charvin, N. H. Dekker, J. F. Allemand, D. Bensimon, and V. Croquette. Stretching of macromolecules and proteins. *Rep. Prog. Phys.*, 66:1–45, 2003.

[60] T. R. Strick, J. F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette. The elasticity of a single supercoiled DNA molecule. *Science*, 271(5257):1835–1837, 1996.

[61] T. R. Strick, J. F. Allemand, D. Bensimon, and V. Croquette. Behavior of supercoiled DNA. *Biophys. J.*, 74(4):2016–2028, 1998.

[62] T. R. Strick, V. Croquette, and D. Bensimon. Single-molecule analysis of DNA uncoiling by a type II topoisomerase. *Nature*, 404(6780):901–904, 2000.

[63] H. Meng, J. Bosman, T. van der Heijden, and J. van Noort. Coexistence of twisted, plectonemic, and melted DNA in small topological domains. *Biophys. J.*, 106(5):1174–1181, 2014.

[64] Z. Bryant, M. D. Stone, J. Gore, S. B. Smith, N. R. Cozzarelli, and C. Bustamante. Structural transitions and elasticity from torque measurements on DNA. *Nature*, 424(6946):338–341, 2003.

[65] J. Lipfert, J. W. J. Kerssemakers, T. Jager, and N. H. Dekker. Magnetic torque tweezers: measuring torsional stiffness in DNA and RecA-DNA filaments. *Nat. Methods*, 7(12):977–980, 2010.

[66] J. Lipfert, M. Wiggin, J. W. J. Kerssemakers, F. Pedaci, and N. H. Dekker. Freely orbiting magnetic tweezers to directly monitor changes in the twist of nucleic acids. *Nat. Commun.*, 2:439, 2011.

[67] F. C. Oberstrass, L. E. Fernandes, and Z. Bryant. Torque measurements reveal sequence-specific cooperative transitions in supercoiled DNA. *Proc. Natl. Acad. Sci. USA*, 109(16):6106–6111, 2012.

[68] T. Ramreddy, R. Sachidanandam, and T. R. Strick. Real-time detection of cruciform extrusion by single-molecule DNA nanomanipulation. *Nucleic Acids Res.*, 39:4275–4283, 2011.

[69] J. A. Abels, F. Moreno-Herrero, T. van der Heijden, C. Dekker, and N. H. Dekker. Single-molecule measurements of the persistence length of double-stranded RNA. *Biophys. J.*, 88(4):2737–2744, 2005.

[70] J. Lipfert, G. M. Skinner, J. M. Keegstra, T. Hensgens, T. Jager, D. Dulin, S. P. Donkers, F.-C. Chou, R. Das, and N. H. Dekker. Double-Stranded RNA under Force and Torque: Similarities and Striking Differences to Double-Stranded DNA. *submitted*, 2014.

[71] K. Howan, A. J. Smith, L. F. Westblade, N. Joly, W. Grange, S. Zorman, S. A. Darst, N. J. Savery, and T. R. Strick. Initiation of transcription-coupled repair characterized at single-molecule resolution. *Nature*, 490(7420):431–434, 2012.

[72] M. Manosas, M. M. Spiering, F. Ding, V. Croquette, and S. J. Benkovic. Collaborative coupling between polymerase and helicase for leading-strand synthesis. *Nucleic Acids Res.*, 40(13):6187–6198, 2012.

[73] D. A. Koster, A. Crut, S. Shuman, M.-A. Bjornsti, and N. H. Dekker. Cellular strategies for regulating DNA supercoiling: a single-molecule perspective. *Cell*, 142(4):519–530, 2010.

[74] M. Nöllmann, M. D. Stone, Z. Bryant, J. Gore, N. J. Crisona, S.-C. Hong, S. Mitelheiser, A. Maxwell, C. Bustamante, and N. R. Cozzarelli. Multiple modes of Escherichia coli DNA gyrase activity revealed by force and torque. *Nat. Struct. Mol. Biol.*, 14(4):264–271, 2007.

[75] Optical tweezer setup. Block Lab website, Stanford University, 2014.

[76] C. Joo, H. Balci, Y. Ishitsuka, C. Buranachai, and T. Ha. Advances in single-molecule fluorescence methods for molecular biology. *Annu. Rev. Biochem.*, 77:51–76, 2008.

[77] P. Gross, N. Laurens, L.B Oddershede, U. Bockelmann, E.J.G. Peterman, and G. J. L. Wuite. Quantifying how DNA stretches, melts and changes twist under tension. *Nature Physics*, 7:731–736, 2011.

[78] G. A. King, P. Gross, U. Bockelmann, M. Modesti, G. J. L. Wuite, and E. J. G. Peterman. Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching using fluorescence microscopy. *Proc. Natl. Acad. Sci. U S A*, 110(10):3859–3864, 2013.

[79] A. L. Forget and S. C. Kowalczykowski. Single-molecule imaging of DNA pairing by RecA reveals a three-dimensional homology search. *Nature*, 482(7385):423–427, 2012.

[80] S. M. Hamdan, J. J. Loparo, M. Takahashi, C. C. Richardson, and A. M van Oijen. Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis. *Nature*, 457(7227):336–339, 2009.

[81] A. M. van Oijen and J. J. Loparo. Single-molecule studies of the replisome. *Annu. Rev. Biophys.*, 39:429–448, 2010.

[82] R. Vafabakhsh and T. Ha. Extreme bendability of DNA less than 100 base pairs long revealed by single-molecule cyclization. *Science*, 337(6098):1097–1101, 2012.

[83] J. Y. Lee, B. Okumus, D. S. Kim, and T. Ha. Extreme conformational diversity in human telomeric DNA. *Proc. Natl. Acad. Sci. U S A*, 102(52):18938–18943, 2005.

[84] P. Gross, G. Farge, E. J. G. Peterman, and G. J. L. Wuite. Combining optical tweezers, single-molecule fluorescence microscopy, and microfluidics for studies of DNA-protein interactions. *Methods. Enzymol.*, 475:427–453, 2010.

[85] S. Hohng, R. Zhou, M. K. Nahas, J. Yu, K. Schulten, D. M. J. Lilley, and T. Ha. Fluorescence-force spectroscopy maps two-dimensional reaction landscape of the holliday junction. *Science*, 318(5848):279–283, 2007.

[86] M. Lee, S. H. Kim, and S.-C. Hong. Minute negative superhelicity is sufficient to induce the B-Z transition in the presence of low tension. *Proc. Natl. Acad. Sci. U S A*, 107(11):4985–4990, 2010.

[87] T. Cordes, Y. Santoso, A. I. Tomescu, K. Gryte, L. C. Hwang, B. Camar, S. Wigneshweraraj, and A. N. Kapanidis. Sensing DNA opening in transcription using quenchable Förster resonance energy transfer. *Biochemistry*, 49(43):9171–9180, 2010.

[88] J. E. Coats, Y. Lin, E. Rueter, L. J. Maher, and I. Rasnik. Single-molecule FRET analysis of DNA binding and bending by yeast HMGB protein Nhp6A. *Nucleic Acids Res.*, 41(2):1372–1381, 2013.

[89] S. Liu, G. Bokinsky, N. G. Walter, and X. Zhuang. Dissecting the multistep reaction pathway of an RNA enzyme by single-molecule kinetic "fingerprinting". *Proc. Natl. Acad. Sci. U S A*, 104(31):12634–12639, 2007.

[90] C. Bustamante, W. Cheng, and Y. X. Mejia. Revisiting the central dogma one molecule at a time. *Cell*, 144(4):480–497, 2011.

[91] S. Hohng, S. Lee, J. Lee, and M. H. Jo. Maximizing information content of single-molecule FRET experiments: multi-color FRET and FRET combined with force or torque. *Chem. Soc. Rev.*, 43(4):1007–1013, 2014.

[92] X. S. Xie, P. J. Choi, G.-W. Li, N. K. Lee, and G. Lia. Single-molecule approach to molecular biology in living bacterial cells. *Annu. Rev. Biophys.*, 37:417–444, 2008.

[93] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.*, 117:51795197, 1995.

[94] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. CHARMM: the biomolecular simulation program. *J. Comput. Chem.*, 30(10):1545–1614, 2009.

[95] M. Orozco, A. Pérez, A. Noy, and F. J. Luque. Theoretical methods for the simulation of nucleic acids. *Chem. Soc. Rev.*, 32(6):350–364, 2003.

[96] S. A. Harris. Modelling the biomechanical properties of DNA using computer simulation. *Phil. Trans. A*, 364(1849):3319–3334, 2006.

[97] R. Lavery, K. Zakrzewska, D. Beveridge, T. C. Bishop, D. A. Case, T. Cheatham, S. Dixit, B. Jayaram, F. Lankaš, C. Laughton, J. H. Maddocks, A. Michon, R. Osman, M. Orozco, A. Perez, T. Singh, N. Spackova, and J. Sponer. A systematic molecular dynamics study of nearest-neighbor

effects on base pair and base pair step conformations and fluctuations in B-DNA. *Nucleic Acids Res.*, 38(1):299–313, 2010.

[98] A. Noy and R. Golestanian. The chirality of DNA: elasticity cross-terms at base-pair level including A-tracts and the influence of ionic strength. *J. Phys. Chem. B*, 114(23):8022–8031, 2010.

[99] A. Noy and R. Golestanian. Length scale dependence of DNA mechanical properties. *Phys. Rev. Lett.*, 109(22):228101, 2012.

[100] S. Kannan and M. Zacharias. Simulation of DNA double-strand dissociation and formation during replica-exchange molecular dynamics simulations. *Phys. Chem. Chem. Phys.*, 11(45):10589–10595, 2009.

[101] A. Pérez, F. J. Luque, and M. Orozco. Frontiers in Molecular Dynamics Simulations of DNA. *Acc. Chem. Res.*, 45:196–205, 2012.

[102] C. A. Laughton and S. A. Harris. The atomistic simulation of DNA. *WIREs Comput. Mol. Sci.*, 1:590–600, 2011.

[103] F. Rico, L. Gonzalez, I. Casuso, M. Puig-Vidal, and S. Scheuring. High-speed force spectroscopy unfolds titin at the velocity of molecular dynamics simulations. *Science*, 342(6159):741–743, 2013.

[104] P. M. D. Severin, X. Zou, H. E. Gaub, and K. Schulten. Cytosine methylation alters DNA mechanical properties. *Nucleic Acids Res.*, 39(20):8740–8751, 2011.

[105] G. L. Randall, L. Zechiedrich, and B. M. Pettitt. In the absence of writhe, DNA relieves torsional stress with localized, sequence-dependent structural failure to preserve B-form. *Nucleic Acids Res.*, 37(16):5568–5577, 2009.

[106] J. Wereszczynski and I. Andricioaei. On structural transitions, thermodynamic equilibrium, and the phase diagram of DNA and RNA duplexes under torque and tension. *Proc. Natl. Acad. Sci. U S A*, 103(44):16200–16205, 2006.

[107] S. A. Harris, Z. A. Sands, and C. A. Laughton. Molecular dynamics simulations of duplex stretching reveal the importance of entropy in determining the biomechanical properties of DNA. *Biophys. J.*, 88(3):1684–1691, 2005.

[108] S. Piana. Structure and energy of a DNA dodecamer under tensile load. *Nucleic Acids Res.*, 33(22):7029–7038, 2005.

[109] T. B. Liverpool, S. A. Harris, and C. A. Laughton. Supercoiling and denaturation of DNA loops. *Phys. Rev. Lett.*, 100(23):238103, 2008.

[110] J. S. Mitchell, C. A. Laughton, and S. A. Harris. Atomistic simulations reveal bubbles, kinks and wrinkles in supercoiled DNA. *Nucleic Acids Res.*, 39(9):3928–3938, 2011.

[111] O. Kratky and G. Porod. Röntgenuntersuchung gelöster Fadenmoleküle. *Rec. Trav. Chim. Pays-Bas.*, 68:1106–1123, 1949.

[112] J. F. Marko and E. D. Siggia. Statistical mechanics of supercoiled DNA. *Phys. Rev. E*, 52(3):2912–2938, 1995.

[113] J. D. Moroz and P. Nelson. Torsional directed walks, entropic elasticity, and DNA twist stiffness. *Proc. Natl. Acad. Sci. USA*, 94(26):14418–14422, 1997.

[114] C. Bouchiat and M. Mezard. Elasticity model of a supercoiled DNA molecule. *Phys. Rev. Lett.*, 80:1556, 1998.

[115] T. Odijk. Stiff Chains and Filaments under Tension. *Macromolecules*, 28:7016–7018, 1995.

[116] J. D. Moroz and P. Nelson. Entropic Elasticity of Twist-Storing Polymers. *Macromolecules*, 31:6333 – 6347, 1998.

[117] J. F. Marko. Torque and dynamics of linking number relaxation in stretched supercoiled DNA. *Phys. Rev. E*, 76(2):021926, 2007.

[118] J. F. Marko and S. Neukirch. Competition between curls and plectonemes near the buckling transition of stretched supercoiled DNA. *Phys. Rev. E*, 85(1):011908, 2012.

[119] J. F. Marko and S. Neukirch. Global force-torque phase diagram for the DNA double helix: structural transitions, triple points, and collapsed plectonemes. *Phys. Rev. E*, 88(6):062722, 2013.

[120] S. Neukirch and J. F. Marko. Analytical description of extension, torque, and supercoiling radius of a stretched twisted DNA. *Phys. Rev. Lett.*, 106(13):138104, 2011.

[121] D. Argudo and P. K. Purohit. The dependence of DNA supercoiling on solution electrostatics. *Acta Biomater.*, 8(6):2133–2143, 2012.

[122] J. F. Marko. Stretching must twist DNA. *Europhys. Lett.*, 38:183–188, 1997.

[123] S. Goyal and N.C. Perkins. Looping mechanics of rods and DNA with non-homogeneous and discontinuous stiffness. *Int. J. Nonlinear Mech.*, 43:1121–1129, 2008.

[124] R. Metzler, T. Ambjörnsson, A. Hanke, and H. C. Fogedby. Single DNA denaturation and bubble dynamics. *J. Phys. Condens. Matter*, 21(3):034111, 2009.

[125] D. Poland and H. A. Scheraga. Occurrence of a phase transition in nucleic acid models. *J. Chem. Phys.*, 45(5):1464–1469, 1966.

[126] Y. Kafri, D. Mukamel, and L. Peliti. Why is the DNA denaturation transition first order? *Phys. Rev. Lett.*, 85(23):4988–4991, 2000.

[127] E. Carlon, E. Orlandini, and A. L. Stella. The roles of stiffness and excluded volume in DNA denaturation. *Phys. Rev. Lett.*, 88:198101, 2002.

[128] A. Kabakçioğlu, E. Orlandini, and D. Mukamel. Supercoil formation in DNA denaturation. *Phys. Rev. E*, 80(1 Pt 1):010903, 2009.

[129] T. Ambjörnsson, S. K. Banik, M. A. Lomholt, and R. Metzler. Master equation approach to DNA breathing in heteropolymer DNA. *Phys. Rev. E*, 75(2 Pt 1):021908, 2007.

[130] A. Bar, Y. Kafri, and D. Mukamel. Loop dynamics in DNA denaturation. *Phys. Rev. Lett.*, 98(3):038103, 2007.

[131] J. SantaLucia and D. Hicks. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.*, 33:415–440, 2004.

[132] C. J. Benham. Theoretical analysis of heteropolymeric transitions in superhelical DNA molecules of specified sequence. *J. Chem. Phys.*, 92:6294–6305, 1990.

[133] C. J. Benham. Energetics of the strand separation transition in superhelical DNA. *J. Mol. Biol.*, 225(3):835–847, 1992.

[134] W. R. Bauer and C. J. Benham. The free energy, enthalpy and entropy of native and of partially denatured closed circular DNA. *J. Mol. Biol.*, 234(4):1184–1196, 1993.

[135] R. M. Fye and C. J. Benham. Exact method for numerically analyzing a model of local denaturation in superhelically stressed DNA. *Phys. Rev. E*, 59, 1999.

[136] D. Jost, A. Zubair, and R. Everaers. Bubble statistics and positioning in superhelically stressed DNA. *Phys. Rev. E*, 84(3), 2011.

[137] S. Cocco and R. Monasson. Statistical Mechanics of Torque Induced Denaturation of DNA. *Phys. Rev. Lett.*, 83:5178–5181, 1999.

[138] C. Nisoli and A. R. Bishop. Thermomechanics of DNA: theory of thermal stability under load. *Phys. Rev. Lett.*, 107(6):068102, 2011.

[139] M. Zoli. Thermodynamics of twisted DNA with solvent interaction. *J. Chem. Phys.*, 135(11):115101, 2011.

[140] A. V. Vologodskii, S. D. Levene, K. V. Klenin, M. Frank-Kamenetskii, and N. R. Cozzarelli. Conformational and thermodynamic properties of supercoiled DNA. *J. Mol. Biol.*, 227(4):1224–1243, 1992.

[141] K. V. Klenin, A. V. Vologodskii, V. V. Anshelevich, A. M. Dykhne, and M. D. Frank-Kamenetskii. Computer simulation of DNA supercoiling. *J. Mol. Biol.*, 217(3):413–419, 1991.

[142] K. V. Klenin, M. D. Frank-Kamenetskii, and J. Langowski. Modulation of intramolecular interactions in superhelical DNA by curved sequences: a Monte Carlo simulation study. *Biophys. J.*, 68(1):81–88, 1995.

[143] A. V. Vologodskii and J. F. Marko. Extension of torsionally stressed DNA by external force. *Biophys. J.*, 73(1):123–132, 1997.

[144] G. Charvin, A. Vologodskii, D. Bensimon, and V. Croquette. Braiding DNA: experiments, simulations, and models. *Biophys. J.*, 88(6):4124–4136, 2005.

[145] B. C. Daniels and J. P. Sethna. Nucleation at the DNA supercoiling transition. *Phys. Rev. E*, 83(4):041924, 2011.

[146] T. Schlick and W. K. Olson. Supercoiled DNA energetics and dynamics by computer simulation. *J. Mol. Biol.*, 223(4):1089–1119, 1992.

[147] T. Schlick. Modeling superhelical DNA: recent analytical and dynamic approaches. *Curr. Opin. Struct. Biol.*, 5(2):245–262, 1995.

[148] J. P. K. Doye, T. E. Ouldridge, A. A. Louis, F. Romano, P. Šulc, C. Matek, B. E. K. Snodin, L. Rovigatti, J. S. Schreck, R. M. Harrison, and W. P. J. Smith. Coarse-graining DNA for simulations of DNA nanotechnology. *Phys. Chem. Chem. Phys.*, 15:20395–20414, 2013.

[149] T. E. Ouldridge. *Coarse-grained modelling of DNA and DNA self-assembly.* PhD thesis, University of Oxford, 2011.

[150] S. P. Mielke, N. Grønbech-Jensen, V. V. Krishnan, W. H. Fink, and C. J. Benham. Brownian dynamics simulations of sequence-dependent duplex denaturation in dynamically superhelical DNA. *J. Chem. Phys.*, 123(12):124911, 2005.

[151] F. Trovato and V. Tozzini. Supercoiling and local denaturation of plasmids with a minimalist DNA model. *J. Phys. Chem. B*, 112(42):13197–13200, 2008.

[152] P. C. Nelson. *Biological Physics: Energy, Information, Life.* W. H. Freeman, 2013.

[153] J. Gore, Z. Bryant, M. Nöllmann, M. U. Le, N. R. Cozzarelli, and C. Bustamante. DNA overwinds when stretched. *Nature*, 442(7104):836–839, 2006.

[154] T. Lionnet, S. Joubaud, R. Lavery, D. Bensimon, and V. Croquette. Wringing out DNA. *Phys. Rev. Lett.*, 96(17):178102, 2006.

[155] F.-C. Chou, J. Lipfert, and R. Das. Blind predictions of DNA and RNA tweezers experiments with force and torque. *PLoS Comput. Biol.*, 10(8):e1003756, 2014.

[156] J. Lipfert, S. Doniach, R. Das, and D. Herschlag. Understanding nucleic Acid-ion interactions. *Annu. Rev. Biochem.*, 83:813–841, 2014.

[157] R. D. Kamien, T. C. Lubensky, P. Nelson, and C. S. O'Hern. Direct determination of DNA twist-stretch coupling. *Europhys. Lett.*, 38:237–242, 1997.

[158] R. D. Kamien. The geometry of soft materials: a primer. *Rev. Mod. Phys.*, 74:953–971, 2002.

[159] F. B. Fuller. The writhing number of a space curve. *Proc. Natl. Acad. Sci. U S A*, 68(4):815–819, 1971.

[160] F. B. Fuller. Decomposition of the linking number of a closed ribbon: A problem from molecular biology. *Proc. Natl. Acad. Sci. USA*, 75(8):3557–3561, 1978.

[161] J. H. White. Self-linking and the Gauss Integral in Higher Dimensions. *Am. J. Math.*, 91:693–728, 1969.

[162] G. Călugăreanu. L' Intégrale de Gauss et l'analyse des Noeuds tridimensionnels. *Rev. Roum. Math. Pures et Appl.*, 4:5–20, 1959.

[163] M. Epple. Orbits of Asteroids, a Braid, and the First Link Invariant. *The Mathematical Intelligencer*, 20:45–52, 1998.

[164] M. A. Berger and C. Prior. The Writhe of Open and Closed Curves. *J. Phys. A: Math. Gen.*, 39:8321–8348, 2006.

[165] V. Rossetto and A. C. Maggs. Writhing geometry of open DNA. *J. Chem. Phys.*, 118(21):9864–9874, 2003.

[166] J. F. Marko. Supercoiled and braided DNA under tension. *Phys. Rev. E*, 55:1758–1772, 1997.

[167] K. Klenin and J. Langowski. Computation of writhe in modeling of supercoiled DNA. *Biopolymers*, 54(5):307–317, 2000.

[168] M. Sayar, B. Avşaroğlu, and A. Kabakçioğlu. Twist-writhe partitioning in a coarse-grained DNA minicircle model. *Phys. Rev. E*, 81:041916, 2010.

[169] J. Cantarella. On Comparing the Writhe of a Smooth Curve to the Writhe of an Inscribed Polygon. *SIAM J. Numer. Anal.*, 42:18461861, 2005.

[170] A. V. Vologodskii, V. V. Anshelevich, A. V. Lukashin, and M. D. Frank-Kamenetskii. Statistical mechanics of supercoils and the torsional stiffness of the DNA double helix. *Nature*, 280(5720):294–298, 1979.

[171] R. Everaers, S. Kumar, and C. Simm. A unifi

ed description of poly- and oligonucleotide DNA melting: Nearest-neighbor, Poland-Sheraga, and lattice models. *Phys. Rev. E*, 75:041918041930, 2007.

[172] T. E. Ouldridge, A. A. Louis, and J. P. K. Doye. Structural, mechanical, and thermodynamic properties of a coarse-grained DNA model. *J. Chem. Phys.*, 134(8):085101, 2011.

[173] P. Šulc, F. Romano, T. E. Ouldridge, L. Rovigatti, J. P. K. Doye, and A. A. Louis. Sequence-dependent thermodynamics of a coarse-grained DNA model. *J. Chem. Phys.*, 137(13):135101, 2012.

[174] C. Matek, T. E. Ouldridge, A. Levy, J. P. K. Doye, and A. A. Louis. DNA cruciform arms nucleate through a correlated but asynchronous cooperative mechanism. *J. Phys. Chem. B*, 116(38):11616–11625, 2012.

[175] T. E. Ouldridge, P. Šulc, F. Romano, J. P. K. Doye, and A. A. Louis. DNA hybridization kinetics: zippering, internal displacement and sequence dependence. *Nucleic Acids Res.*, 41(19):8886–8895, 2013.

[176] F. Romano, D. Chakraborty, J. P. K. Doye, T. E. Ouldridge, and A. A. Louis. Coarse-grained simulations of DNA overstretching. *J. Chem. Phys.*, 138(8):085101, 2013.

[177] T. E. Ouldridge, A. A. Louis, and J. P. K. Doye. DNA nanotweezers studied with a coarse-grained model of DNA. *Phys. Rev. Lett.*, 104(17):178101, 2010.

[178] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. UCSF chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.*, 25(13):1605–1612, 2004.

[179] F. Romano, A. Hudson, J. P. K. Doye, T. E. Ouldridge, and A. A. Louis. The effect of topology on the structure and free energy landscape of DNA kissing complexes. *J. Chem. Phys.*, 136(21):215102, 2012.

[180] T. E. Ouldridge, R. L. Hoare, A. A. Louis, J. P. K. Doye, J. Bath, and A. J. Turberfield. Optimizing DNA nanotechnology through coarse-grained modeling: a two-footed DNA walker. *ACS Nano*, 7(3):2479–2490, 2013.

[181] N. Srinivas, T. E. Ouldridge, P. Sulc, J. M. Schaeffer, B. Yurke, A. A. Louis, J. P. K. Doye, and E. Winfree. On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Res.*, 41(22):10641–10658, 2013.

[182] P. Šulc, F. Romano, T. E. Ouldridge, J. P. K. Doye, and A. A. Louis. A nucleotide-level coarse-grained model of RNA. *J. Chem. Phys.*, 140(23):235102, 2014.

[183] P. Šulc. *Coarse-grained modelling of nucleic acids.* PhD thesis, University of Oxford, 2014.

[184] J. Russo, P. Tartaglia, and F. Sciortino. Reversible gels of patchy particles: role of the valence. *J. Chem. Phys.*, 131(1):014504, 2009.

[185] A. A. Louis. Coarse-graining dynamics by telescoping down time-scales: comment for Faraday FD144. *Faraday Discuss.*, 133:323, 2010.

[186] C. Maffeo, R. Schöpflin, H. Brutzer, R. Stehr, A. Aksimentiev, G. Wedemann, and R. Seidel. DNA-DNA interactions in tight supercoils are described by a small effective charge density. *Phys. Rev. Lett.*, 105(15):158101, 2010.

[187] T. E. Haran and U. Mohanty. The unique structure of A-tracts and intrinsic DNA bending. *Q. Rev. Biophys.*, 42(1):41–81, 2009.

[188] E. Segal and J. Widom. Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.*, 19(1):65–71, 2009.

[189] T. Dršata, N. Špačková, P. Jurečka, M. Zgarbová, J. Šponer, and F. Lankaš. Mechanical properties of symmetric and asymmetric DNA A-tracts: implications for looping and nucleosome positioning. *Nucleic Acids Res.*, 42(11):7383–7394, 2014.

[190] S. Forth, M. Y. Sheinin, J. Inman, and M. D. Wang. Torque measurement at the single-molecule level. *Ann. Rev. Biophys.*, 42:583–604, 2013.

[191] T. Lionnet and F. Lankaš. Sequence-dependent twist-stretch coupling in DNA. *Biophys. J.*, 92(4):L30–L32, 2007.

[192] M. Ghorbani and F. Mohammad-Rafiee. Twist-stretch correlation of DNA. *Phys. Rev. E*, 78(6 Pt 1):060901, 2008.

[193] M. Emanuel, G. Lanzani, and H. Schiessel. Multi-plectoneme phase of double-stranded DNA under torsion. *Phys. Rev. E*, 88:022706–1–20, 2013.

[194] T. R. Strick, V. Croquette, and D. Bensimon. Homologous pairing in stretched supercoiled DNA. *Proc. Natl. Acad. Sci. U S A*, 95(18):10579–10583, 1998.

[195] R. T. Dame, C. Wyman, and N. Goosen. H-NS mediated compaction of DNA visualised by atomic force microscopy. *Nucleic Acids Res.*, 28(18):3504–3510, 2000.

[196] C. J. Lim, L. J. Kenney, and J. Yan. Single-molecule studies on the mechanical interplay between DNA supercoiling and H-NS DNA architectural properties. *Nucleic Acids Res.*, 42(13):8369–8378, 2014.

[197] G. Altan-Bonnet, A. Libchaber, and O. Krichevsky. Bubble dynamics in double-stranded DNA. *Phys. Rev. Lett.*, 90(13):138101, 2003.

[198] J. Roca. The torsional state of DNA within the chromosome. *Chromosoma*, 120(4):323–334, 2011.

[199] L. Baranello, D. Levens, A. Gupta, and F. Kouzine. The importance of being supercoiled: how DNA mechanics regulate dynamic processes. *Biochim. Biophys. Acta*, 1819(7):632–638, 2012.

[200] R. T. Dame. The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin. *Mol. Microbiol.*, 56(4):858–870, 2005.

[201] A. Travers and G. Muskhelishvili. A common topology for bacterial and eukaryotic transcription initiation? *EMBO Rep.*, 8(2):147–151, 2007.

[202] C. H. Laundon and J. D. Griffith. Curved helix segments can uniquely orient the topology of supertwisted DNA. *Cell*, 52(4):545–549, 1988.

[203] S. Maurer, J. Fritz, G. Muskhelishvili, and A. Travers. RNA polymerase and an activator form discrete subcomplexes in a transcription initiation complex. *EMBO J.*, 25(16):3784–3790, 2006.

[204] B. ten Heggeler-Bordier, W. Wahli, M. Adrian, A. Stasiak, and J. Dubochet. The apical localization of transcribing RNA polymerases on supercoiled DNA prevents their rotation around the template. *EMBO J.*, 11(2):667–672, 1992.

[205] S. T. Smale and J. T. Kadonaga. The RNA polymerase II core promoter. *Annu. Rev. Biochem.*, 72:449–479, 2003.

[206] C. B. Harley and R. P. Reynolds. Analysis of E. coli promoter sequences. *Nucleic Acids Res.*, 15(5):2343–2361, 1987.

[207] H. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore, and J. Darnell. *Molecular Cell Biology.* Freeman (New York), 2000.

[208] R. P. Sinha and D. P. Häder. UV-induced DNA damage and repair: a review. *Photochem. Photobiol. Sci.*, 1(4):225–236, 2002.

[209] E. C. Friedberg, G. C. Walker, W. Siede, R. D. Wood, R. A. Schultz, and T. Ellenberger. *DNA Repair and Mutagenesis.* ASM Press, 2006.

[210] H. Park, K. Zhang, Y. Ren, S. Nadji, N. Sinha, J.-S. Taylor, and C. Kang. Crystal structure of a DNA decamer containing a cis-syn thymine dimer. *Proc. Natl. Acad. Sci. U S A*, 99(25):15965–15970, 2002.

[211] P. G. de Gennes. Tight knots. *Macromolecules*, 17(4):703–704, 1984.

[212] R. D. Hills, L. Lu, and G. A. Voth. Multiscale coarse-graining of the protein energy landscape. *PLoS Comput. Biol.*, 6(6):e1000827, 2010.

[213] T. E. Cloutier and J. Widom. Spontaneous sharp bending of double-stranded DNA. *Mol. Cell.*, 14(3):355–362, 2004.

[214] T. E. Cloutier and J. Widom. DNA twisting flexibility and the formation of sharply looped protein-DNA complexes. *Proc. Natl. Acad. Sci. U S A*, 102(10):3645–3650, 2005.

[215] J. Yan and J. F. Marko. Localized single-stranded bubble mechanism for cyclization of short double helix DNA. *Phys. Rev. Lett.*, 93(10):108108, 2004.

[216] Q. Du, C. Smith, N. Shiffeldrim, M. Vologodskaia, and A. Vologodskii. Cyclization of short DNA fragments and bending fluctuations of the double helix. *Proc. Natl. Acad. Sci. U S A*, 102(15):5397–5402, 2005.

[217] Q. Du, A. Kotlyar, and A. Vologodskii. Kinking the double helix by bending deformation. *Nucleic Acids Res.*, 36(4):1120–1128, 2008.

[218] D. Demurtas, A. Amzallag, E. J. Rawdon, J. H. Maddocks, J. Dubochet, and A. Stasiak. Bending modes of DNA directly addressed by cryo-electron microscopy of DNA minicircles. *Nucleic Acids Res.*, 37(9):2882–2893, 2009.

[219] T. A. Lionberger, D. Demurtas, G. Witz, J. Dorier, T. Lillian, E. Meyhfer, and A. Stasiak. Cooperative kinking at distant sites in mechanically stressed DNA. *Nucleic Acids Res.*, 39(22):9820–9832, 2011.

[220] F. Lankaš, R. Lavery, and J. H. Maddocks. Kinking occurs during molecular dynamics simulations of small DNA minicircles. *Structure*, 14(10):1527–1534, 2006.

[221] J. M. Fogg, N. Kolmakova, I. Rees, S. Magonov, H. Hansma, J. J. Perona, and E. L. Zechiedrich. Exploring writhe in supercoiled minicircle DNA. *J. Phys. Condens. Matter*, 18(14):S145–S159, 2006.

[222] D. Norouzi, F. Mohammad-Rafiee, and R. Golestanian. Effect of bending anisotropy on the 3D conformation of short DNA loops. *Phys. Rev. Lett.*, 101(16):168103, 2008.

[223] F. Kouzine, S. Sanford, Z. Elisha-Feil, and D. Levens. The functional response of upstream DNA to dynamic supercoiling in vivo. *Nat. Struct. Mol. Biol.*, 15(2):146–154, 2008.

[224] E. Guitter and S. Leibler. On supercoiling instability in closed DNA. *Europhys. Lett.*, 17:643–648, 1992.

[225] S. A. Allison. Brownian dynamics simulation of wormlike chains. Fluorescence depolarization and depolarized light scattering. *Macromolecules*, 19:118–124, 1986.

[226] M. Krepl, M. Zgarbová, P. Stadlbauer, M. Otyepka, P. Banáš, Jaroslav Koča, T. E. Cheatham, P. Jurečka, and J. Sponer. Reference simulations of non-canonical nucleic acids with different chi variants of the AMBER force field: quadruplex DNA, quadruplex RNA and Z-DNA. *J. Chem. Theory Comput.*, 8(7):2506–2520, 2012.

[227] D. Frenkel and B. Smit. *Understanding Molecular Simulation, second edition.* Academic Press, 2002.

[228] J. Delambre. *Mem. Acad. Turin*, 5:143, 1791.

[229] L. Verlet. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.*, 159:98–103, 1967.

[230] A. Dullweber, B. Leimkuhler, and R. McLachlan. Symplectic splitting methods for rigid body molecular dynamics. *J. Chem. Phys.*, 107:5840–5851, 1997.

[231] H. C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.*, 72:2384–2393, 1980.

[232] R. Tacutu, T. Craig, A. Budovsky, D. Wuttke, G. Lehmann, D. Taranukha, J. Costa, V. E. Fraifeld, and J. P. de Magalhães. Human Ageing Genomic Resources: integrated databases and tools for the biology and genetics of ageing. *Nucleic Acids Res.*, 41(Database issue):D1027–D1033, 2013.

[233] P. Lamb and L. Crawford. Characterization of the human p53 gene. *Mol. Cell. Biol.*, 6(5):1379–1385, 1986.

Appendix A

Simulation methods

A.1 Verlet integrator

In classical molecular dynamics simulation codes, NEWTON's equations have to be numerically integrated in order to obtain the time evolution of the system studied [227]. Both coarse-grained models used in this work, oxDNA and oxRNA, have been designed with differential functional forms of their interaction potentials, in order to allow for the use of standard integration techniques.

One of the most common numerical integrators for Molecular Dynamics simulations has been introduced by DELABMRE already in 1791 [228], but has become widely known as the VERLET algorithm in the context of computational physics [229]. It can be shown that this algorithm conserves the symplectic phase space structure of HAMILTONIAN rigid body models [227, 230], and minimizes energy drift in the course of a simulation. Throughout this work, a variant of the VERLET algorithm described in Ref. [230] was used to propagate both the linear and angular degrees of freedom of the system in time.

Mass m and moment of interia I of the rigid bodies are set to 1 in simulation units for both model used in this thesis. Furthermore, it was assumed that the form of the moment of inertia of a nucleotide corresponds to that of a sphere. The experimental average mass of a nucleotide is 315.75 AMU for DNA and 321.4 AMU for RNA respectively. Energy in simulations was measured such that one unit of energy was equal to $E = 4.142 \times 10^{-20}$J. Hence, correspondingly, $k_B T = 0.1$ simulation units at $T = 300$K. Finally, the simulation unit of length corresponds

to 0.8518nm. Together, setting units in this way fixes the simulation time unit to 3.03×10^{-12}s for oxRNA and 3.06×10^{-12}s for oxRNA.

A.2 Thermostat

While the VERLET algorithm described in Sec. A.1 propagates the system in time according to NEWTON's equation, using it in isolation would lead to ballistic particle motion, and is therefore inappropriate for implicit-solvent models such as oxDNA and oxRNA. In order to be able to perform molecular dynamics simulations in the NVT ensemble, the system has to be coupled to a heat bath at temperature T using a thermostat. For the simulations presented in this thesis, a variant of the thermostat introduced by ANDERSEN [231] was used, which is described in more detail in Ref. [184]. Briefly, the thermostat acts after a number N_{Newt} of standard steps performed by the integrator described in Sec. A.1 by resetting the linear and angular velocities of each particle in the system with probabilities p_v and p_ω respectively. The values to which the velocities are set are sampled from a BOLTZMANN distribution at the given temperature T. The default settings chosen for simulations presented in this thesis were $N_{\mathrm{Newt}} = 103$, and a translational diffusion coefficient of $D_{\mathrm{lin}} = 2.5$ in simulation units, corresponding to SI unit values of $6.0 \times 10^{-7}\mathrm{m}^2\mathrm{s}^{-1}$ for oxDNA and $5.8 \times 10^{-7}\mathrm{m}^2\mathrm{s}^{-1}$ in oxRNA. The rotational diffusion coefficient was set to $D_{\mathrm{rot}} = 3D_{\mathrm{lin}}$. Following Ref. [184], this leads to the following values for the velocity resetting probabilities:

$$p_v = \frac{2k_BTN_{\mathrm{Newt}}\delta t}{k_BTN_{\mathrm{Newt}}\delta t + 2mD_{\mathrm{lin}}} \tag{A.1}$$

$$p_\omega = \frac{2k_BTN_{\mathrm{Newt}}\delta t}{k_BTN_{\mathrm{Newt}}\delta t + 2ID_{\mathrm{rot}}} \tag{A.2}$$

Here, δt is the integration time step, which was chosen as 0.004 simulation time units, corresponding to 12.2fs. Due to thermostat action, the simulated bodies behave diffusively for time scales longer than $N_{\mathrm{Newt}}\delta t/p_v$.

Appendix B

Measurement of writhe

To quantify the topological properties of minicircle systems in chapter 6, the quantities twist (Tw) and writhe (Wr) were determined. In order to calculate these values from microscopic configurations, the discretised formulae introduced in Sec. 2.1.2.2 were used. In most cases, this discretisation routine yields reasonable values that fulfill the Fuller-White-Călugăreanu relation with a deviation of less than 1%, as shown in Sec. 6.2.2. However, in very rare cases, single, unphysical outliers for the value of Wr occur, as can be seen from the time series shown in Fig. B.1.

The outlier values were traced back to stem from individual terms contributing very high values to the discretised formula for Wr (Eq. 2.11). In all cases, these terms occurred in denatured regions, where the double strand centre line $\mathbf{r}(s) = [\mathbf{r}_1(s) + \mathbf{r}_2(s)]/2$ is very flexible and can form tightly wound loops of very high curvature. At high curvature, the discrete approximation of the Wr double integral might be expected to break down, as the tangent vectors $\hat{\mathbf{t}}$ change very quickly, leading to high approximation errors. Indeed, most previous applications of Eq. 2.11 were in systems where no denatured regions occur. Ultimately, this difficulty stems from the ambiguity of assigning a centre line to a denatured double strand region which can temporarily assume very tightly coiled configurations.

This issue might be fixed in a number of ways, either by using a different interpolation algorithm to determine the centre line, smoothing, or a better discrete approximation scheme for Eq. 2.5. Here, a pragmatic approach was chosen, and configurations discarded from averages in which the absolute value of a single term

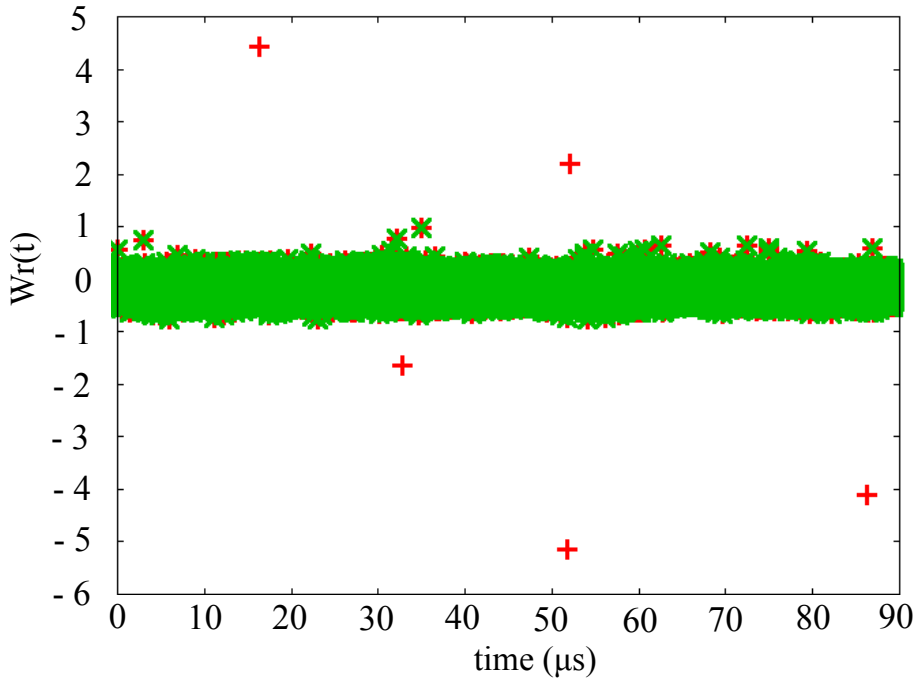


Figure B.1: Times series of Wr as measured for the minicircle carrying the FUSE102 sequence. Red crosses show the calculated values without the cutoff, and green symbols the values with cutoff imposed, as described in the text. Introducing the cutoff on individual terms reliably removes outliers.

in Eq. 2.11 exceeded 1. This removes all the outliers from the time traces (see Fig. B.1), and affected only less than 0.05% of all cases. It was checked that neither means nor standard deviation values shown in Table 6.1 were significantly affected by this procedure.

Appendix C

Sequences

C.1 Random sequences

In this work, sequence-dependent properties of plectonemes were studied using the sequence-dependent parametrisation of oxDNA [173]. The random sequences used in the study of sequence-dependent behaviour are listed here:

Random sequence:

5'-AGAGTACTTAGGCTTGACGATTTCGCGCCTGAACTTCTGATAACTCAGTCTGAGAGACTAAGTTGA
CGTTCTATCCATCATCAGGTGGGCTCAGAGATTGTGCGGCAGACTTAAGTGTAGTACCAGCTGCTGGTC
AATTTGATCTATGCTGATCCGCTCGGAACGGGCCGTGAAAGAAGTACTCTCGCCTATAGAACGGTTAGT
GCTACGACTTTTGCGCGACACAATGTGGTAGTTATCTTCTGTTTTCCTGAATAGTGAGCCTACCAGAAG
AGGCCACCGACAAATCTGATGAGATAGACGGGAACACGGTTTGCGGAGCCTCTGAAACGCTTGTTTATG
AGCAAGAGAGGTGCGGTGGGTATGACCGCCGTAGAAGTACCGTATTCTTCCGGGCTCGGTGGCAATGAA
CACTTAAGGGGCCGACACATTCTGAAGTCAATCGATGGACGGACCTCAACCGTGCACCCTTCTATATAC
GTGTGGCTAGGATACTCTAGCGTTTACCCGCCGTCTTCCACGATGCCGAATATAAGCCGAGGATAAAGG
TGCAGACAAATATCAGGCTTCGCAGTTGTGTAACTTCCTGTATTGTTGTGC-3'

This sequence was chosen at random, with a 25% probability for each possible base identity.

Block-random sequence:

5'-AGACTCGACCGACCCGCGAGATCGGCTCCAGTCTCTGCGCCAAGTGCCGTTGCCCCGTCTCGTGGG
CCGGTCGGTGAACCTTCATACGGTGGGAGTCGCTGAGGCCGTCCCATTTGTACCCATCGAACTCTTATT
TTGTATTTTTTGGACATCCTCAGCTAACCACACGAGCCAAGCTATAGATCAGATTTGGGTATTCGGCGA
TCTTTCTAATCAACTGTATCCGATGCTATACAGATACTTTATTCTAAGGCGGTCCGCGATGCGCCCAGT
CCGTTGACCGGGCGAGTCATGTCAGAGTCGGCAATTATCGGGCACGTCGCCGGGGTGATACGTCCCTGT
GTCACTAGGCATAGGTCGTAACATATGATTATATATACTTTCCACTTTATGTATATCATTTGCAAGTTA
GACATAATAAGGATATATAATATAAGAATCTCTTCACCTCTAAAGTGAGTGATTGGAATATAAGTATT
TGCGCCACTACCCGGCCGAAAGCCCGCGGCTCCTCGCGGGTAGGTTGCCGGGGACCCGCGTGAAGAAAG
GATGAAGCACCCGGACGCCCGCCTGCGAGTTGGCCACGGGCCCATAACGGCG-3'

Above, stretches of 120 bp length are colored, with blue regions having a high GC content and red regions a high AT content. In an abbreviated form, the block-random sequence can be written as 5'-S$_1$$W_1$$S_2$$W_2$$S_3$-3'. The overall GC content of the block-random sequence is 52%. The GC contents of the individual 120 bp stretches is given in Table C.1.

stretch	GC content	AT content
S_1	65.0%	35.0%
W_1	39.2%	60.8%
S_2	60.0%	40.0%
W_2	24.2%	75.8%
S_3	71.7%	28.3%

Table C.1: GC-contents of 120 bp stretches in the block-random sequence

C.2 Biological sequence: p53 promoter region

A nonrandom biological sequence from the promoter region of the p53 gene was used to observe sequence-dependent localisation behaviour of plectonemes. The sequence was retrieved from the HAGR database [232], where it is sourced from Ref. [233]. 5 CG base pairs were added at each end of the sequence, representing the particules

on which the boundary traps act. The full sequence used is:

5'-GCGCGCTCTTACTTGCTACCCAGCACTGATATAGGCACTCAGGAATACAACAATGAATAAGATA
GTAGAAAAATTCTATATCCTCATAAGGCTTACGTTTCCATGTACTGAAAGCAATGAACAAATAAATCTT
ATCAGAGTGATAAGGGTTGTGAAGGAGATTAAATAAGATGGTGTGATATAAAGTATCTGGGAGAAAACG
TTAGGGTGTGATATTACGGAAAGCCTTCCTAAAAAATGACATTTTAACTGATGAGAAGAAAGGATCCAG
CTGAGAGCAAACGCAAAAGCTTTCTTCCTTCCACCCTTCATATTTGACACAATGCAGGATTCCTCCAAA
ATGATTTCCACCAATTCTGCCCTCACAGCTCTGGCTTGCAGAATTTTCCACCCCAAAATGTTAGTATCT
ACGGCACCAGGTCGGCGAGAATCCTGACTCTGCACCCTCCTCCCCAACTCCATTTCCTTTGCTTCCTCC
GGCAGGCGGATTACTTGCCCTTACTTGTCATGGCGACTGTCCAGCTTTGTGCCAGGAGCCTCGCAGGGG
TTGATGGGATTGGGGTTTTCCCCTCCCATGTGCTCAAGACTGGCGCTAAAAGTTTTGAGCGCG-3'

C.3 Minicircle sequences

In the simulations of DNA minicircles discussed in chapter 6 three different basic sequences were used. The identical sequence was used in atomistic simulations with the AMBER package and for denaturation calculations in the model by Benham and co-workers.

An artificially designed sequence with long AT- and GC-rich stretches was employed to study the generic denaturation behaviour. The full designed sequence is:

5'-CGCGCGCGTATATATATATACGCGCGCGCGCGCATATATATATATCGCGCGCGCGCGCAAAAAA
AAAAAACGCGCGCGCGCGCCACACACACACACGCGCGC-3'

The second sequence studied is the random sequence used by Du *et al.*[217] in an experimental study of minicircle formation. The full random sequence is:

5'-ATCTTGCGGCAGTTAATCGAACAAGACCCGTGCAATGCTATCGACATCAAGGCCTATCGTATTA
CGGGGTTGGGAGTCAATGGGTTCAGGATGCAGGTGAGG-3'

Finally, a sequence was studied in the minicircle system in which 50bp of the random sequence given above were replaced by the FUSE sequence found upstream from the human MYC gene. The full FUSE-containing sequence is:

5'-ATCTATATTTAATATATAATGTATATTCCCTCGGGATTTTTTATTTTGTGTTATTTATCGCTAT
TACGGGGTTGGGAGTCAATGGGTTCAGGATGCAGGTGA-3'

Systems of a length 106bp and 108bp were also studied to ensure the correct level of twist in the corresponding AMBER simulations. To attain this system size, 4 and 6 base pair were added to the above sequences respectively, in agreement with the setup used in the corresponding atomistic simulations.
The explicit sequences are given below.

DESIGN106:

5'-CGCGCGCTATATATATATACGCGCGCGCGCGCGCATATATATATATCGCGCGCGCGCGCGCAAA
AAAAAAAAACGCGCGCGCGCGCGCCACACACACACACGCGCG-3'

FUSE106:

5'-ATCTATATTTAATATATAATGTATATTCCCTCGGGATTTTTTATTTTGTGTTATTTATCGCTAT
TACGGGGTTGGGAGTCAATGGGTTCAGGATGCAGGTGAGGAT-3'

RANDOM106:

5'-ATCTTTGCGGCAGTTAATCGAACAAGACCCGTGCAATGCTATCGACATCAAGGCCTATCGCTAT
TACGGGGTTGGGAGTCAATGGGTTCAGGATGCAGGTGAGGAT-3'

DESIGN108:

5’-CGCGCGCTATATATATATACGCGCGCGCGCGCGCATATATATATATCGCGCGCGCGCGCGCAAA
AAAAAAAAACGCGCGCGCGCGCGCCACACACACACACGCGCGCG-3’

FUSE108:

5’-ATCTATATTTAATATATAATGTATATTCCCTCGGGATTTTTTATTTTGTGTTATTTATCGCTAT
TACGGGGTTGGGAGTCAATGGGTTCAGGATGCAGGTGAGGATAT-3’

RANDOM108:

5’-ATCTTTGCGGCAGTTAATCGAACAAGACCCGTGCAATGCTATCGACATCAAGGCCTATCGCTAT
TACGGGGTTGGGAGTCAATGGGTTCAGGATGCAGGTGAGGATAT-3’

C.4 Denaturation distributions for other circle lengths

The denaturation distributions for sequences of length 106bp and 108bp are given in Figs. C.1 to C.6, showing similar behaviour as the 102bp system (see Figs. 6.2-6.4)

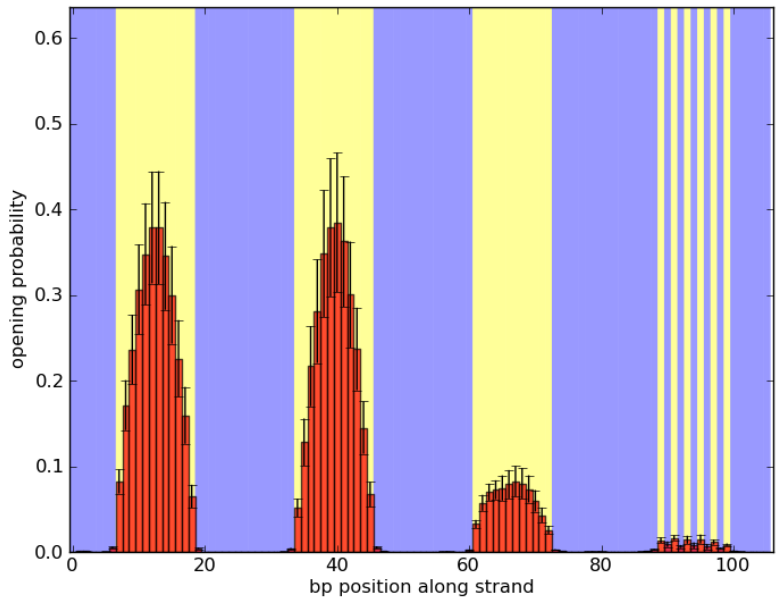


Figure C.1: Denaturation distributions of minicircles of length 106bp carrying the DESIGN sequence, analogous to Fig. 6.2 of the main text.

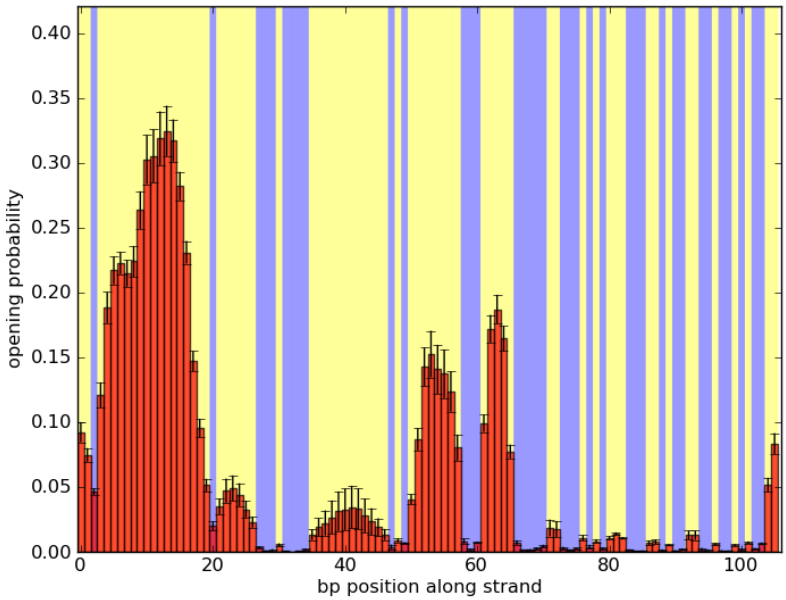


Figure C.2: Denaturation distributions of minicircles of length 106bp carrying the FUSE sequence, analogous to Fig. 6.3 of the main text.

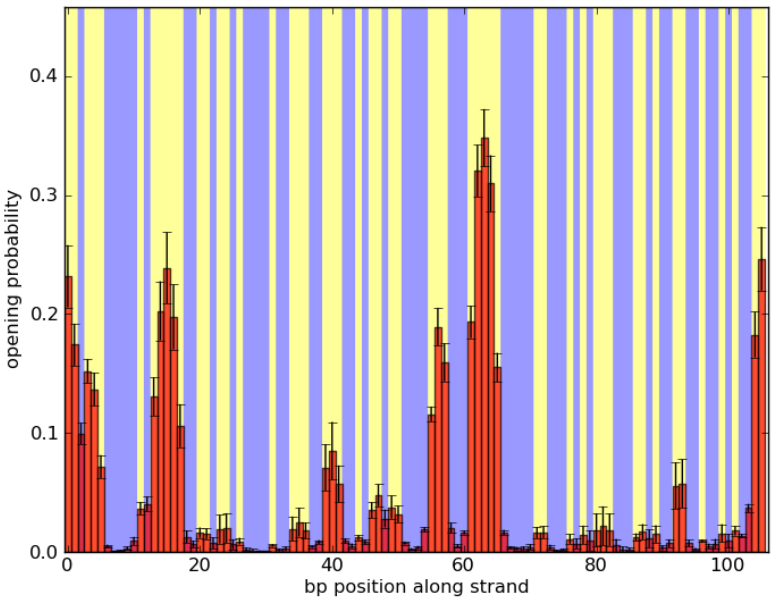


Figure C.3: Denaturation distributions of minicircles of length 106bp carrying the RANDOM sequence, analogous to Fig. 6.4 of the main text.

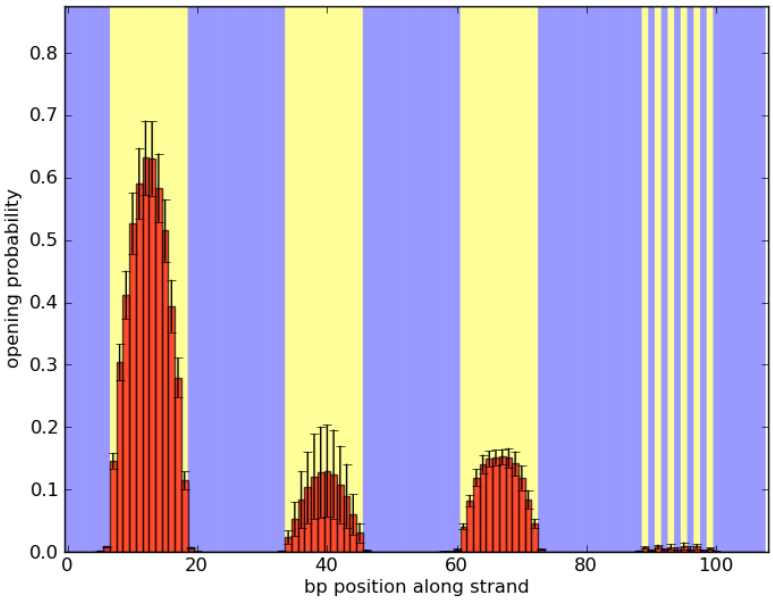


Figure C.4: Denaturation distributions of minicircles of length 108bp carrying the DESIGN sequence, analogous to Fig. 6.2 of the main text.

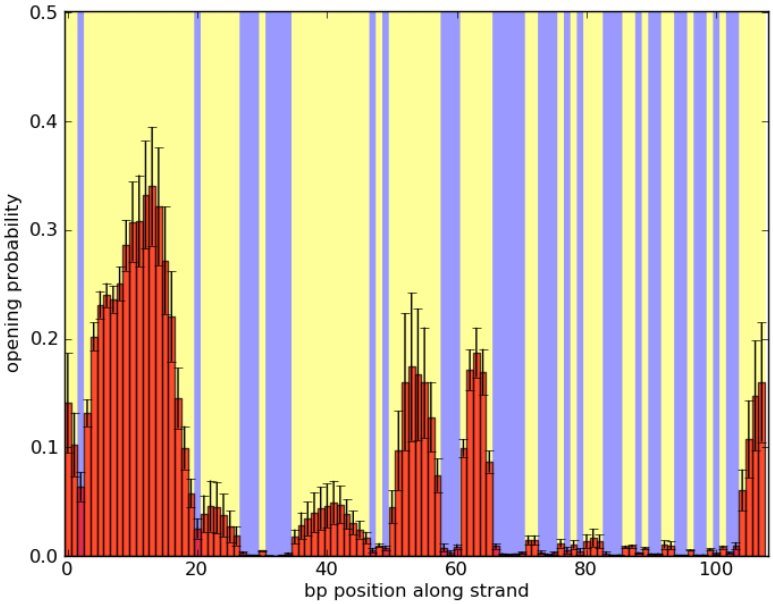


Figure C.5: Denaturation distributions of minicircles of length 108bp carrying the FUSE sequence, analogous to Fig. 6.3 of the main text.

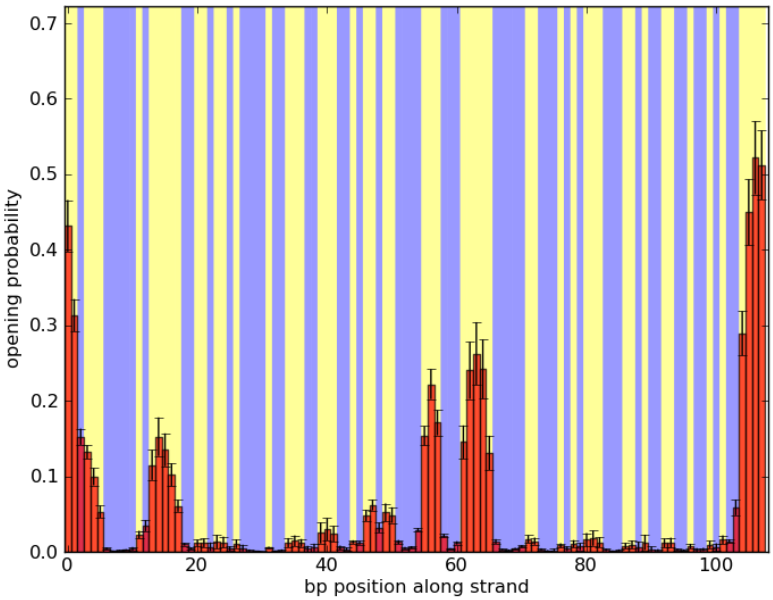


Figure C.6: Denaturation distributions of minicircles of length 108bp carrying the RANDOM sequence, analogous to Fig. 6.4 of the main text.