NORTHWESTERN UNIVERSITY

DNA Topology and Mechanics

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Physics & Astronomy

By

Sumitabha Brahmachari

EVANSTON, ILLINOIS
September, 2018

Abstract

The overall goal of my thesis is to enhance our quantitative understanding of the biophysical properties of DNA – a long polynucleotide chain, present in every living cell, that embodies the genetic information. The existence of DNA has been known to us for over a century, however, our understanding of its physiochemical nature and spatial organization inside our cells is ever evolving. Since the momentous discovery of the double-helical structure of DNA, the emergent view of DNA as a long biopolymer that proteins manipulate via physical interactions has been very successful in explaining experimental observations and proposing biological mechanisms. Since DNA resides in an environment where thermal fluctuations are omnipresent, statistical-mechanical properties of DNA play an indispensable role in protein-DNA interactions. Topological constraints are an essential feature of cellular DNA. Active control of both the intra-DNA topology, arising from the double-helix structure, and the inter-DNA topology due to its long length and self-avoidance, is an important characteristic of various biological functions. This dissertation contains theoretical models of double-helix DNA and other biologically-relevant DNA structures, such as two intertwined DNAs, where we view the double helix as a semi-flexible elastic rod or a worm-like chain with an inherent twist stiffness. Our results explain torsion-induced buckling in stretched double-helix DNAs and intertwined DNAs that are in good quantitative agreement with existing experiments. New experimental data, resulting from collaborations, that successfully verified theoretical predictions are also reported. Some of our novel findings shed light on the role of certain structural defects in modulating DNA-buckling behavior, and the influence structural bulkiness may have on the stability of buckled DNAs. We also address a long-standing question of topological simplification of cellular chromosomes via modeling chromosome as a polymer bottle-brush or a cylindrical array of DNA loops. Our finding, inter-chromosome entanglements can be minimized by an optimal-loop length, may suggest that chromosome domains in interphase nuclei or "Topologically Associating Domains" (TADs) play a role in entanglement minimization. Loop extrusion, that has been recently proposed as a mechanism to compact chromosomes during the cell cycle, provides an active process to

control the compaction state of chromosomes within our model. Our model of loop-extruded chromosome is quantitatively consistent with experiments reporting the rigidity modulus in chromosomes, which in our case, derives from a cylindrical core of densely packed DNA. The theoretical models described here make testable predictions, that we hope, will help design future experiments and uncover new phenomena, creating a need for novel theoretical techniques.

Acknowledgment

My dissertation sums up a body of work that would not have been accomplished if not for the active participation of many individuals, to whom I want to extend my heartiest thanks. The crucial contribution of these individuals in making this work possible demands the use of "we" as the preferred personal pronoun, that I gleefully conform to. First and foremost, I intend to acknowledge my adviser, Prof. John Marko, who has been a constant source of motivation and knowledge. *The first step to solving a physics problem is to draw a figure* – was one of his many good advices that were instrumental in shaping the work presented here. I am thankful to my collaborators, Prof. Alfonso Mondragón, Dr. Kathryn Gunn, Dr. Andrew Dittmore, and Prof. Keir Neuman, who devised the experimental component of the work discussed here. I am also grateful to the members of the Marko Lab, Dr. Edward Banigan, Dr. Andrew Stephens, and Mr. Ronald Biggs for many interesting discussions and inputs. Thanks to the members of my thesis committee, Prof. Monica Olvera de la Cruz and Prof. Michelle Driscoll, for their helpful comments. I want to acknowledge the Molecular Biosciences Training Program at Northwestern University for funding during the early years of my PhD. Last but not by any means the least, my parents, without whose selfless moral support none of this would have been possible.

Brief Contents

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Complexity in the biological world has diverse manifestations: it can be in form of intricate intracellular processes occurring at the nanometer (nm) length scale, or collective behavior of many cells at upwards of many micrometers (μm). Here, our focus will be on the complex intracellular organization and the associated statistical mechanics of nucleic acids or the genome: a polynucleotide chain, where each nucleotide monomer is constituted of a phosphate group, a sugar-ring or ribose, and one of the four nitrogenous bases: adenine (A), guanine (G), cytosine (C), and thymine (T). Nucleic acids can be DNA (deoxyribonucleic acid), a polymeric chain of paired bases (A=T or G≡C)[1] that stores genetic information as a well-defined sequence of complementary bases. DNA is transcribed or read by a synchrony of proteins to make another commonly found nucleic acid, RNA (ribonucleic acid), a polymeric sequence of single bases: A or C or G or U (uracil)[2]. RNA chains are codes that prompt synthesis of specific proteins in the cell. Proteins are macromolecular structures made up of amino acids. The RNA code is translated by ribosomes to make a protein with a specific sequence of amino acids. Thus, DNA, that stores the genetic code of an organism, is read and then translated to synthesize proteins according to the DNA code.

The concepts of modern molecular biology were largely established following the discovery

[1] Owing to structural constraints, adenine only pairs with thymine using two hydrogen bonds, while guanine pairs with cytosine using three hydrogen bonds.

[2] RNA base uracil usually takes the place of DNA base thymine that has a methyl group lacking in uracil.

of the double-helical structure of DNA in 1953 by J. Watson and F. Crick [1]. Nonetheless, it is worthwhile to indulge in a short historical discourse. In 1880, W. Flemming provided a detailed study of cell division, where he was intrigued by the "thread figures" appearing inside the nucleus [2] [see Fig. 1.4, chromosomes appear as cylindrical objects in mitosis]. He also coined the term "chromatin", which led to chromosomes, the nuclear figures that are passed from the mother to daughter cells. We now know chromatin as a polymer of protein-bound DNA segments that is compacted to form chromosomes, however, Flemming's observation and characterization of chromatin using staining dyes are relevant till date.

> "Therefore, we will designate as *chromatin that substance*, in the *nucleus, which upon treatment with dyes known as nuclear stains does absorb the dye*. From my description of the results of staining resting and dividing cells it follows that the chromatin is distributed *throughout the whole resting nucleus*, mostly in the nucleoli, the network and the membrane, but also in the ground-substance. In *nuclear division* it accumulates exclusively in the *thread figures*."
>
> — W. Flemming (1880) [2]

DNA was extracted as the nuclear substance or "nuclein" by F. Miescher in 1869, who characterized its high phosphate content [3]. In 1889, R. Altmann termed the phosphate-rich nuclear substance as "nucleic acid", which was later studied by A. Kossel and coworkers, in early 1900s, to identify the five constituting nucleotides: A, T, G, C, and U.

In 1886, G. Mendel, unaware of Flemming's work, developed a quantitative framework to analyze hereditary inheritance in hybrid plants [4]. He proposed that traits or "characters" are passed on from the parent to progeny via "inheritable units" that make a certain trait either present or absent in the progeny, as opposed to the notion that the progeny contains a blend of parental traits. Although the mechanistic basis of Mendel's arguments were limited to the abstract concept of inherited units, by the end of nineteenth century, when Mendel's work garnered renewed enthusiasm, the possibility that chromosomes are the nuclear material linked with hereditary inheritance gained prominence. In 1903, W. Sutton, put forward "the

chromosome theory of inheritance", which postulated chromosomes as the bearer of hereditary traits [5].

Around the same time, in early 1900s, W. Johannsen proposed "genes" as the unit in Mendelian inheritance, and highlighted a distinction between the "genotype": "the sum total of all the *genes*", and the "phenotype" of an organism: "*types* of organisms, distinguishable by direct inspection" [6]. The exclusive correspondence of genes to DNA became clear after a series of studies, notably, the Avery-MacLeod-McCarty experiment (1943)[7], and the Hershey-Chase experiment (1952) [8], concluded that nucleic acids and not proteins are the genetic material.

In 1953, Watson and Crick proposed the currently-accepted double-helix model for DNA [1], which proved instrumental in understanding the molecular nature of DNA. Watson and Crick's proposal of the three dimensional double-helical structure of DNA, with hydrogen-bonding interactions between purine (A, G) and pyrimidine bases (C, T) holding the structure together, was consistent with the X-ray crystallographic data published by R. Franklin and M. Wilkins, as well as the base-pairing "rules" proposed by E. Chargaff, which stated that the purine-to-pyrimidine ratio is a constant close to unity [9]. The complementarity of double-helix bases (A can only pair with T, while G can only pair with C) indicated that one strand can act as a template to synthesize the other, and provided a possible copying mechanism for the genetic material [1], called the semi-conservative replication scheme, which is now known to be the predominant mode of DNA replication in cells.

Discovery of the double-helix DNA structure qualified the abstract notion of genes into something with a known chemical and physical identity. This ushered in an era of biophysical research where the cellular DNA or chromosomes are viewed as physical objects that are constantly interacting with various proteins to facilitate vital life processes, such as DNA replication, gene transcription and DNA compaction. Continued efforts by researchers has enhanced our molecular understanding of biological processes, enabling us not only to ask more detailed questions probing the fundamentals of the processes, but also to apply our understanding in controlling certain biological processes, such as treatment of diseases.

In early 1990s, it became possible to experimentally study the mechanical response of short DNA fragments to stretching and twisting perturbations [10]. These precise experimental studies, along with theoretical modeling led to a worm-like chain or a stiff polymer picture of DNA at short length scales ($\approx \mu$m) [11, 12]. Quantification of DNA mechanics at the short length scales provided a tool to quantitatively study single-molecule protein-DNA interactions using DNA as a calibrated substrate [13–16]. These studies also highlighted the intimate connection between DNA mechanics and its topology, derived from the double-helical structure, that led to a better understanding of many biological processes where the interplay between mechanics and topology takes a center stage. For example, the essential role of DNA topoisomerase proteins became clearer: relieving DNA torsional stress resulting from DNA replication, as well as disentangle the highly intertwined replicated DNAs [17]. In this dissertation, we will delve deeper into DNA mechanics, the interplay with topology, and its possible cellular ramifications, with a focus on modeling DNA as a semiflexible polymer.

An enhanced understanding of the microscopic nature of DNA and protein-DNA interactions makes us better equipped to study the large scale organization of the genome and its compaction-decompaction cycle during cell division. Recently, lengthwise compaction via loop extrusion has been proposed as a mechanism to compact the genome [18–22]. Active extrusion of DNA loops by proteins provides a microscopic mechanism to manipulate the large scale genome structure and compaction state to form cylindrical chromosomal structures: the "thread figures", as remarked by Flemming more than a century ago [2]. We will also discuss modeling chromosomes as cylindrical polymer brushes, a consequence of loop extrusion, and the implications of the brush morphology to chromosome entanglements. In the rest of this chapter, we will intimate the reader to structure and topology of DNA, and the electrostatic aspects of the aqueous ionic buffer that our cells live in.

1.1 DNA structure

1.1.A DNA is composed of two polynucleotide chains

The building blocks of DNA are nucleotide monomers. Each nucleotide contains a phosphate group attached to a sugar ring that is connected to a nucleobase: adenine (A), thymine (T), guanine (G), and cytosine (C) [Fig. 1.1(a)]. DNA contains two polynucleotide strands where the sugar and the phosphate groups are arranged on the outside with the nucleobases projected inwards, such that two nucleobases from the opposite strands are facing each other, allowing them to form hydrogen bonds.

The nucleobases are flat heterocyclic rings containing carbon and nitrogen atoms. The bases fall into two categories: purines (adenine and guanine), a double ring structure, and pyrimidines (cytosine and thymine) that have a single ring structure [Fig. 1.1(a)].

1.1.B The two strands of DNA have complementary sequences

Structural constraints of the nucleotides only allow a purine-pyrimidine base pair between the two strands, i.e., the allowed base pairs are A=T, and G≡C. This complementarity of base pairs implies one strand can be used as a template to make a copy of the DNA or replicate the DNA for cell division.

1.1.C DNA has a double-helical structure

The two polynucleotide strands of the DNA are helically wrapped, where the sugar-phosphate part of each nucleotide sits on the outside of the helix forming the backbone of the double helix. While the bases on either strands project inwards and noncovalently bind with each other. Specificity of base pairing is important for hydrogen bonding of the bases of the two strands. DNA has two unequal grooves: the major and minor grooves, which is a direct consequence of the double-helical structure.

The most common form of cellular DNA is B-DNA, which is a right-handed double helix. We will always refer to the B-form of DNA unless otherwise mentioned. A consequence of the

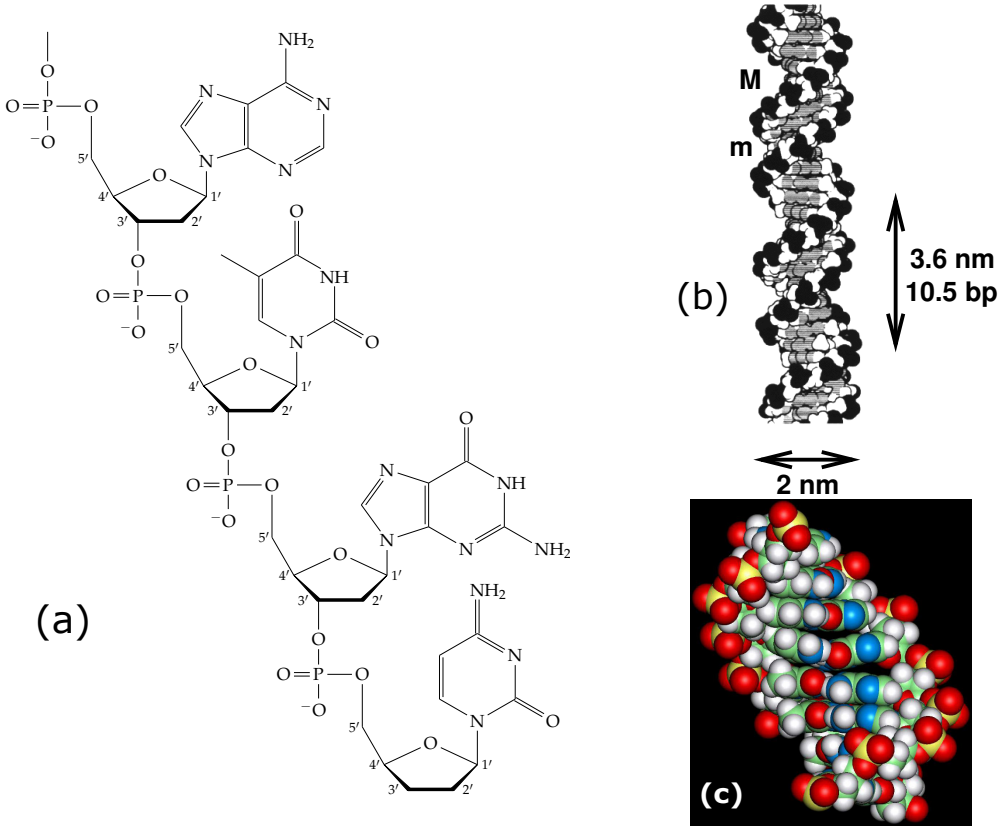

Figure 1.1: (a) Chemical structure of one DNA chain, showing the deoxyribose sugars (note numbered carbons) and charged phosphates along the backbone, and the attached bases (A, T, G and C following the 5′ to 3′ direction from top to bottom). (b) Space-filling diagram of the double helix. Two complementary-sequence strands as in (a) noncovalently bind together via base-pairing and stacking interactions, and coil around one another to form a regular helix. The two strands can be seen to have directed chemical structures, and are oppositely directed. Note the different sizes of the major (M) and minor (m) grooves, and the negatively charged phosphates along the backbones (dark groups). The helix repeat is 3.6 nm, and the DNA cross-sectional diameter is 2 nm. Image reproduced from Ref. [23]. (c) Molecular-dynamics snapshot suggestive of a typical double-helix DNA conformation for a short 10 bp molecule in solution at room temperature. Reproduced from Ref. [24].

helical nature of DNA is its periodicity. Each base pair is rotated from its neighbor by $\sim 35°$, meaning, the helix contour makes one full rotation about its axis every ≈ 10.5 bp, which is the helix-repeat length of DNA.

1.1.D The double helix is stabilized by base pairing and base stacking

The hydrogen bonds between the complementary bases is essential for the double helix, however, noncovalent bonds alone are not enough to explain the thermodynamic stability of double-helix DNA. In absence of complementary base pairing, the nucleobases form hydrogen bonds with the water molecules,which also stabilizes the bases. An important contribution to stability of the double helix is from the stacking interactions between adjacent bases. The bases are flat, allowing them to stack on top of each other, which is favored because of van der Waals attraction between the stacked bases.

1.1.E Conformations of the double helix

DNA can adopt configurations other than the canonical B-from: a right-handed helix with a 10.5 bp helix repeat length. In low humidity conditions or in many protein-DNA complexes, DNA adopts the A-form, which is also a right-handed helix but with a wider major groove and 11 bp per turn. DNA is also known to conform to a left-handed double-helical structure, called Z-DNA, an intermediate structure that occurs during gene transcription.

1.2 Mechanical properties of DNA

The stacked nature of the bases makes the double-helix a stiff polymer, allowing only a few degrees of lateral bending per base pair. One degree of lateral bend corresponds to roughly 0.03 nm of separation between the adjacent bases. However, one may expect to see occasional large bends arising from correlated distortions over many base pairs. In this section, we develop a quantitative understanding of how the double helix responds to mechanical perturbations in a thermal environment.

1.2.A DNA is a stiff polymer

DNA is a semiflexible polymer with a bending stiffness similar to an elastic rod. The thermal persistence length of DNA is $A \approx 50$ nm, or $\approx$150 bp. This means that, first, DNA polymers

longer than a few hundred nanometers will be severely distorted from a straight line configuration due to thermal fluctuations; and second, the configurations of the fluctuating DNA will feature correlated bends over ≈ 50 nm segments.

1.2.B Statistical mechanics of DNA

The fact that thermal fluctuations will bend DNA indicates that the end-to-end extension of a long piece of DNA is typically smaller than its contour length. The ensemble-averaged end-to-end distance of semiflexible DNA, computed from a partition function [25], gives the following:

$$\langle R_{ee}^2 \rangle = 2AL + 2A^2 \left(e^{-L/A} - 1 \right) \tag{1.1}$$

where L is the total contour length. The limit $L \approx A$, furnishes $R_{ee} \approx L$, the rod-like limit of DNA. While, $L \gg A$ limit gives the Gaussian polymer behavior: $R_{ee} \approx \sqrt{AL}$.

1.2.C Entropic elasticity of stretched DNA

If a stretching force is applied to the ends of a long DNA molecule, the end-to-end extension increases. This is due to suppression of the entropic degrees of freedom perpendicular to the force direction. In other words, work needs to be done in stretching the ends of a polymer.

The characteristic force delineating the strong stretching regime from weak stretching is set by the thermal persistence length. Forces in the range $f < k_B T/A \approx 0.1$ pN (recall 1 $k_B T \approx 4$ pN-nm) are weak perturbations to the equilibrium Gaussian-polymer conformation. While, forces $f \ll k_B T/A$ strongly stretch the DNA polymer.

Polymer elasticity at low forces. In the weak stretching regime $f < k_B T/A$, the polymer has a linear force response, furnishing a spring constant $k = k_B T / \langle R_{ee} \rangle^2 = k_B T/(AL)$. This corresponds to a Gaussian polymer, where the spring constant is inversely proportional to polymer length, and $f \approx k \langle R_{ee} \rangle$, sets the order of magnitude of the restoring force.

As the length of DNA is increased, the self-avoidance of the polymer plays an important role that makes the force response nonlinear [26]. However, for double-helix DNA, the narrow

effective thickness ($\approx$ 3.5 nm at 100 mM univalent salt including electrostatic effects [27]) of the double helix compared to its segment length ($b = 2A \approx 100$ nm) leads to quite weak self-avoidance, and makes dsDNA elasticity quite close to that of an ideal polymer for DNA lengths (< 50 kb ≈ 16 μm) studied experimentally [12].

Strongly stretched DNA exhibits semiflexible-polymer behavior. DNA under high forces ($f \gg k_BT/A$) approaches a fully extended state, where transverse fluctuations are small and can be treated as a perturbation. The total energy has a force-extension energy component that is negative since work has to be done to stretch the molecule, along with the bending energy term that scales with the square of the local curvature, as shown in the following [12].

$$\beta E = \int_0^L ds \left[\frac{A}{2} \left(\frac{d\hat{\mathbf{t}}}{ds} \right)^2 - \beta f \hat{\mathbf{z}} \cdot \hat{\mathbf{t}} \right] \qquad \Rightarrow \frac{\langle z \rangle}{L} = \langle \hat{\mathbf{z}} \cdot \hat{\mathbf{t}} \rangle = 1 - \sqrt{\frac{k_BT}{4Af}} + \cdots \tag{1.2}$$

The first equation shows the total energy whereas the second one shows the ensemble-averaged extension derived from a canonical partition function. The characteristic reciprocal square-root dependence of extension on force for a semiflexible polymer in the regime $f \gg k_BT/A$ is observed in single-molecule experiments on double-helix DNA for forces from about 0.1 up to 10 pN [10].

1.2.D DNA denaturation by stress

From DNA "melting" studies, we know that the energy required to separate the helically stacked single-stranded DNAs (ssDNA) is $g \approx 2.5 k_BT$ per base pair [28]. The secondary structure of DNA, which is held together by weak non-covalent bonds of binding energy $\approx k_BT$, is expected to strongly deform under highly stressed conditions. This has been observed in a few different ways.

Unzipping: Pulling the two strands of the DNA in opposite directions leads to unzipping of the double helix DNA strands. The helical arclength associated with each base pair is $\ell \approx 1$ nm, which is the length released upon unzipping. Hence, the force, at which the required work to procure ℓ length of ssDNA from a double helix equals the base-pairing energy, gives a simple estimate of the unzipping force: $f_{\text{unzip}} \approx g/\ell = 10$ pN. The experimentally observed unzipping

force ranges from 8 to 15 pN, depending on DNA sequence [29–31]. The variations in unzipping force has been proposed to be used to analyze DNA sequence.

Overstretching: Under a large applied force a long dsDNA undergoes a structural transition, where the double helix length per base pair increases from 0.34 nm to 0.6 nm. Again using DNA strand separation energy as the free energy scale, we estimate the overstretching force: $f_{\text{overstretch}} \approx 2.5 k_{\text{B}}T/(0.2 \text{ nm}) \approx 50$ pN. Experimentally observed overstretching transition occurs at a well-defined force 65 pN [32–34].

Unwinding: One might imagine an applied torque with a negative helicity (double helix DNA has positive helicity) will unwrap the two single strands of the DNA. Unwinding the DNA releases ≈ 0.6 rad/bp (2π radians per 10.5 bp), which along with the base-pairing energy of $2.5 k_B T$/bp gives an estimate of the critical unwinding torque: $\tau_{\text{unwind}} \approx -2.5 k_{\text{B}}T/(0.6 \text{ rad}) \approx -16$ pN·nm (the sign reflects helicity or handedness). The experimentally observed unwinding torque is ≈ -10 pN·nm (a slightly lower torque than the above estimate occurs since there is left-handed wrapping resulting after denaturation).

Experimental observations and more detailed theoretical work has resulted in development of a force-torque "phase-diagram" for the double helix, with a variety of structural states [35, 36].

1.3 Topology of DNA

The two helically wrapped strands of a DNA are *linked*, i.e., for a circular DNA the two strands cannot be separated or unlinked from one another without breaking one of them. This gives rise to an internal linking number for the double helix, which is closely connected to its twist response.

Topology of a polymer refers to linking or entanglement of the polymer. Topology is invariant under smooth geometric deformations, and only changes when one polymer passes through another. A simple example is the linking of two rings; they can be linked or unlinked, and one cannot pass from the linked to the unlinked state without breaking one of the rings.

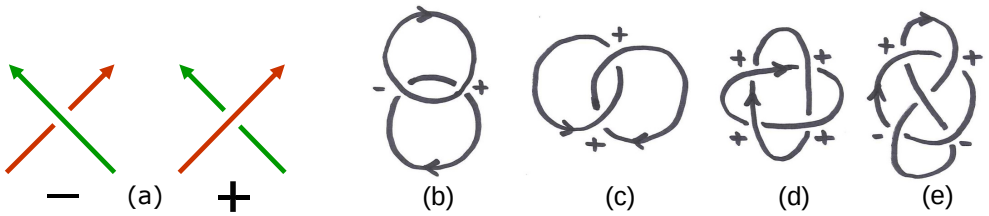

Figure 1.2: (a)Sign convention for computation of linking number using crossings. Left: left-handed (-1) crossing. Right: right-handed ($+1$) crossing. Simple links of oriented loops. Lk for each pair is computed by adding up the signs of the crossings and dividing the sum by 2. (b) unlinked rings; the signs of the crossings cancel, so Lk $= 0$. (c) the Hopf link; the signs of the crossings add, so Lk $= +1$ (Lk would be -1 if the orientation of one of the loops were reversed). (d) for this link (sometimes called "Solomon's knot") the signs of the crossings again add, making Lk $= +2$. (e) the Whitehead link has canceling signs of its crossings, and has Lk $= 0$ despite being a nontrivial link. Image reproduced from Ref. [37].

1.3.A Linking number

The *linking number* of two oriented closed curves can be computed by counting their signed crossings, according to the rules shown in Fig. 1.2. Dividing the total crossing number by two gives an integer, the linking number Lk of the two curves. This quantity can only change when one curve is passed through another. [3]

The Gauss invariant computes the same quantity, but determines it from the geometry of the two curves:

$$\mathrm{Lk} = \frac{1}{4\pi} \oint_{C_1} \oint_{C_2} \frac{d\mathbf{r}_1 \times d\mathbf{r}_2 \cdot (\mathbf{r}_1 - \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|^3} \tag{1.3}$$

For DNA, we can distinguish between *external* linking of two double helix molecules together, and the *internal* linking property of the double helix itself.

[3] Linking topology is perfectly well defined only for closed curves or polymers. However, it is sometimes useful to define linkage of open curves, using suitably defined closure boundary conditions, e.g., closing chains at infinity by extending them with long straight paths. This introduces small corrections to the properties of entanglement of interest here (primarily estimates of linking number). Qualitatively this can be understood by considering the definition of linking number in terms of signed crossings (Fig. 1.2). If we imagine deforming part of one of the links of Fig. 1.2 so that it closes far from the other crossings (not introducing any new crossings in the process) the topology and linking number of the polymer will be unchanged. This will be true for all closure paths that do not introduce additional strand crossings, indicating a rather weak dependence of linking number on closure boundary conditions, and further allowing us to talk about the topology of the region of the polymers *not including* the closure in a reasonably well-defined way.

1.3.A.i Internal double-helix linking number Lk

The two strands of a double-helix DNA are wrapped around each other in a right-handed manner, with a preferred helix-repeat of one turn every $n_h \approx 10.5$ bp, or every $h \approx 3.6$ nm. This causes *linking* of the two strands, resulting in a net linking number associated with the double helix structure: $\mathrm{Lk} \approx \mathrm{Lk}_0 = L/h = N/n_h$, for a double helix of length L or N base pairs. However, Lk is an integer for a closed double helix, and is not in general equal to Lk_0.

The difference between double helix linking number and the preferred linking number, $\Delta\mathrm{Lk} = \mathrm{Lk} - \mathrm{Lk}_0$, is often expressed as a fraction of the preferred linking number (linking number density), $\sigma \equiv \Delta\mathrm{Lk}/\mathrm{Lk}_0$ (the excess linking number per DNA length is $\Delta\mathrm{Lk}/L = \sigma/h$). In *E. coli* and many other species of bacteria, circular DNA molecules are maintained in a state of appreciably perturbed Lk, with $\sigma \approx -0.05$. This is a sufficient perturbation to drive the DNA to *supercoil*, or wrap around itself in the manner of a twisted extension cord, due to competition between bending and twisting elasticity of the double helix.

1.3.A.ii Perturbation of internal linking number generates DNA torsion

If Lk is sufficiently different from Lk_0, then there will be a buildup of twist in the DNA, such that the DNA twist energy can be expressed as,

$$\beta E_{\mathrm{twist}} = \frac{C}{2L}\Theta^2 \tag{1.4}$$

where Θ is the net twist angle along the double helix. This is just the form of the twisting energy for a uniform elastic rod. Experimentally, this simple linear model has been observed to have a surprisingly wide range of validity for DNA, for $C \approx 100$ nm [35].

In the absence of other constraints, thermal fluctuations of twist give rise to a fluctuation: $\langle\Theta^2\rangle = \frac{L}{C}$, suggesting the interpretation of C as a characteristic length for twist fluctuations. For the double helix, this *twist persistence length* is $C \approx 100$ nm.

The derivative of E_{twist} with respect to Θ is the torque or "torsional stress" in the DNA that

increases linearly with the applied twist:

$$\tau = \frac{\partial E_{\text{twist}}}{\partial \Theta} = \frac{k_B T C}{L}\Theta \tag{1.5}$$

If there is no bending, then any excess linking number ΔLk goes entirely into twisting the double helix: $\Theta = 2\pi\Delta\text{Lk}$ (or $\sigma = \Theta/[2\pi L/h]$). The mechanical torque in DNA will be $\tau = 2\pi k_B T C \Delta\text{Lk}/L = (2\pi k_B T C/h)\sigma$. The parameter $2\pi C/h \approx 175$ sets the scale for when the linking number density will start to appreciably perturb DNA conformation, i.e., when $|\tau| \approx k_B T$. This level of torque occurs for $|\sigma| \approx 0.005$.

1.3.A.iii Decomposition of double helix Lk into twist Tw and writhe Wr

The previous computation supposed that there was no bending, in which case all of the ΔLk is put into twisting the double helix. This DNA twisting can be quantified through the twist angle Θ, or equivalently through the twisting number [4].

If DNA bending occurs, there may be *nonlocal* crossings of the double helix over itself. These nonlocal crossings contribute to double-helix linking number, and the separation of length scales between DNA thickness and the longer scale of DNA self crossing (controlled by the persistence length A) allows linking number to be decomposed into local (twist) and nonlocal (writhe) crossing contributions:

$$\text{Lk} = \text{Tw} + \text{Wr} \tag{1.6}$$

or equivalently, $\Delta\text{Lk} = \Delta\text{Tw} + \text{Wr}$. This is known as White's Theorem. While Lk is a topological property and is quantized for a covalently closed double helix, Wr and Tw are geometrical, and change value smoothly as the molecule is distorted.

One can demonstrate this with a thin strip of paper (30 cm by 1 cm works well). Put one twist into the strip, closing it in a ring. The two edges of the strip are linked together once. Now without opening the ring, let it assume a figure-8 shape; you will see that you can make

[4] The total twist of a DNA molecule is often written as the excess twist ΔTw plus the intrinsic twist, or $\text{Tw} = \Delta\text{Tw} + \text{Lk}_0 = \Delta\text{Tw} + L/h$, where $\Delta\text{Tw} = \Theta/(2\pi)$.

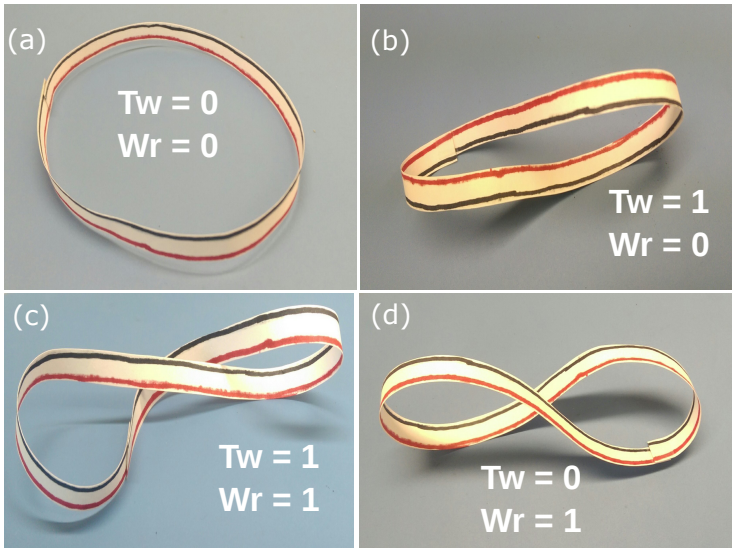

Figure 1.3: Conformations of a ribbon (or a piece of paper cut out in the shape of a ribbon as shown in the images) with different linking of the two edges, see the red and black colored lines showing the edges. (a) Untwisted ribbon. (b) Twisted ribbon with a unit twist, which can be achieved by rotating one end of the ribbon by 360° before joining the two ends. (c) Ribbon with two links between the edges, where the shown conformation has, roughly, one unit of twist and one unit of writhe. (d) The linking number between the edges is unity, the same as in (b). However, note the altered conformation where the entire linking number is stored in form of writhe, following White's theorem [Eq. (1.6)].

the twist go away: in this state there is only writhe (Fig. 1.3).

DNA supercoiling. The torque developed in the DNA due to twisting of the double helix can lead to a mechanical response in form of chiral bending. This response is often a wrapping of the double helix around itself, a phenomenon known as *supercoiling*. One can observe this by taking a stiff cord and twisting it. This behavior arises from a competition between the bending energy and the elastic twist energy. A bent configuration gives rise to nonlocal crossings and has a writhe contribution. Hence, when the twist energy is very high the total energy can be lowered by bending, because the writhe contribution of the bent structure lowers the torque in the double helix. However, note that bent structure should have a well defined chiral to contribute maximum writhe, and consequently provide maximum torsional stability. A commonly found bent configuration that the double helix succumbs to under torsional stress is that of superhelix

or a plectoneme, where the double helix wraps around itself in a helical fashion. We will indulge in a detailed discussion of plectonemes and the corresponding buckling transition in Chapter 3.

1.3.B Catenation or inter-DNA linking number

DNA due to long length often gets intertwined that leads to topological entanglements. This is inter-DNA topology, which we call entanglement or catenation, and is distinct from the linking topology of the two strands of the double helix. Torsionally relaxed DNA, e.g., DNA with one base missing along its contour or nicked DNA, cannot be twisted, however they can get entangled due to inter-DNA topology constraints. Here, we will discuss entanglement or catenation of torsionally relaxed DNA. Interplay of topology between double-helix linking number and inter-DNA catenation number is an interesting and poorly understood topic [38], and will be left for future work.

1.3.B.i Knotting of DNA

DNA can get knotted,which is typically detrimental to the cell, because these knots, unless resolved, can lead to problems in gene transcription and DNA replication. The probability that a polymer will be knotted will, in general, increase with its contour length, i.e., a longer polymer is expected to be more knotted. The characteristic length scale above which knotting is probable has been shown to be quite large (about 300-500 segment lengths, which for DNA $\approx$ 100 kb) [19]. This estimate is based on numerical simulations of ideal polymers, and self-avoiding interaction makes the characteristic knotting length much longer ($\approx 10^6$ segment lengths). The remarkably low knotting probability of DNA, being based on numerical studies, lacks fundamental understanding, however, suggests that knotting of cellular DNA is rare.

1.3.B.ii DNA topoisomerase

The topology of cellular DNA, both intra- and intermolecular linking, is actively manipulated by proteins called DNA topoisomerases (Topo). These proteins, Type-I Topoisomerase (TopoI),

can alter the DNA linking number by over/under-winding the double helix. TopoI mediates a transient cut in one of the strands of the double-helix backbone and relaxes the topology by swiveling the cut strand about the intact one. DNA in most bacteria, e.g., *E. coli*, are maintained in a twisted state by the action of TopoI.

Type-II DNA topoisomerases (TopoII) drives inter-DNA topology manipulation. TopoII mediates a cleavage of both the DNA strands allowing another double-helix segment to pass through the cut, before resealing the severed backbone. Thus, TopoII allows changes in the inter-DNA topology or DNA catenation. TopoII is thought to perform selective decatenation in order to suppress the equilibrium probabilities of knotted DNA states, which is consistent with the fact that topo II mediated decatenation requires ATP hydrolysis (the requirement of ATP seems to ensure that the second molecule is passed through the gap in a specific direction). However, the mechanism underlying active suppression of entanglements via selective decatenation is not fully understood [39–41].

TopoII drives topology fluctuations. TopoII are locally-acting enzymes that are unaware of the global topology of long DNA polymers in the cell. Thus inter-DNA topology changes by TopoII in the cell are uncorrelated events and can be thought of as a mechanism to maintain a fluctuating topology, and as we shall see, these topology fluctuations are crucial in compacting and segregating the genome.

Topoisomerases are present in all living cells, bacteria and eukarya, and are essential for a living cell. Inhibition of topoisomerases lead to stalling of the cell cycle, and eventually cell death, which makes topoisomerase inhibition an important biomedical tool to stop cell proliferation. This has led to the use to topoisomerase inhibitors, like etoposide, as drugs against cancerous cells and infectious bacteria [42].

1.4 Electrostatics in the cellular environment

Cells live in an aqueous solution of salts, i.e., an ionic aqueous buffer, where the most abundant are the monovalent ions, like Na^+ and K^+. The typical cellular concentration of monovalent

ions is ≈ 0.1 M[5]. DNA has a negatively charged backbone, due to the negative charges on the phosphate group, which adds to $2e^-$ per base pair[6], i.e., a linear charge density of $\approx 6e^-$/nm of the double helix. This charge on the backbone, however, is strongly screened by the local arrangement of counter ions around the backbone. The free ions in the solution form a charge layer encapsulating the double helix that screens the charges on the backbone, as a result, the electrostatic potential decays much more rapidly than the inverse-distance behavior expected from Coulomb's law.

1.4.A Debye-Hückel theory

The charge distribution in an ionic solution follows Maxwell-Boltzmann distribution, where the configurations with higher electrostatic energy, due to close proximity of similar charges, are less probable. This leads to a Poisson-Boltzmann equation for the electrostatic potential Ψ (see Appendix A). Assuming that the number of ions per unit volume is low enough that thermal energy dominates the electrostatic energy ($e\Psi \ll k_BT$), we can linearize the Poisson-Boltzmann equation to write the Debye-Hückel equation:

$$\lambda_D^2 \nabla^2 \Psi = \Psi \tag{1.7}$$

where λ_D is the Debye-Hückel screening length, a characteristic length scale in the system over which the electrostatic potential is screened. The screening length is higher for lower molarity of the ionic buffer: $\lambda_D \approx 0.3/\sqrt{M}$ nm, where M is the molar concentration of the solution (see Table 2.1). The solution of the above equation gives an exponentially decaying potential: $\Psi(r) \sim r^{-1} \exp(-r/\lambda_D)$. The electrostatic potential at short distances ($r \approx \lambda_D$) follow Coulomb's law: $\sim r^{-1}$; however, at long distances ($r \gg \lambda_D$) the potential decays exponentially with the characteristic length scale λ_D, due to screening by the counter-ion charge distribution. Thus, effectively, DNA fragments in solution may be considered to be cylindrical units with a radius

[5] One molar concentration (1M) means there is 1 mol per litre, where 1 mol of a substance has 6.023×10^{23} units of the substance, which can be molecules or ions or atoms.

[6] units of electrostatic charge: $1e = 1.6 \times 10^{-19}$ Coulomb

set by the electrostatic screening length of the solution, which is larger for lower salts.

Geometry of the linear charge density associated with DNA backbone modifies the boundary conditions and leads to a different short-distance behavior, while at long distances geometry is irrelevant and the electrostatic potential decays exponentially. See Appendix A, where we calculate the electrostatic potential of various relevant DNA geometries.

1.5 Cellular DNA: a confined polymer solution

DNA in our cells reside in confined spaces; for eukaryotes, the confinement is provided by the nuclear envelope, while bacterial DNA is confined by the cell wall itself. Such confinement may play a variety of roles, ranging from protection of our genome from possible damage by other large organelles, to compartmentalization, enabling an efficient processing of the genetic information, e.g., RNA splicing [7].

Our entire genome, containing all the DNA segments inside the cell (or nuclei, in case of eukaryotic cells), is not always a single long polymer, but maybe divided into a few chains or chromosomes. The number of chromosomes in a cell defines the *karyotype* of the cell. A haploid human genome has 23 chromosomes, whereas, bacteria has only one chromosome, see Table 4.1 for a list of various karyotypes.

Chromosomes in our cell are typically not bare-DNA segments, but are composed of protein-bound DNA. Eukaryotic chromosomes are constituted of nucleosomes [8], a histone-DNA complex, that is roughly 10 nm in diameter. A string of nucleosomes is called a chromatin, where the name derives from the ability of chromatin to absorb certain dyes, as remarked by W. Flemming [2]. While, bacterial chromosomes do not have nucleosomes, their DNA is coated with various

[7] DNA is transcribed to make pre-mRNA (precursor to messenger RNA), the RNA molecule may undergo partial deletion and editing to mature the molecule into an mRNA ready to be translated by Ribosomes to synthesize proteins. Splicing of nuclear genes is carried out by a specialized macromolecular machinery called Spliceosomes. There are clusters enriched with spliceosomes and RNA inside the nucleus, called splicing speckles, where most of the RNA splicing occurs. Nuclear confinement provides an enclosed compartment for splicing of nuclear genes, following which mRNA is transported outside the nucleus where Ribosomes translate the message encoded in the RNA molecule.

[8] A nucleosome contains ≈ 200 bp of DNA, where an octamer of histone proteins tightly wrap ≈ 150 bp around themselves to form a histone core. Two 20 bp DNA fragments are "linker" DNA that is not wrapped around the core, but as the name suggests, acts as a connection between adjacent nucleosomes.

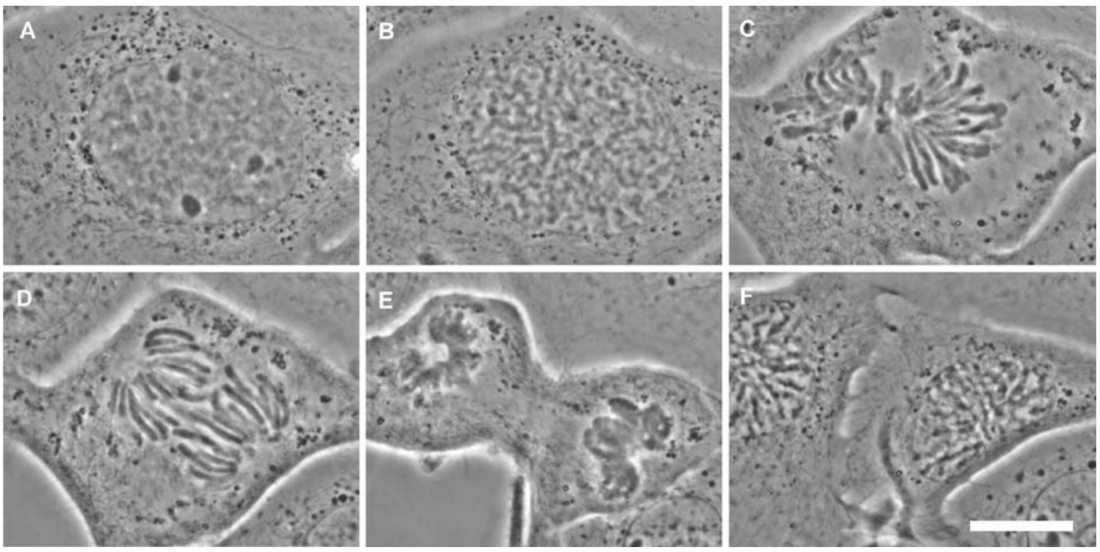

Figure 1.4: Animal cell division. (A) Late interphase showing dark nucleoli and still uncondensed chromosomes; (B) prophase showing long but condensed prophase chromosomes; (C) spindle-aligned and shortened metaphase chromosomes; (D) separation of chromatids at anaphase; (E) telophase chromosomes beginning to decondense; (F) interphase nuclei in daughter cells. Bar is 20 μm (Image reproduced from Ref. [43]).

DNA-bending proteins, like IHF and HU.

1.5.A Chromosomes are self-avoiding polymers

We view chromatin as a polymer constituted of spherical nucleosome monomers of diameter $a \approx 10$ nm [9]. In a physiological solution (0.1 M Na^+), chromatin is a self-avoiding polymer, i.e., the volume occupied by a nucleosome is inaccessible to the others. Genome size or the total number of nucleosomes N, corresponds to a contour length $L = Na$ such that the end-to-end distance (radius of gyration) of the genome scales positively with the number of nucleosomes via the 3D-Flory exponent [26, 44]:

$$R_F(N) = a^{2/5}(wa^3)^{1/5}N^{3/5} \tag{1.8}$$

[9] Our ideas are applicable to bacterial DNA also, as we shall see in Chapter 4. Bacteria do not have nucleosomes and the microscopic monomers constituting the chromosomes are cylinders of height given by the DNA persistence length $a \approx 50$ nm and width $b \approx 4$ nm corresponding to a protein-bound DNA fragment.

where the excluded volume per monomer is wa^3. *Good solvent* ($w > 0$) refers to the condition where the free energy associated with monomer-monomer contacts is higher than that of monomer-solvent contacts. When the monomers have no self-adhesion, we have $w = 1$, the polymer expands due to self-avoidance. Weak self-adhesive interactions can lead to $0 < w < 1$, where the end-to-end extension is lower due to less self-avoidance among the monomers. Strong monomer self-adhesion may overcome their self-avoidance and lead to a collapse of the polymer, which corresponds to *poor solvents* ($w < 0$). Eq. (1.8), that is applicable only for self-avoiding polymers, breaks down in the poor solvent limit ($w < 0$). However, the effect of a deteriorating solvent quality may be analyzed using Eq. (1.8) by lowering the excluded volume parameter w between 0 and 1. We shall see that a globally-operative self-adhesive interaction that facilitates non-specific sticking among nucleosomes ($w < 1$), increases inter-chromosomal entanglements [45].

1.5.B Volume fraction of confined chromosomes

The volume fraction of chromosomes in confinement, i.e., the fraction of the total confinement volume occupied by all the genomic segments, is typically high enough to promote overlap between different chromosomes ($\phi \approx 0.01 - 0.1$, see Table 4.1) [26]. The high degree of confinement is important for gene regulation, as physical proximity of genomic segments is crucial to turn on/off expression of certain genes (e.g., via promoter-enhancer contacts). However, strong confinement also has an undesirable aspect of a higher degree of entanglement among different chromosomes. As we shall see in Chapter 4, the cell has devised a way to disentangle its chromosomes, namely, via chromosome compaction.

1.5.C Closely packed blobs of chromatin

Self-avoiding chromatin polymer may be viewed as a sequence of "de Gennes' blobs", where the size of the blob ξ, also called the correlation length, is the length scale associated with screening of polymer correlations [26]. In a dilute polymer solution, where chains are isolated,

the correlation length is given by the Flory radius [Eq. (1.8)] and the entire chain is one blob. At a higher volume fraction, different chains overlap and the correlation length is smaller than the whole chain's radius of gyration, due to stronger screening of polymer interactions at higher concentrations [26]. The volume fraction of chromatin inside the nucleus ϕ (≈ 0.07 for humans, see Table 4.1), is typically much higher than the critical volume fraction above which chains overlap $\phi^\star \sim N^{-4/5}$ ($\approx 10^{-6}$ for humans, Table 4.1). A higher volume fraction also means that nuclear chromatin may be considered as a closely packed system of mutually excluding blobs or a *melt of blobs*, where the polymer correlations inside a blob is that of a self-avoiding polymer; and the polymer considered as a string of blobs obeys Gaussian statistics, akin to an ideal polymer [26, 44].

1.5.D Cell cycle dependent (de)compaction of chromosomes

The mitotic cell cycle is composed of interphase and mitosis. During interphase the cell is transcribing genes to make proteins, replicating its DNA for cell division. Interphase is followed by mitosis, where the chromosomes undergo compaction. During mitosis, chromosomes emerge as individualized rod like structures, composed to densely compacted chromatin. The compact chromosomes then undergo segregation, followed by their equal division into the two daughter cells. Following mitosis, the chromosomes in daughter cells undergo decompaction in the interphase of the daughter cells.

While the decompacted state of the genome is essential to allow physical access to DNA for transcription and replication, the compacted state during mitosis helps drive segregation of the chromosomes. We will delve deeper into disentanglement and segregation of mitotic chromosomes in Chapter 4.

1.6 Summary

DNA is long polymer made up of di-nucleotide monomers or base-pairs that exhibits semi-flexibility at short length scales (≈ 100 nm). DNA stiffness originates from the double-helical

stacking of its base pairs. DNA has an intrinsic topology due to helical linking of the two sugar-phosphate strands bearing the nucleotide monomers, and perturbation of topology is inherently related to mechanics. Mechanical stress can lead to twisting, stretching and denaturation of the double-helix, which shape the free energy landscape of protein-DNA interactions in the cell. Chromosomes, the bearer of hereditary units, are DNA coated with proteins. Cellular chromosomes are under strong confinement forcing inter-chromosome overlap and entanglement. Chromosomes undergo cell-cycle synchronized compaction and disentanglement (or segregation), a process where the global structure and topology of chromosomes are actively controlled by locally-acting proteins.

In Chap. 2, we model the mechanics of two intertwined DNA double helices under applied stretching that elucidates its buckling behavior shaped by the torque from DNA wrapping. The findings of this chapter are published in Refs. [46, 47]. Chap. 3 contains the mechanics of stretched-twisted DNA, with applications to the case of DNA harboring an inhomogeneity, such as a protein-mediated kink. Our model and results for supercoiled DNA and the consequences of a defect on DNA mechanics, as described in Chap. 3 can be found in Refs. [48, 49]. We then study, in Chap. 4, the global compaction and segregation of chromosomes in confinement by the active process of DNA loop extrusion.

Chapter 2

Intertwined double-helix DNAs

This chapter concerns a statistical-mechanical model for the behavior of two intertwined double-helix DNAs or DNA braids, where individual DNAs are torsionally unconstrained or *nicked*[1]. We focus on the torque and extension as a function of braid catenation (linking) number and applied force, as studied in magnetic tweezers experiments.

Braids have two double-helix DNA molecules wrapped around each other in a helical fashion, where the wrapping generates a restoring torque. This restoring torque, that tends to unwrap the braids, increases nonlinearly with the catenation number of the braid. In other words, DNA braids exhibit a catenation-dependent effective twist modulus, distinct from what is observed for twisted individual double-helix DNAs (we discuss individual twisted double helices in Chapter 3). Braid torque drives buckling of the stretched braid, where buckling occurs near the point where experiments have observed a change of slope in the extension versus linking number curves.

The buckled structure is characterized by a "supercoiled" or helically wrapped configuration of the two braids. Writhing of the braid in the buckled state suppresses torque increase and stabilizes the buckled structure. However, the bulkiness of plectonemically supercoiled braid contributes to destabilizing the buckled state. Structural bulkiness derives from the fact that

[1] Nicked DNA implies at least one of the DNA base pairs has a missing base, indicating that DNA torque, originating from double-helix twisting, is relaxed

braid diameter is at least twice the double-helix diameter, this restricts the minimum braid-plectoneme diameter to four times that of the double-helix diameter. The bending energy of a plectoneme structure scales positively with plectoneme radii, resulting in a less stable braid plectoneme.

Destabilization of the bulk braid-plectoneme state leads to a proliferation of multiple buckled domains following the nucleation of the first domain. A buckled domain is defined as a braid-plectoneme with a braid end loop that is a finite-sized structure connecting the self-writhing braid. Nucleation of the braid end loop associates an extension discontinuity at the buckling transition, that we predicted and subsequently followed up experimentally to verify its existence. We also experimentally find coexistence of multiple buckled structures in the braid-plectoneme state, as predicted. The findings of this chapter are published [46, 47].

Significance Catenated or intertwined DNA molecules are a common occurrence in the cell as they are an intermediate in segregation of sister chromatids. Following DNA replication and recombination, the sister chromatids are in an intertwined or braided state due to the remnant linking of their parental double helix [50–53]. Catenated DNA molecules can be mimicked *in vitro* by wrapping or "braiding" two single DNA molecules around each other. At the single-molecule level, DNA braids are important substrates to study the topology-manipulation mechanism of DNA topoisomerases and site-specific DNA recombinases [14–16, 53].

Braided DNAs have also been studied in precise single-molecule manipulation experiments using magnetic tweezers[2]. Experiments of this type have been used to study removal of DNA catenations by Type-II DNA topoisomerases (TopoII) [13–15, 55, 57]. Braided DNAs have also been used to study the decatenation activity of type-I topoisomerases [16, 52], as well as the double-helix segment-exchange activity of site-specific DNA recombinases [53, 58, 59].

[2] Magnetic tweezers are used to apply a constant force to a particle to which two double-helix DNAs are attached; the opposite ends of the DNAs are tethered to a surface [13, 54–56] [Figs. 2.1(a) and 2.8(a)]. The resulting double tether can have a controlled force applied to it using a magnetic field gradient, while at the same time, the total linking number of the two double-helix molecules can be adjusted by rotating the magnetic field so as to rotate the magnetic particle. As a result, one can study the extension of the two DNAs as a function of inter-DNA linking number or catenation number (Ca).

Understanding these kinds of DNA-topology-changing enzyme experiments depend on the understanding of the physical properties of the DNA braids, but this has lagged behind our understanding of the simpler problem of a single twisted stretched double helix [11, 12, 60–66]. The reason for this is that braided DNAs are a more complex physical situation than a single supercoiled DNA; as a result, while there have been prior theoretical studies of helically intertwined DNAs [45, 56, 67–70], those works have not quantitatively analyzed the buckling ("braid supercoiling") behavior. While it has been assumed that the experimentally observed change in the slope of braid extension versus catenation number corresponds to the onset of braid supercoiling [13, 55, 56], the precise location and nature of braid buckling have not been theoretically understood. Two factors that make the problem of braided DNAs distinct from the mechanics of a single twisted DNA under tension are first, the lack of an intrinsic braid twist elastic modulus, and second, the dependence of the braid mechanics on the distance between the tethering points of the two double helices [13, 55, 56].

Outline The layout of the chapter is as follows. Sec. 2.1 contains a detailed description of the mathematical model, where we study the braid Hamiltonian in the thermodynamic limit (Sec. 2.1.A). Free energies corresponding to a tethered braid and thermal averaging of fluctuations are discussed in Sec. 2.1.B and Sec. 2.1.C, respectively. The results and predictions are contained in Sec. 2.2, where we study braids at physiological salt (Sec. 2.2.A), as well as the effect of varying salt concentration (Sec. 2.2.B) and other finite-size effects (Sec. **??** and 2.2.D).

2.1 Model

We build a free energy model for braids considering double-helix DNAs as electrically charged semi-flexible polymers residing in an ionic solution. Fig. 2.1(a) shows how we view a braided DNA structure. The ends of two *nicked* DNA molecules are tethered to a fixed wall and a rotating bead respectively, such that the intertether distance on either end is d. This scenario is similar to the setup for tweezer experiments [13, 54–56]. The beads used in experiments are

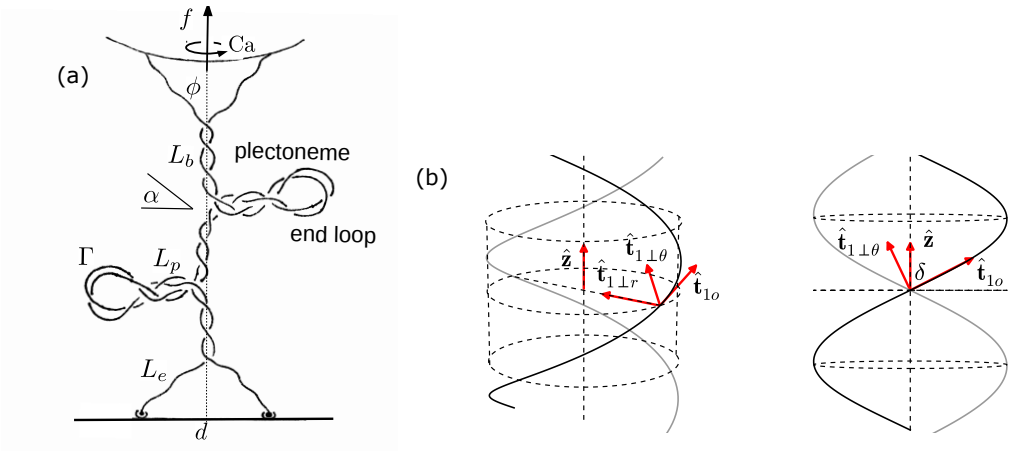

Figure 2.1: (a) Schematic of a DNA braid under torsional stress, showing coexistence of straight and plectonemically buckled states. The individual dsDNAs are able to swivel around their contact with the wall, keeping the dsDNAs from twisting. (b) Two duplex DNAs (dark and gray shaded) helically braided in a right-handed manner on the surface of a cylinder of radius R oriented parallel to the $\hat{\mathbf{z}}$-axis, viewed from two angles. The orthonormal triad $(\hat{\mathbf{t}}_{io}, \hat{\mathbf{t}}_{i\perp r}, \hat{\mathbf{t}}_{i\perp\theta})$, where $i \in \{1, 2\}$ is shown for one of the curves. $\hat{\mathbf{t}}_{1o}$ is in the direction of the tangent to the helix, $\hat{\mathbf{t}}_{1\perp r}$ is oriented radially inward, and $\hat{\mathbf{t}}_{1\perp\theta} \equiv \hat{\mathbf{t}}_{1o} \times \hat{\mathbf{t}}_{1\perp r}$. The projection of the triad on the $\hat{\mathbf{z}}$ axis is a constant dependent on the helix parameters [Eq. (2.3)].

large enough to safely assume no leakage of catenation number via looping of the DNA over the beads. By applying a constant force to the rotationally constrained bead it is possible to study DNA braids in a fixed force and fixed catenation ensemble.

2.1.A The Hamiltonian

We express the Hamiltonian $\mathcal{H}$ associated with two nicked double-helix DNA molecules of length L, held at a fixed inter-DNA linking or catenation number and under a constant applied force $f\hat{\mathbf{z}}$ as follows.

$$\frac{\mathcal{H}}{k_B T} = \int_0^{L/A} d\xi \left[\frac{1}{2} \left(\left| \frac{d\hat{\mathbf{t}}_1}{d\xi} \right|^2 + \left| \frac{d\hat{\mathbf{t}}_2}{d\xi} \right|^2 \right) - \frac{fA}{k_B T} \hat{\mathbf{z}} \cdot \frac{1}{2} (\hat{\mathbf{t}}_1 + \hat{\mathbf{t}}_2) + U(\mathbf{r}_1, \mathbf{r}_2) \right] \tag{2.1}$$

where A is the DNA persistence length and ξ is the dimensionless arc length. $\mathbf{r}_i(\xi)$ and $\mathbf{t}_i(\xi) \equiv (1/A)(\partial \mathbf{r}_i / d\xi)$, are respectively the position vector and the tangent of the i-th braiding strand

for $i \in \{1, 2\}$. The first term in the integrand containing the sum of the squares of the local curvature of the two strands correspond to elastic bending energy of the two double helices. The second term containing the external force $f\hat{\mathbf{z}}$ corresponds to the entropic elasticity of the two chains. The electrostatic part of the Hamiltonian is represented by $U(\mathbf{r}_1, \mathbf{r}_2)$. Since we only consider nicked double-helix DNAs, there are no DNA-twist-energy terms in the above Hamiltonian.

Two catenated elastic rods, many persistence lengths long, under a high stretching force form coaxial helices. We take the average shape of the braiding strands to be that of a regular helix oriented parallel to the direction of the external force $f\hat{\mathbf{z}}$ [Fig. 2.1(a)], and propose a perturbative expansion of the braid Hamiltonian [Eq. (2.1)] around a mean-field solution parameterized by radius R, and pitch $2\pi P$ of the helix.

2.1.A.i Oscillating reference frame

We expand the tangent vectors $\hat{\mathbf{t}}_i$ ($i \in \{1, 2\}$) in Eq. (2.1) about a mean-field direction $\mathbf{t}_{io}$ [Fig. 2.1(b)]:

$$\hat{\mathbf{t}}_i = \left[1 - \frac{\mathrm{t}_{i\perp}^2}{2} + \mathcal{O}(\mathrm{t}_{i\perp}^4)\right] \hat{\mathbf{t}}_{io} + \mathbf{t}_{i\perp} \tag{2.2}$$

where $\mathbf{t}_{i\perp} = \mathbf{t}_{i\perp r} + \mathbf{t}_{i\perp\theta}$. We introduce two rotating right-handed orthonormal triads [Fig. 2.1(b)]: $(\hat{\mathbf{t}}_{io}, \hat{\mathbf{t}}_{i\perp r}, \hat{\mathbf{t}}_{i\perp\theta})$, where $i \in \{1, 2\}$, such that the unit vector $\hat{\mathbf{t}}_{io}$ points along the tangent to the mean-field helix corresponding to the i-th strand, $\hat{\mathbf{t}}_{i\perp r}$ points along the radially-inward direction, and $\hat{\mathbf{t}}_{i\perp\theta} \equiv \hat{\mathbf{t}}_{io} \times \hat{\mathbf{t}}_{i\perp r}$. Note that the $\hat{\mathbf{z}}$-projection of the basis vectors depend only on the helix parameters:

$$\hat{\mathbf{z}} \cdot \hat{\mathbf{t}}_{io} = \cos\delta \; ; \qquad \hat{\mathbf{z}} \cdot \hat{\mathbf{t}}_{i\perp\theta} = \sin\delta \; ; \qquad \hat{\mathbf{z}} \cdot \hat{\mathbf{t}}_{i\perp r} = 0 \; ; \tag{2.3}$$

where $\delta \equiv \arctan(R/P)$, is the braiding angle.

The derivatives of the orthonormal basis with respect to normalized arc length ξ are given

by the following equations:

$$\frac{d}{d\xi}\begin{bmatrix}\hat{\mathbf{t}}_{io}\\ \hat{\mathbf{t}}_{i\perp r}\\ \hat{\mathbf{t}}_{i\perp\theta}\end{bmatrix} = \begin{bmatrix}0 & \kappa & 0\\ -\kappa & 0 & \kappa a\\ 0 & -\kappa a & 0\end{bmatrix}\begin{bmatrix}\hat{\mathbf{t}}_{io}\\ \hat{\mathbf{t}}_{i\perp r}\\ \hat{\mathbf{t}}_{i\perp\theta}\end{bmatrix} \tag{2.4}$$

where $\kappa \equiv AR/(R^2+P^2)$, is the total mean-field curvature per unit persistence length of the strands and $a \equiv P/R$. The above set of equations [Eq. (2.4)] are also known as the Frenet-Serret formulas. We neglect any space-varying component of the mean-field helix curvature [68, 71–73], which is a good approximation in the thermodynamic limit of long braids.

Using the above equations, the square of the local curvature is written as

$$\left|\frac{d\hat{\mathbf{t}}_i}{d\xi}\right|^2 = (1-\mathrm{t}_{i\perp}^2)\kappa^2 + \left|\frac{d\mathbf{t}_{i\perp}}{d\xi}\right|^2 + 2\kappa\hat{\mathbf{t}}_{i\perp r}\cdot\frac{d\mathbf{t}_{i\perp}}{d\xi} \tag{2.5}$$

where we have neglected $\mathcal{O}(\mathrm{t}_{i\perp}^3)$ terms.

2.1.A.ii Electrostatic interaction

Physical micromanipulation experiments on DNA have been performed in varied concentrations of aqueous buffers, whereas ≈ 100 mM Na^+ or K^+ is the physiologically relevant range of salt. Counterion condensation on the negatively charged DNA backbone ($2e^-$ per base pair) results in a screened Coulomb potential over a characteristic length scale called the Debye screening length, $\lambda_D \approx 0.3\ \mathrm{nm}/\sqrt{\mathrm{M}}$ for M molar univalent salt. DNA-DNA repulsion over a few screening lengths is that of the Debye-Hückel type, *i.e.*, the electrostatic potential decays exponentially at large distances and diverges like the Coulomb potential at distances shorter than λ_D.

The electrostatic potential due to a close proximity of two parallel DNAs has been shown to be described by Debye-Hückel interaction of uniformly charged rods [74, 75]. For helically wrapped DNA chains, also the case for plectonemes in single supercoiled DNA an empirical modification of the parallel rod potential has been shown to account for the enhancement due to helical bends in the structure [76]. Furthermore, there is an electrostatic contribution from

the self interaction of the braiding helices, for which we propose an empirical form that agrees with the numerical solution of the Debye-Hückel-type self interaction (Appendix A.3). The total electrostatic potential energy per unit length A of the braid, in k_BT units is given by

$$\mathcal{U}(R,P) = \zeta K_0\left(\frac{2R}{\lambda_D}\right) Z(a) + \frac{\zeta c_1 \lambda_D^2}{R^2 a^2 (1+c_2 a^2)} \tag{2.6}$$

where $a = \cot\delta$ and $Z(a) \equiv 1 + m_1/a^2 + m_2/a^4$.

The first term in Eq. (2.6) is the electrostatic interaction potential between the braiding strands and Z is the correction factor for helix curvature with $m_1 = 0.828$ and $m_2 = 0.864$ [76]. The second term corresponds to the self-electrostatic energy of a helix, where $c_1 = 0.042$, and $c_2 = 0.312$, are chosen to closely match the numerical solution [Fig. A.2]. $\zeta \equiv 2A\ell_B\nu^2$, is the amplitude of the Debye-Hückel potential, where $\ell_B = e^2/(\epsilon k_B T)$, is the Bjerrum length of the solution with dielectric constant ϵ and ν is the effective linear charge density of the double-helix DNA, which is a parameter used to satisfy the near-to-surface boundary conditions for the far-field Debye-Hückel solution [12, 63, 74–77]. We use $\ell_B = 0.7$ nm corresponding to water at 290 K; numerical values of the effective charge ν and the Debye screening length λ_D, used for various salt concentrations are given in Table 2.1. The bending persistence length of DNA is also known to slightly modify on changing the salt concentration [78, 79], but we neglect such small changes as they are inconsequential to our qualitative results.

We approximate the total braid electrostatic potential as the average potential arising from self and mutual repulsion of two coaxial helices, and consider radial fluctuations in the braid in the asymptotic limit of parallel chains. We consider small uniform deviations in the braid radius $A\mathbf{w}(\xi)$ such that,

$$\mathbf{w}(\xi) = \int_0^\xi \frac{1}{2}[\mathbf{t}_{1\perp r} - \mathbf{t}_{2\perp r}]d\bar{\xi} = \hat{\mathbf{t}}_{1\perp r}\int_0^\xi \frac{1}{2}[\mathrm{t}_{1\perp r} + \mathrm{t}_{2\perp r}]d\bar{\xi} \tag{2.7}$$

where $\mathrm{t}_{i\perp r}$ are given by Eq. (2.2) and we assume the boundary condition $\mathbf{w}(0) = 0$. The above definition of normalized radial deformations $\mathbf{w}(\xi)$ assumes a parallel configuration of the two

strands. We define the electrostatic part of the Hamiltonian as:

$$U(\mathbf{r_1},\mathbf{r_2}) \equiv \mathcal{U}(R + A\mathrm{w}(\xi), P) = \mathcal{U}_0 + g\mathrm{w} + \eta\mathrm{w}^2 + \mathcal{O}(\mathrm{w}^3) \tag{2.8}$$

where $\mathcal{U}_0 \equiv \mathcal{U}(R,P)$, is given by Eq. (2.6). $g \equiv A\partial\mathcal{U}/\partial R$, and $\eta \equiv (A^2/2)(\partial^2\mathcal{U}/\partial R^2)$, where η is the effective electrostatic modulus of uniform radial deformations in the braid. The first term gives the mean electrostatic energy per unit length A of braid with fixed radius and pitch, while the subsequent terms are corrections for small uniform derivation in braid radius. We neglect expansion of the electrostatic potential in the pitch of the braid because the fluctuations in the pitch are predominantly controlled by the external tension.

2.1.A.iii Perturbative expansion of the Hamiltonian

We expand the total Hamiltonian [Eq. (2.1)] to the quadratic order in transverse-tangent fluctuations:

$$\begin{aligned}\beta\mathcal{H} =& \frac{L}{A}\left[\kappa^2 - 2\mu + \mathcal{U}_0\right] + \int_0^{L/A} d\xi \left[\kappa\left\{\hat{\mathbf{t}}_{1\perp r}\cdot\frac{d\mathbf{t}_{1\perp}}{d\xi} + \hat{\mathbf{t}}_{2\perp r}\cdot\frac{d\mathbf{t}_{2\perp}}{d\xi}\right\} - (\mu/a)(\mathrm{t}_{1\perp\theta} + \mathrm{t}_{2\perp\theta}) + g\mathrm{w}\right] \\ &+ \frac{1}{2}\int_0^{L/A} d\xi \left[\left|\frac{d\mathbf{t}_{1\perp}}{d\xi}\right|^2 + \left|\frac{d\mathbf{t}_{2\perp}}{d\xi}\right|^2 + (\mu-\kappa^2)\left(|\mathbf{t}_{1\perp}|^2 + |\mathbf{t}_{2\perp}|^2\right) + 2\eta\mathrm{w}^2\right] + \mathcal{O}(\mathrm{t}_{1\perp}^3, \mathrm{t}_{2\perp}^3)\end{aligned} \tag{2.9}$$

where $\mu \equiv (\beta A f\cos\delta)/2$, is the dimensionless effective tension in each strand of the braid. The first term, associated with $\hat{\mathbf{t}}_{io}$ ($i \in \{1,2\}$) component of the tangent vectors is the leading order term that gives the total mean-field energy of the braid.

We represent the real-space components of the transverse-tangents as a sum over dimensionless Fourier modes q:

$$\mathrm{t}_{j\perp k}(\xi) = \frac{A}{L}\sum_q e^{iq\xi}\tilde{\mathrm{t}}_{j\perp k}(q) \tag{2.10}$$

where $i = \sqrt{-1}$, $j \in \{1,2\}$ and $k \in \{r,\theta\}$. We set the reference of the fluctuation free energy

by setting the amplitude of zero-momentum transverse fluctuations to zero: $\tilde{\mathrm{t}}_{j\perp k}(0) = 0$, where $j \in \{1, 2\}$ and $k \in \{r, \theta\}$. The contribution from zero momentum is accounted for by the mean-field parameters, and this boundary condition precludes order-mean-field perturbations. Also, subject to the zero-momentum boundary condition, the second term in Eq. (2.9) (terms linear in $\mathrm{t}_{j\perp k}$) vanishes.

The third term in Eq. (2.9), containing quadratic transverse tangents, accounts for the free energy contribution due to Gaussian fluctuations of the two braided strands about their average helical shapes. We write the third term as a sum over the Fourier modes q:

$$\beta\Delta\mathcal{H} = \frac{1}{2}\sum_{q}\omega^{\dagger}\mathcal{M}\omega \tag{2.11}$$

where $\omega(q)$ is a 4×1 column vector and $\mathcal{M}(q)$ is a 4×4 Hermitian matrix:

$$\mathcal{M}\omega = \frac{A}{L}\begin{bmatrix} X_q + \frac{1}{2}\eta q^{-2} & 2iq\kappa a & \frac{1}{2}\eta q^{-2} & 0 \\ -2iq\kappa a & X_q - \kappa^2 & 0 & 0 \\ \frac{1}{2}\eta q^{-2} & 0 & X_q + \frac{1}{2}\eta q^{-2} & 2iq\kappa a \\ 0 & 0 & -2iq\kappa a & X_q - \kappa^2 \end{bmatrix}\begin{bmatrix} \tilde{\mathrm{t}}_{1\perp r} \\ \tilde{\mathrm{t}}_{1\perp\theta} \\ \tilde{\mathrm{t}}_{2\perp r} \\ \tilde{\mathrm{t}}_{2\perp\theta} \end{bmatrix} \tag{2.12}$$

where $X_q \equiv q^2 + \mu + \kappa^2 a^2$. We compute the fluctuation free energy in the limit of zero curvature ($\kappa \to 0$), which simplifies the configuration to that for two fluctuating parallel chains and makes the problem analytically tractable. Also, note that in our scheme to include fluctuations in the electrostatic part of the Hamiltonian [Eq. (2.7)] we have already assumed zero curvature.

2.1.A.iv Partition function

We construct the canonical partition function for two fluctuating parallel strands:

$$\mathscr{Z} = \prod_{q}\int d\tilde{\mathrm{t}}_{1\perp r}\int d\tilde{\mathrm{t}}_{1\perp\theta}\int d\tilde{\mathrm{t}}_{2\perp r}\int d\tilde{\mathrm{t}}_{2\perp\theta}\; e^{-\frac{1}{2}\omega^{\dagger}\mathcal{M}\omega} = \prod_{q}\frac{(2\pi L/A)^2 q}{\sqrt{(q^2+\mu)^3(q^4+\mu q^2+\eta)}} \tag{2.13}$$

and obtain the fluctuation correction to the mean-field free energy from the partition function:

$$-\frac{\ln \mathscr{Z}}{L/A} = \frac{3}{2}\sqrt{\mu} + \eta^{1/4}\cos\left(\frac{1}{2}\tan^{-1}\sqrt{\frac{4\eta}{\mu^2}-1}\right) \tag{2.14}$$

where we drop constants dependent on only the ultraviolet cutoff. The RHS of Eq. (2.14), which is real and positive for all positive values of μ and η gives the fluctuation free energy per unit length A of the braid. There are four degrees of freedom for transverse fluctuations in a stretched braid; three of them ($\hat{\mathbf{t}}_{1\perp\theta}$, $\hat{\mathbf{t}}_{2\perp\theta}$, and one in $\hat{\mathbf{t}}_{1\perp r}$ where there is no relative displacement between the strands) are controlled solely by the external tension, as seen in the first term of Eq. (2.14). The second term accounts for the fluctuations in $\hat{\mathbf{t}}_{1\perp r}$ that correspond to displacement of the two strands relative to one another, and is controlled by both the external tension and the electrostatic forces.

2.1.A.v Radial fluctuations

The average energy corresponding to radial fluctuations for each wavenumber q can be obtained from the fluctuation Hamiltonian [Eq. (2.9)]:

$$\langle|\tilde{\mathrm{t}}_{1\perp r}(q) + \tilde{\mathrm{t}}_{2\perp r}(q)|^2\rangle = \frac{2q^2L}{A(q^4+\mu q^2+\eta)} \tag{2.15}$$

The two-point correlation function associated with radial undulations can be computed as follows,

$$\langle \mathrm{w}(0)\mathrm{w}(\Delta\xi)\rangle = \frac{A}{L}\int_{-\infty}^{\infty}\frac{dq}{2\pi}e^{iq\Delta\xi}\frac{\langle|\tilde{\mathrm{t}}_{1\perp r}(q)+\tilde{\mathrm{t}}_{2\perp r}(q)|^2\rangle}{4q^2} = \int_{-\infty}^{\infty}\frac{dq}{4\pi}\frac{e^{iq\Delta\xi}}{q^4+\mu q^2+\eta} \tag{2.16}$$

where $\Delta\xi$ is the distance between the two points. Note, the two-point correlation of radial fluctuations decays exponentially with the distance between the points: $\langle \mathrm{w}(0)\mathrm{w}(\Delta\xi)\rangle \sim \exp(-k\Delta\xi)$, where $k \sim \mathcal{O}(\sqrt{\mu})$, which is typical of Gaussian fluctuations.

We obtain the radial fluctuations in the braid from the zero-distance behavior of the above

correlation function:

$$\sigma_R^2 = A^2\langle|\mathrm{w}(0)|^2\rangle = \frac{A^2}{8}\eta^{-3/4}\left[\cos\left(\frac{1}{2}\tan^{-1}\sqrt{\frac{4\eta}{\mu^2}-1}\right)\right]^{-1} \tag{2.17}$$

where σ_R is the fluctuation in braid radius. The probability distribution of braid radii is given as follows.

$$P(R) = \frac{1}{\sqrt{2\pi\sigma_R^2}}\exp\left[-\frac{(R-\langle R\rangle)^2}{2\sigma_R^2}\right] \tag{2.18}$$

Fluctuations in the radius of the braid (σ_R) decrease with increasing electrostatic modulus of radial fluctuations (η): $\sigma_R \sim \eta^{-3/8}$ [Eq. (2.17)], which implies a scaling of the fluctuation free energy with the radial fluctuations: $-\ln\mathscr{Z} \sim \sigma_R^{-2/3}$. A similar scaling relation appears for the confinement entropy of a worm-like chain trapped in a rigid cylindrical tube, where the confinement entropy scales with the radius of the tube: $\Delta F \sim \langle R\rangle^{-2/3}$ [80]. Theoretical studies of supercoiled DNA have used the confinement entropy scaling to account for strand undulations in a plectoneme structure [12, 65]. Again in the context of plectonemic DNA, the scaling ansatz was modified: for Gaussian fluctuations of a worm-like chain trapped in a potential well, the average radius $\langle R\rangle$ could be replaced by the radial fluctuation σ_R, which was then chosen to be the Debye length of the solution [66, 76, 81]. Indeed, we find that σ_R of the free energy minimized braid is of the order of the Debye length [Fig. 2.3(b), Table 2.1]. The existing literature on plectonemic and braided DNAs, to the best of our knowledge uses the confinement entropy scaling approach to account for strand undulations [12, 65–70, 76]. Our calculations treat fluctuations systematically and without a scaling ansatz, and produce the previously assumed scaling behavior.

Pitch fluctuations. We assume a spatially-uniform mean-field curvature in the braiding helices, and the introduction of a space-varying component to the helical curvature will be an interesting addition to the model. The effect of the spatially-varying helical pitch has been studied for loaded plies at zero temperature [71–73] as well as for straight DNA braids [68, 69].

Variable pitch solutions for many persistence-lengths long braided helices feature a constant helical angle inside the braid, which is smaller than the end angle that connects the helices with the end regions [68]. We note that we determine the end angle ϕ from free energy minimization, and we indeed find that an end angle larger than the helical angle δ is energetically favored, consistent with the effect observed in Ref. [68].

2.1.B Mean-field theory

In this section, we develop the free energy expressions for a finite-sized tethered braid [Fig. 2.1(a)], where the total length of each of the braiding molecules is partitioned into a force-extended state (straight phase), and a plectonemically buckled state (plectoneme phase). The plectoneme state also consists of a braid "end loop", a teardrop-shaped loop at the end of every plectoneme structure [Fig. 2.1(a)].

2.1.B.i Catenation number

The amount of catenation per helical repeat of the DNA molecules is defined as the catenation density in braids, $\sigma_c \equiv \mathrm{Ca}/\mathrm{Lk}_0$ ($\mathrm{Lk}_0 = L/h$, where L is the contour length of each DNA and $h = 3.6$ nm, is the length of one helical repeat of double-helix DNA). Total catenation (Ca) is divided between the straight phase (Ca_s) and the plectoneme phase (Ca_p),

$$\mathrm{Ca} = \mathrm{Ca}_s + \mathrm{Ca}_p \tag{2.19}$$

which are further redistributed between twist and writhe as dictated by minimization of the total free energy.

2.1.B.ii Straight braid

The length of each double helix in the straight phase L_s is divided into two parts: (1) the helical intermolecular wrappings of length L_b, such that $L_b = 2\pi \mathrm{Ca_s}\sqrt{R_s^2 + P_s^2}$, where R_s and $2\pi P_s$ are respectively the radius and the pitch of the helical interwounds; and (2) the end regions

of length L_e ($L_e = L_s - L_b$), which do not contain any inter-molecular links and connect the helical wrappings to the tethered points (Fig. 2.1). The mean-field energy of the straight braid is obtained using the leading order term in the expansion of the Hamiltonian [the first term in Eq. (2.9)] as follows.

$$\beta\mathcal{E}_s = L_b\left[\frac{AR_s^2}{(R_s^2+P_s^2)^2} - \beta f\cos\delta_s + \frac{1}{A}\mathcal{U}_0(R_s,P_s)\right] - \beta f L_e\cos\phi \tag{2.20}$$

The first term (with the brackets) corresponds to the helical region of the straight phase, which is a sum of free energy contributions from elastic bending, force-extension, and electrostatic repulsion respectively. Here δ_s is the braiding angle ($\tan\delta_s = R_s/P_s$) in the straight phase. The second term contains the force-extension free energy of the end regions, where ϕ is the opening angle at the end of the braid ($\sin\phi = d/L_e$, where d is the intertether distance, see Fig. 2.1). Note that for a given length (L_s) and catenation ($\mathrm{Ca_s}$), the radius (R_s) and the pitch ($2\pi P_s$) of the braid are the only free parameters in the free energy of the straight phase, minimizing which we obtain the equilibrium state.

The fluctuation correction to the mean-field free energy in the straight phase is obtained from Eq. (2.14),

$$\beta\Delta\mathcal{F}_s = \frac{L_b}{A}\left[\frac{3}{2}\sqrt{\mu_s} + \eta_s^{1/4}\cos\left(\frac{1}{2}\tan^{-1}\sqrt{\frac{4\eta_s}{\mu_s^2}-1}\right)\right] + L_e\sqrt{\frac{2\beta f\cos\phi}{A}} \tag{2.21}$$

The first term (with the brackets) corresponds to the fluctuation contribution to the free energy of the helically wrapped section of the straight phase, where $\mu_s = \mu(R_s,P_s)$ and $\eta_s = \eta(R_s,P_s)$ [Eqs. (2.8) and (2.9)]. The second term corresponds to the worm-like-chain fluctuations of the end regions and is obtained by plugging $\mu = (\beta A f\cos\phi)/2$ and $\eta = 0$ in Eq. (2.13).

2.1.B.iii Plectonemically buckled braid

The plectonemic braid is a buckled structure where the braid centerline writhes around itself (Fig. 2.1). Buckling achieves a lower energy state by the release of torque in the braid, due to an

increase in the writhe contribution to the total linking number. Total writhe in the plectoneme formed of superhelical wrappings of the braid with total double helix length L_p, superhelix opening angle α, and superhelical radius $\mathscr{R}_p$ is given by

$$\mathrm{Wr}_p = L_p \cos\delta_p \frac{\sin 2\alpha}{4\pi\mathscr{R}_p} \tag{2.22}$$

where δ_p is the helix angle of the braid in the plectoneme [12, 76, 82].

2.1.B.iv Braided end loop

Every plectoneme domain is accompanied by a finite-sized loop-shaped structure where the braid bends back (Fig. 2.1). The braid end loop presents an energy cost to nucleation of a plectoneme domain; thermodynamically, the situation is similar to plectoneme nucleation in supercoiled single DNAs [64, 65, 83, 84]. The equilibrium size of the end loops Γ is obtained by separately minimizing the elastic energy cost of forming them:

$$\beta\mathcal{E}_\Gamma = 2\frac{\varepsilon A}{\Gamma} + \beta f\Gamma \quad \Rightarrow \Gamma = \sqrt{\frac{2\varepsilon A}{\beta f}} \tag{2.23}$$

where the first term in the above elastic energy equation is the bending contribution, and the second term is the work done in decoupling the plectoneme end loop from the external tension f [64, 65]. We use $\varepsilon = 16$, corresponding to a "teardrop" geometry of the loop [85–87].

2.1.B.v Partition of plectoneme catenation number

Considering the writhe contribution of an end loop to be unity ($\mathrm{Wr}_\Gamma \approx 1$), the total catenation in the plectoneme phase (Ca_p) made up of m domains is partitioned into the twist (Tw_p), containing the local twisting of the braid, and the writhe (Wr_p), reflecting the overall structure of the plectoneme.

$$\mathrm{Ca}_p = \mathrm{Tw}_p + \mathrm{Wr}_p + m\mathrm{Wr}_\Gamma = (L_p + m\Gamma)\frac{\sin\delta_p}{2\pi R_p} + L_p\cos\delta_p\frac{\sin 2\alpha}{4\pi\mathscr{R}_p} + m \tag{2.24}$$

where R_p is the radius of the braid in the plectoneme state. Considering a simple geometric picture of closely-packed braids, we set the plectoneme superhelical radius to be twice the braid radius in the plectoneme: $\mathscr{R}_p = 2R_p$. As a simplifying assumption, we ignore local structural rearrangements in the plectoneme that may lead to spatially-varying mean-field superhelical radii.

The mean-field free energy of the plectoneme phase is given by

$$\begin{aligned}\beta\mathcal{E}_p =&(L_p + m\Gamma)\left[\frac{A\sin^4\delta_p}{R_p^2} + \frac{1}{A}\mathcal{U}_0(R_p, P_p)\right] + L_p\left[\cos\delta_p\frac{A\sin^4\alpha}{4{R_p}^2} + \frac{2}{A}\mathcal{U}_0(\mathscr{R}_p, \mathscr{P}_p)\right]\\ &+ m\sqrt{2\varepsilon\beta A f} - \ln\Omega(m)\end{aligned} \tag{2.25}$$

where the first bracketed term is the sum total of elastic bending energy and electrostatic energy of the braid inside the plectoneme, obtained from the mean-field term in Eq. (2.9). The second bracketed term is the energy contribution from elastic bending and electrostatic repulsion in the superhelix, where $2\pi P_p$ and $2\pi\mathscr{P}_p$ are the braid pitch and the superhelical pitch in the plectoneme respectively. The factor of 2 multiplying the superhelix electrostatic term is because the length of the superhelix is half of that of the DNA length in the plectoneme while the effective charge is two times that of the double helix. The third term corresponds to the elastic energy of m braid end loops [Eq. (2.23)], and finally, the logarithm term is the free energy associated with the configuration entropy of m loops. The origin of this entropy is from the following two sources: (1) sliding of a plectoneme domain along the braid contour, and (2) exchange of DNA length among the plectonemic domains. We define the arc length corresponding to unit twist in the braid ($\sqrt{R^2 + P^2} \sim 10$ nm at 100 mM Na^+) as the characteristic length distinguishing these energetically degenerate but structurally distinct states. The total number of such states (Ω) for a plectoneme phase constituted of m domains (where $m \geq 1$) can be written as a product of two combinatorial factors [65]:

$$\Omega(m) = \frac{(2\pi \mathrm{Ca}_s)^m (2\pi \mathrm{Tw}_p)^{m-1}}{m!(m-1)!} \tag{2.26}$$

where the first term corresponds to the sliding entropy of m loops ($2\pi \mathrm{Ca}_s$ is the total number of possible plectoneme nucleation sites) and the second term is the number of distinct configurations associated with the exchange of DNA length among the domains. Note, for a plectoneme of given length (L_p), catenation (Ca_p) and number of domains (m), the total free energy has two free parameters [namely, δ_p, α and R_p constrained by Eq. (2.24)] that determine the equilibrium structure.

Similar to the straight braid case [Eq. (2.21)], the fluctuation free energy correction to the mean-field energy of the braid in the plectoneme is obtained from Eq. (2.14), using $\mu_p = \mu(\delta_p)$ and $\eta_p = \eta(\delta_p, \alpha)$:

$$\beta\Delta\mathcal{F}_p = \frac{(L_p + m\Gamma)}{A}\left[\frac{3}{2}\sqrt{\mu_p} + \eta_p^{1/4}\cos\left(\frac{1}{2}\tan^{-1}\sqrt{\frac{4\eta_p}{\mu_p^2} - 1}\right)\right] \tag{2.27}$$

The above expression gives the total free energy associated with worm-like-chain fluctuations in the braiding strands forming the plectoneme structure.

2.1.C Numerical partition function

In an ensemble of fixed catenation and fixed force, the total free energy of the braid can be obtained by minimizing the sum total of the straight and the plectoneme phase energy. We thermally average over states with all possible plectoneme lengths and number of domains, where the free energy in each state is minimized with respect to the partition of the total linking number, thus ensuring torque balance between the two structural phases. The free energy of the braid for each fixed L_p and m is obtained via numerical minimization over Ca_s [Eqs. (2.20), (2.21), (2.25), (2.27)].

$$F(L_p, m) = \min_{\mathrm{Ca}_s}\left(\mathcal{E}_s + \Delta\mathcal{F}_s + \mathcal{E}_p + \Delta\mathcal{F}_p\right) \tag{2.28}$$

The above minimization is constrained by conservation of total catenation ($\mathrm{Ca} = \mathrm{Ca}_s + \mathrm{Ca}_p$) and total DNA length ($L = L_s + L_p + m\Gamma$). The states described by Eq. (2.28) for all possible

values of L_p and m are then thermally averaged over to construct a partition function:

$$\mathcal{Z}(\mathrm{Ca}, f) = e^{-\beta F(0,0)} + \sum_{m=1,2\ldots} \sum_{L_p=\Delta}^{L-d} e^{-\beta F(L_p,m)} \tag{2.29}$$

where the first term is the purely straight phase, and the second term corresponds to a sum over all possible co-existence states. The sum over L_p in Eq. (2.29) is done numerically using a $\Delta = 1$ nm mesh. An averaging scheme as described above takes into account various thermally accessible equilibrium states which allow the possibility of torque fluctuations in the fixed catenation ensemble. A similar approach was taken in Ref. [65] to study supercoiled single DNAs.

2.1.C.i Ensemble-averaged observables

Equilibrium values of the end-to-end distance (z), the torque in the braid (τ), and the size of the helical wrappings in the straight phase (L_b) are obtained from the partition function [Eq. (2.29)]:

$$\langle z \rangle = -\frac{1}{\mathcal{Z}} \left[\frac{\partial F(0,0)}{\partial f} e^{-\beta F(0,0)} + \sum_m \sum_{L_p} \frac{\partial F(L_p,m)}{\partial f} e^{-\beta F(L_p,m)} \right] \tag{2.30}$$

$$\langle \beta\tau \rangle = -\frac{1}{2\pi} \frac{\partial \ln \mathcal{Z}}{\partial \mathrm{Ca}} \tag{2.31}$$

$$\langle L_b \rangle = 2\pi \langle \mathrm{Ca}_s \rangle \sqrt{\langle R_s \rangle^2 + \langle P_s \rangle^2} \tag{2.32}$$

The average values of the pure-state free variables X, and the coexistence-state free variables Y are computed from Eq. (2.29) as follows:

$$\langle X \rangle = \frac{1}{\mathcal{Z}} \left[X e^{-\beta F(0,0)} + \sum_m \sum_{L_p} X \; e^{-\beta F(L_p,m)} \right] \qquad \langle Y \rangle = \frac{1}{\mathcal{Z}} \sum_m \sum_{L_p} Y \; e^{-\beta F(L_p,m)} \tag{2.33}$$

where $X \in \{R_s, P_s, \mathrm{Ca}_s\}$ and $Y \in \{\delta_p, \alpha, L_p, m\}$.

2.2 Results

2.2.A Braids at 100 mM salt

2.2.A.i Higher catenation at a fixed force decreases extension

Figure 2.2(a) shows the comparison of theoretically calculated braid extension curves for various forces with experimental observation at 100 mM univalent salt concentration [56]. The size of the intertether distance d being comparable to the length of the braiding molecules results in a sharp decrease in extension when the first catenation is added. The decrease is due to the formation of the first helical bend in the braid along with the end-regions from the zero-catenation parallel configuration. The extension shortening is used to estimate the intertether distance by simply using the Pythagorean theorem [13, 55, 56]. Notably, the intertether distance d is a parameter that has not been controlled in experiments to date.

2.2.A.ii Nonlinear braid torque

Further addition of catenation decreases the end-to-end extension of the braid [Fig. 2.2(a)] due to the double-helix length being passed from the end-regions to the helically wrapped section. The size of the straight braid, containing helical intertwines, increases with catenation and reaches a maximum just before the onset of buckling [Fig. 2.2(d)]. Elastic bends in the braiding double helices generate torsional stress, which increases nonlinearly with catenation [Fig. 2.2(b)]. The nonlinearity of torque originates from the adjustable helical geometry, and has been seen in previous models of the straight braid [56, 67, 68]. When the torque reaches a critical value, which mainly depends on thermodynamic parameters such as the external force, nucleation of the first plectonemically buckled domain becomes energetically favorable.

Stretched, supercoiled single DNAs at 100 mM salt also show a buckling transition separating a force-extended phase from a plectoneme-coexistence phase [12, 34, 79], although the mechanical response of braids is fundamentally different than that of supercoiled DNAs (see Chapter 3). The torque in a stretched supercoiled DNA double helix increases *linearly* with the linking number

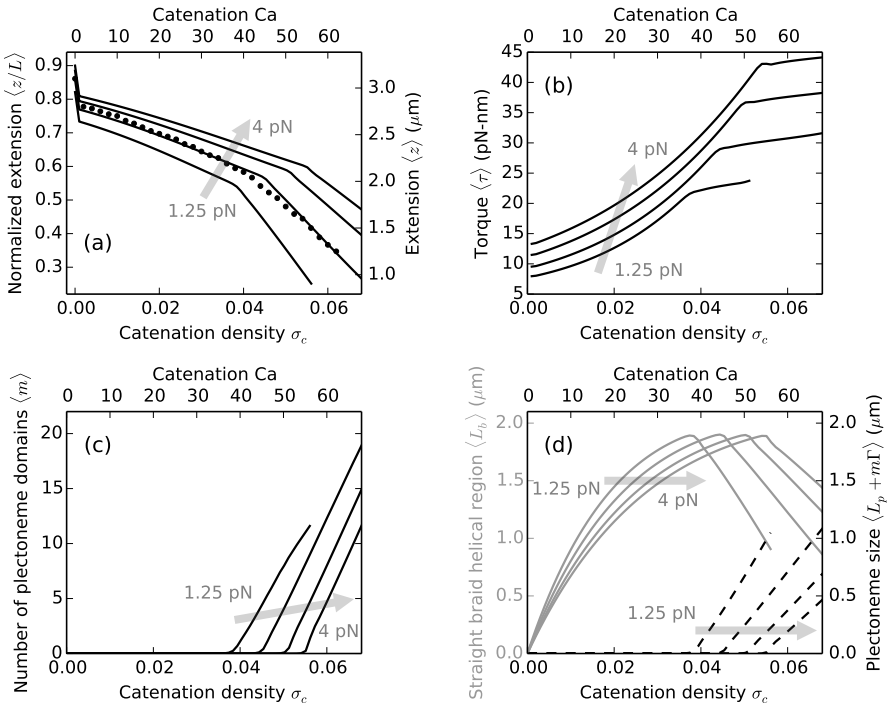

Figure 2.2: DNA braids at 100 mM monovalent salt under various forces, the shaded arrows show the direction of increasing force. Theoretical predictions are for ≈11 kb ($L = 3.6$ μm) long double helices, tethered 1.5 μm ($d = 0.42L$) apart. Catenation (Ca) and catenation density ($\sigma_c = \mathrm{Ca}/\mathrm{Lk}_0$) are plotted on the top and the bottom x-axes respectively. (a) Relative end-to-end distance (left y-axis) or extension (right y-axis) versus catenation. Lines are theoretical predictions for 1.25 (lowest curve), 2, 3 and 4 pN (highest curve) force, while filled circles are experimental data at 2 pN [56]. The change in slope of the lines corresponds to plectonemic buckling transition, which is at a higher catenation for larger external tension due to increased stability of the force-coupled straight state. The kink at the onset of buckling transition is related to the plectoneme-nucleation cost presented by the braid end loop. (b) Torque in the braid shows a non-linear increase in the straight phase, and continues to increase in the coexistence phase but with a much weaker slope. The torsional stress is released in the coexistence phase due to the contribution from plectoneme writhe [Eq. (2.24)]. (c) Number of plectonemic domains versus catenation, showing that the buckled phase is characterized by multiple plectoneme domains. Nucleation of new domains causes the increase of torque in the coexistence phase, as opposed to a constant torque expected in the case of a single plectoneme domain. (d) Plot of the size of the straight-phase helical region $\langle L_b \rangle$ (left y-axis, solid gray curves) and the size of the plectoneme region $\langle L_p + m\Gamma \rangle$ (right y-axis, dashed black curves) as a function of catenation. L_b increases in the straight phase with catenation till the buckling point, after which L_b decreases as DNA length is passed into the plectoneme phase, also seen in the increase in the total size of the plectoneme.

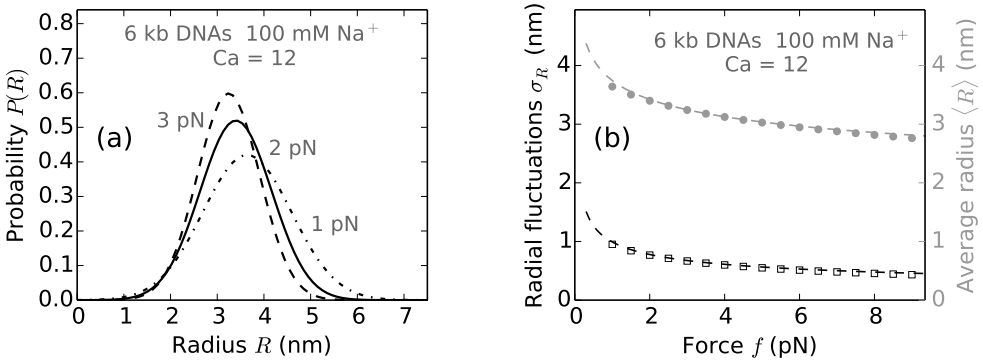

Figure 2.3: Radial fluctuations in braid with $L = 2$ μm, $d = 800$nm and Ca$= 12$. (a) Probability distribution of braid radii [Eq. (2.18)], plotted for applied forces of 1 pN (dot-dashed line), 2 pN (solid line) and 3 pN (dashed line). Note the broadening of the distribution at lower forces indicating higher fluctuations. $R = 1$ nm corresponds to the excluded radius of the DNA molecules due to self avoidance, included while calculating the average radius via free energy minimization but not explicitly taken into account in plotting Eq. (2.18). (b) Variation of the standard deviation of the radial distribution σ_R (black open squares) and the average value of the radius $\langle R\rangle$ (shaded filled circles) with the external tension (f) on the braid. The dashed lines correspond to best-fit equations: $\sigma_R = \sigma_0(\beta A f)^{-0.35}$ and $\langle R\rangle = R_0(\beta A f)^{-0.13}$, where $\sigma_0 = 2.4$ nm and $R_0 = 5.2$ nm. The exact power laws depend on the catenation in the braid, however, both the average value and the fluctuations in braid radius always decrease with increasing force.

[35, 83], as opposed to a non-linear increase in braids. The linearity of torque in supercoiled single DNA arises from a constant twist modulus in the double helix ($C \sim \partial\tau/\partial\sigma_{sc} \approx 100$ nm [60, 79, 83]), which is attributed to the strong base-pairing interactions holding the DNA strands together. Conversely, braids are soft structures (the two braiding molecules are not attached to each other), where twist-stiffening occurs as the catenation is increased, making the twist modulus of braids a quantity that depends on catenation.

2.2.A.iii Abrupt plectonemic buckling in braids

The onset of buckling can be identified as a *knee* in the extension plots [Fig. 2.2(a)], past which DNA length is passed into the force-decoupled buckled phase [Eq. (2.25)], resulting in a steeper decrease of the end-to-end extension. In the coexistence region, the writhe contribution to the total linking number reduces torsional strain in the braid. The torque in the braid shows a small non-monotonic "overshoot" at the buckling transition, and continues to increase with a

small slope in the coexistence phase. The abruptness of the buckling transition owes to the finite-energy cost to nucleating a plectoneme domain, vis-à-vis the plectoneme end loop.

2.2.A.iv Multiple plectonemic domains in buckled braid

The coexistence phase at 100 mM salt is characterized by multiple domains of plectoneme [Fig. 2.2(c)], where the number of domains is equal to the equilibrium number of plectoneme end loops. The total size of the plectoneme phase increases after the buckling transition [Fig. 2.2(d)], where DNA length is transferred to the buckled region from the straight phase. The higher number of domains is a consequence of the structural bulkiness of braids. Braid plectonemes require a higher radii, because of the self-avoidance between the two writhing braids, which destabilizes the plectoneme state relative to the end loops, resulting in a large number of domains in the buckled state.

2.2.A.v Higher force stabilizes straight braid and decreases fluctuations

Higher external tension lowers the total energy of the straight braid [Eqs. (2.20), (2.21)], resulting in a higher end-to-end extension at a given catenation [Fig. 2.2(a)]. The stabilization of the straight phase upon increasing force has an effect of delaying the buckling transition, i.e, buckling occurs at a higher value of catenation. The torque is also higher in both the straight and the buckled braids under a larger external tension [Fig. 2.2(b)].

Figure 2.3(a) shows the probability distribution of braid radius for various forces. Higher forces result in a smaller average radius of the braid with less radial fluctuations. Figure 2.3(b) shows the variation of the mean and the fluctuation in braid radius for various forces at 100 mM salt. Radial fluctuations decrease with increasing force and are much smaller than the average value of the radius. This indicates that the braiding DNAs show small fluctuations about their average shape when stretched under a higher tension.

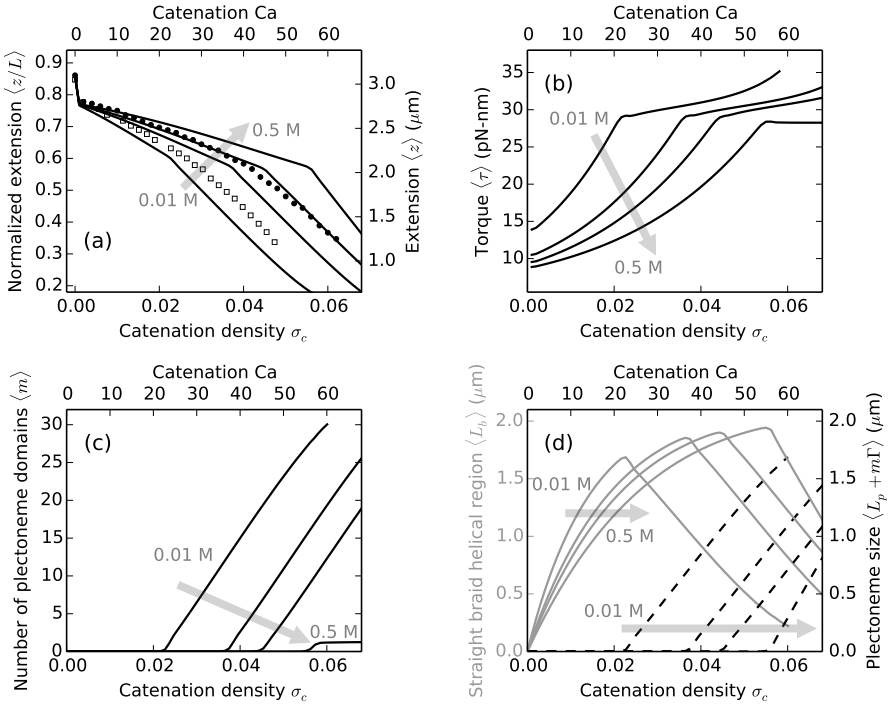

Figure 2.4: Effect of salt concentration on ≈11 kb ($L = 3.6$ μm) DNA braids under 2 pN force. The tether points are 1.5 μm ($d = 0.42L$) apart. Theoretical curves are plotted for (0.01, 0.05, 0.1 and 0.5) M salt concentrations, where the shaded arrows show the direction of increasing salt concentration. Catenation (Ca) and catenation density ($\sigma_c \equiv \mathrm{Ca}/\mathrm{Lk}_0$) are plotted on the top and the bottom x-axes respectively. (a) Relative end-to-end extension (left y-axis) or extension (right y-axis) versus catenation shows smaller extension and buckling at a lower catenation for lower salt concentrations. Low salt increases the effective DNA diameter, which effectively increases the twist elasticity of the braid, thereby decreasing the stability of the straight phase. The filled circles (0.1 M) and open squares (0.01 M) are experimental data reproduced from Ref. [56]. The total length of DNA for 10 mM and 100 mM experimental data sets [56] differ slightly (~ 0.2 μm); we renormalized the length of the 10 mM case (open squares) to be close to that of the 100 mM data (filled circles) for comparison. (b) Torque in the braid shows a non-linear increase in all salt conditions. Also, twist stiffening occurs faster for braids at lower salt concentration due to the larger radius of the braid. The critical buckling torque, being a thermodynamic variable does not vary significantly with the salt concentration. (c) The number of plectoneme domains versus catenation or catenation density, showing nucleation of multiple domains of plectoneme at lower salt concentrations, while a single plectoneme state is favored at higher salts. Smaller excluded diameter of the braid at higher salt makes the superhelical bending in the plectoneme phase favorable over nucleation of new domains. (d) The size of the straight phase helical wrappings $\langle L_b \rangle$ (left y-axis, solid gray curves) and the size of the plectonemic phase $\langle L_p + m\Gamma \rangle$ (right y-axis, dashed black lines) versus catenation. L_b increases faster for lower salt concentrations due to larger braid radii (Table 2.1).

2.2.B Braid mechanics at various salt concentrations

2.2.B.i Electrostatic screening is weaker at lower salt concentrations

Lowering the ionic strength of the solution causes an increase in excluded volume of the double-helix DNA in the solution, due to less screening of the negative charges on the double-helix backbone. At low salt, the Debye length (λ_D) of the solution is higher, causing the Coulomb-repulsion effect to propagate a longer distance before it is cut off, thus increasing the effective diameter of the double helix.

2.2.B.ii Braids buckle at lower catenations for lower salt concentrations

Figure 2.4(a) shows braid extension curves under various salt concentrations. The buckling transition occurs at a lower catenation for braids under lower salt concentrations. The larger effective DNA diameter increases the braid radius at lower salts (Table 2.1), which destabilizes the straight phase and consequently causes buckling at a lower catenation number. The predicted trend of buckling with varying salt has been observed experimentally [56].

In the supercoiled single DNA case, the opposite trend is observed, where lowering the salt concentration of the solution makes plectonemic buckling occur at a higher supercoiling density [84]. In stretched supercoiled DNA, the stability of the force-extended phase is unaltered by changing the ionic strength, but the supercoiled plectoneme phase (containing intra-molecular writhes) is relatively destabilized on decreasing the salt concentration, again due to the increase in effective diameter of the double helix, resulting in the observed trend of buckling at a lower linking number for higher salt concentrations.

The torque in the braid plotted as a function of the catenation [Fig. 2.4(b)] shows a nonlinear increase and a small non-monotonic overshoot at the buckling transition at all salt conditions. For a given catenation, the torque is higher for lower salt concentrations due to effective swelling of the braid. The torsional stress in the braid decreases with increasing salt, a trend also observed in supercoiled DNAs [79].

2.2.B.iii Braid plectonemes are stabilized at higher salt concentrations

Figure 2.4(c) shows that multiple plectoneme domains are favored at lower salt conditions, while single domain plectonemes are favored at higher salts. This is due to the lower excluded volume of DNA at higher salts leading to a more stable plectoneme state. At lower salts, larger braid diameter increases the energy associated with superhelical bending of the braid in the plectoneme, making the formation of looped structures of braid favored over a superhelical structure, consequently favoring formation of multiple domains of plectoneme. At higher salts, the average number of superhelical turns per plectoneme domain is higher, i.e., the average size of each plectoneme domain is larger [Fig. 2.4(d)]. This effect is directly related to the decrease in DNA excluded volume at higher salt concentrations (Table 2.1), which stabilizes the superhelical bends in a braid plectoneme.

2.2.C Braid intertether distance characterizes mechanics

2.2.C.i Braids with a larger intertether distance buckles at a lower catenation

The intertether distance d between the two braided molecules affects the critical catenation, i.e., the catenation at which buckling occurs [Fig. 2.5(a)]. Braiding molecules with larger d makes a helix with a larger aspect ratio (ratio of radius to pitch of the helix), which causes a steeper increase of the torque in the braid [Fig. 2.5(b)]. In effect, the twist modulus of the braid is larger for larger intertether distances and causes buckling at a lower catenation [Fig. 2.5(a)]. The torque at which a braid buckles is a thermodynamic property dependent on the external force and remains roughly the same on changing the intertether distance [Fig. 2.5(b)] or the salt conditions [Fig. 2.4(b)].

2.2.C.ii Asymmetric braid tethers

The difference in extension between the states Ca = 0 and 1 is dependent on the intertether distance, but the two intertether distances at the two ends of the braid need not be the same. In fact, these distances being uncontrolled parameters in experiments, are almost never equal.

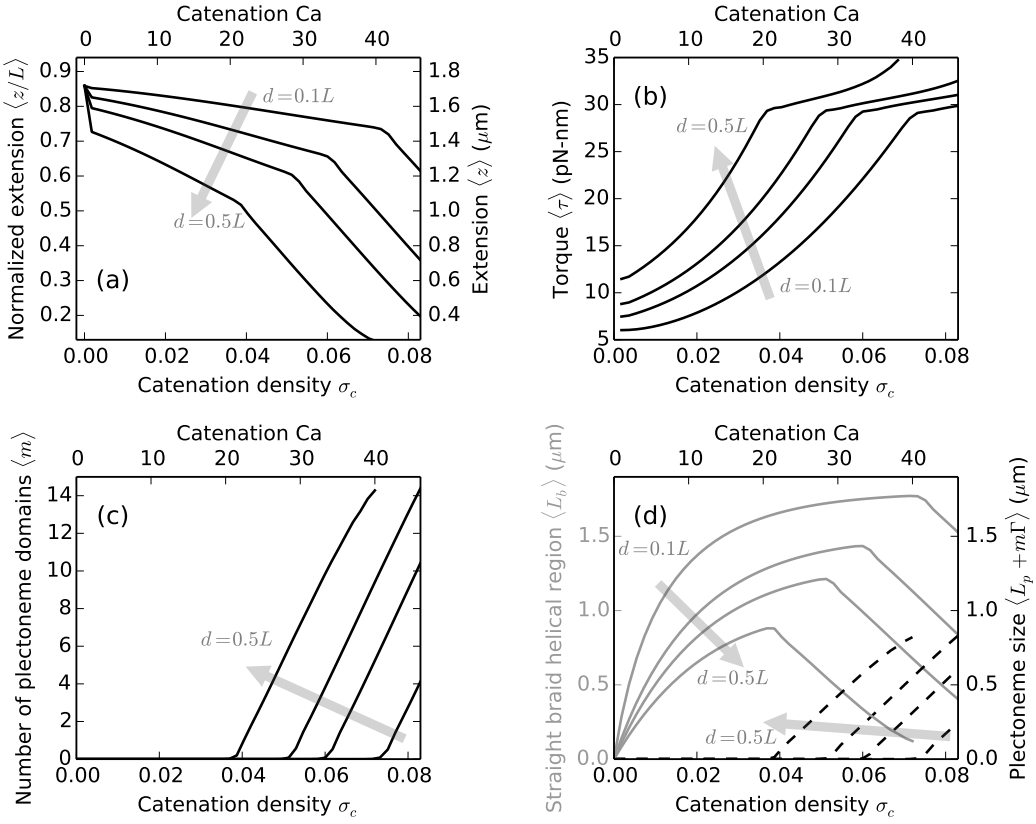

Figure 2.5: Effect of the intertether distance for braids with ≈6 kb (L = 2 μm) long DNAs under 2 pN force at 100 mM salt. The shaded arrows show the direction of increasing intertether distance $d = 0.1L,\ 0.25L,\ 0.35L$ and $0.5L$, where the top and the bottom x-axes show catenation (Ca) and catenation density ($\sigma_c \equiv \mathrm{Ca}/\mathrm{Lk}_0$) respectively. (a) Variation of relative extension (left y-axis) or extension (right y-axis) with catenation in the braid. Larger intertether distance results in a larger initial jump in the extension and lowering of the critical catenation density. (b) The torque in the braid is higher for larger intertether distances. The increase of torque per unit catenation or effective twist modulus of the braid is also higher for larger intertether distance, resulting in buckling at a lower catenation. However, the critical value of torque at which buckling occurs is a bulk property of the braid and remains roughly the same for various intertether distances, ≈ 30 pN-nm at 2 pN external force. (c) The average number of plectonemes as a function of catenation showing the formation of multiple domains at all intertether distances. (d) The size of the helical straight braid (left y-axis, solid gray curves) and the total plectoneme length (right y-axis, dashed black curves) as a function of catenation. For braids with larger intertether distance, buckling occurs at relatively smaller size of the straight braid helical section due to the larger size of the triangular end regions [Figure 2.1(a)].

The theoretically predicted extension plots are a characteristic of the arithmetic mean of the intertether distances at the two ends of the braid; although, not surprisingly, the structure of the

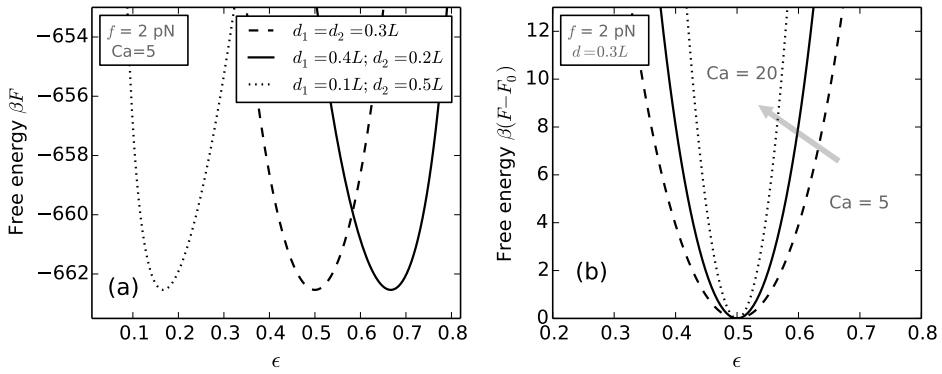

Figure 2.6: Effect of the intertether distance. (a) Comparison of the free energy versus ϵ [Eq. (2.34)] for braids with various choices of the two intertether distances d_1 and d_2, but keeping the arithmetic mean of the two distances the same. The end-region where the two DNAs are closer to each other is smaller in size (i.e., has less DNA in it), and vice-versa. (b) Free energy versus ϵ for braids at various catenations (Ca=5, 12 and 20) with identical intertether distances $d = 0.3L$. The stiffness of the free energy decreases for lower catenation values indicating higher fluctuations in the relative sizes of the two end regions. The free energy curves have been shifted in the y-axis in order to overlay them.

braid depends on the specific values of the two intertether distances [Fig. 2.6(a)]. In the case of unequal intertether distances, we find that the end-region associated with the larger intertether distance is bigger in size, i.e., the helically wrapped region of the braid does not form at the center of the structure but is pushed towards the end with the smaller intertether distance.

We also find that the energy cost of fluctuations in the relative size of the two end regions, i.e., small displacement of the entire helical section away from the equilibrium position are $\mathcal{O}(k_B T)$, hence permissible, especially in the regime of low catenation [Fig. 2.6(b)]. The energy cost increases with increasing catenation, reflecting the sliding of the helical section is energetically expensive when the torque in the braid is higher. Such behavior of a braid may be possible to probe in braiding experiments done on DNA molecules labeled along their length with fluorescent tags.

For a braid with unequal intertether distances d_1 and d_2, we define a dimensionless parameter

ϵ such that

$$\epsilon = \frac{\ell_1}{\ell_1 + \ell_2}, \tag{2.34}$$

where the total length in the end-regions L_e [Eq. (2.20)] is divided into ℓ_1 and ℓ_2 corresponding to the two end-regions. Hence, $\epsilon = 1/2$ indicates the scenario of a braid with symmetric end-regions. Figure 2.6(a) shows the minimized total free energy [Eqs. (2.20), (2.21)] versus ϵ for three pairs of intertether distances, keeping the arithmetic mean of the distances the same in all three choices. The extension plots are a characteristic of the arithmetic mean, but the relative sizes of the end-regions depend on the particular choice of d_1 and d_2. The end-region associated with the larger of the two intertether distances contains more length of double-helix DNA.

Figure 2.6(b) shows minimized free energy versus ϵ for a braid with symmetric ends but various catenations. The stiffness of the potential near the equilibrium value of ϵ increases with increasing catenation in the braid. This suggests the possibility of fluctuation in the relative sizes of the end-regions, which would be higher for lower catenations and becomes energetically expensive in the regime of tight braid or braid with high catenation density. It may be possible to directly visualize the sliding of the braid helical intertwines in a braiding experiment using DNA molecules labeled with fluorescent tags.

2.2.D Braiding short DNA molecules

Braiding DNA molecules shorter than ≈3 kb in size shows the same qualitative trends of extension decrease with catenation and formation of multiple buckled domains past the critical catenation density, as seen for larger molecules. However, due to the small size of the molecules and hence the smaller number of fluctuation states, discrete nucleation of buckled domains may be observed as steps in the extension plot [Fig. 2.7(a)]. The torque also shows multiple overshoots associated with nucleation of new buckled domains [Fig. 2.7(b)-(c)]. Similarly, a step-like behavior can be seen in the length transfer between the straight and the plectoneme phase. The discrete jumps in extension or torque are masked by thermal fluctuations in longer molecules

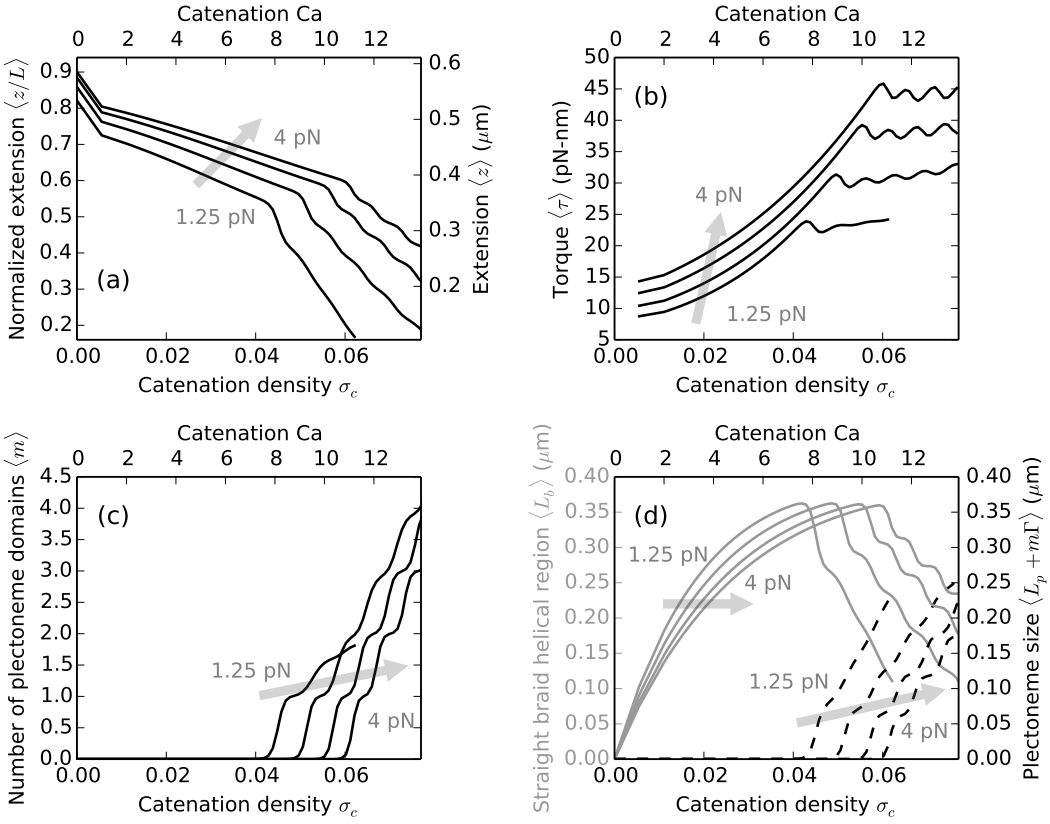

Figure 2.7: Braiding ≈2 kb ($L = 0.65$ μm) DNA at 100 mM Na^+ with forces $f = 1.25$, 2, 3 and 4 pN, where the shaded arrows show the direction of increasing force. The intertether distance is 0.26 μm ($d = 0.4L$), where the top and the bottom x-axes show catenation and catenation density ($\sigma_c \equiv \mathrm{Ca/Lk_0}$) in the braid. (a) Relative extension (left y-axis) or extension (right y-axis) versus catenation for short DNA molecules. (b) Torque vs catenation shows multiple discrete "overshoots", corresponding to the nucleation of new plectoneme domains. (c) Number of plectoneme domains vs catenation. The appearance of new plectoneme domains coincides with the steps in the extension or overshoots in the torque. (d) The size of the straight-phase helical region $\langle L_b \rangle$ (left y-axis, solid gray curves) and the size of the plectoneme phase $\langle L_p + m\Gamma \rangle$ (right y-axis, dashed black curves) versus catenation, where steps are associated with the formation of finite-length braid end loops. The successive nucleation events are smoothed out by thermal fluctuations in braids made up of long DNA molecules (> 4 kb).

(> 4 kb or 1.3 μm). Similarly, at lower forces, large fluctuations make the average extension of the braid decrease more uniformly [Fig. 2.7(a)].

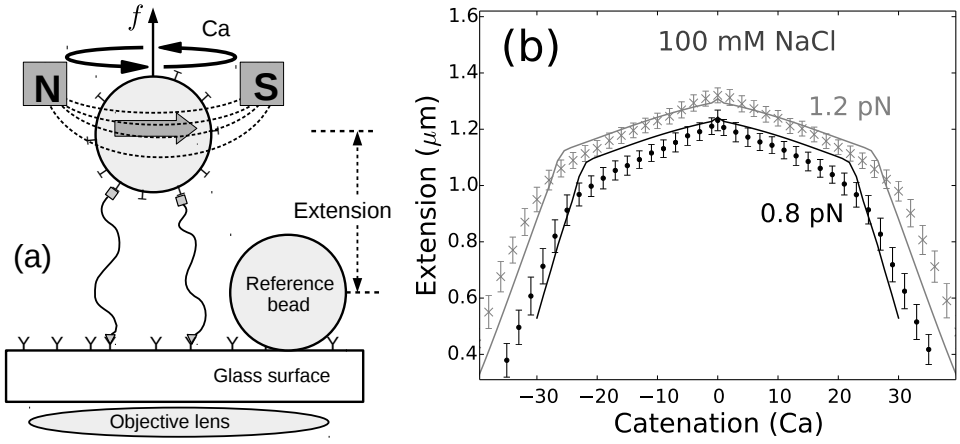

Figure 2.8: (a) Schematic of the magnetic tweezers setup. Two DNA molecules are attached (using only one of the DNA strands as attachments) to the glass surface and a paramagnetic bead via digoxigenin-antidigoxigenin and biotin-streptavidin interactions, respectively. The single-strand attachments ensure no twisting of individual DNA molecules upon rotation of the bead [16, 88]. (b) Experimentally measured end-to-end extension of a braid as a function of catenation number for 0.8 (black points) and 1.2 pN (gray cross marks) applied forces in 100 mM NaCl, where the error bars represent standard deviation. The solid lines represent theoretical predictions for 1.6 μm DNA molecules, with intertether distance 0.19 μm, and under 0.8 (black line) and 1.2 pN (gray line) force at 100 mM monovalent salt. The small peak in extension at zero catenation is due to the small intertether distance, *i.e.*, the close proximity of the two DNA molecules for this particular DNA pair [46].

2.3 Comparison with experiments

2.3.A Braid-extension measurement using magnetic tweezers

We used bright-field magnetic tweezers [88] to study braided DNAs, where we attached one pair of ends of the two double-helical DNAs to a glass surface and the other pair of ends to a one-micron paramagnetic bead [16] [Fig. 2.8(a)]. The inter-DNA linking (or "catenation") number in the braid is controlled by rotating the bead using the magnet, whereas the applied force is controlled by varying the distance between the magnet and the bead. To ensure that each DNA is not subject to double-helix twisting torque, only the 5′-ends of the DNAs were attached to the surfaces, allowing swiveling of the free strand about the tethered one, which makes the number of turns of the bead a direct measure of catenation in the braid. All experiments were carried out in 100 mM NaCl buffer.

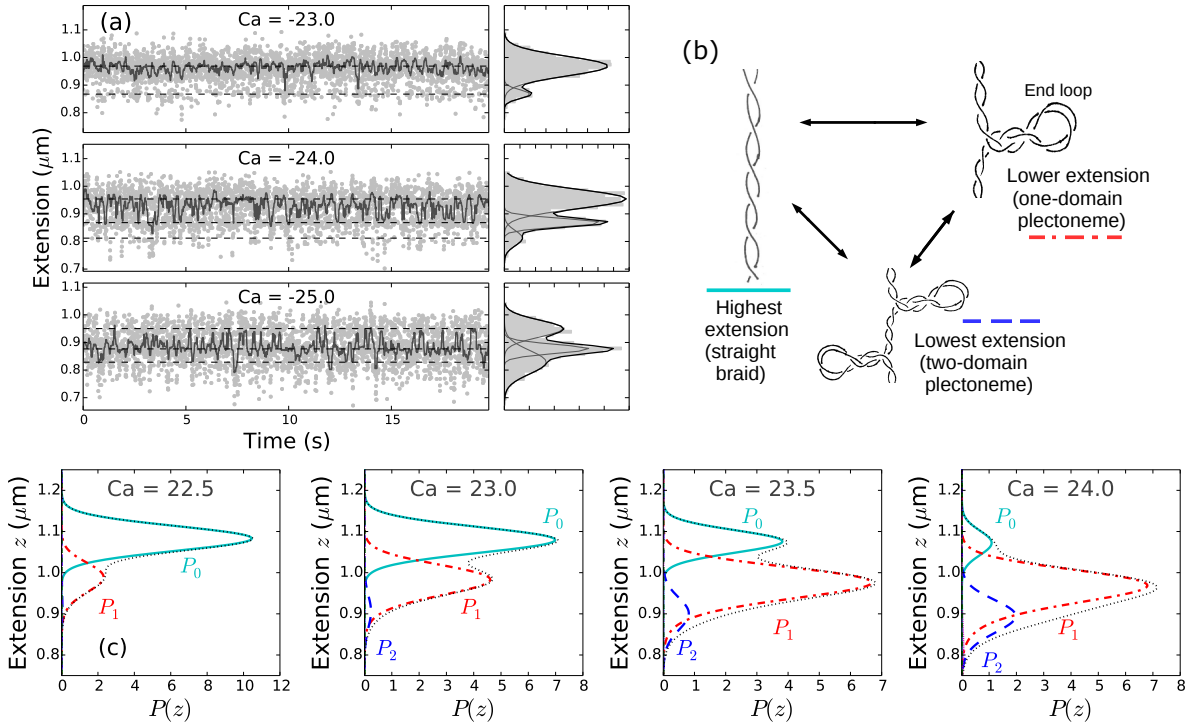

Figure 2.9: (a) Time series of end-to-end extension of a braid constituted of two torsionally-unconstrained 6 kb DNAs, under 0.8 pN force at 100 mM NaCl [Fig. 2.8(b)], for three catenation numbers (Ca = -23, -24, and -25) near the buckling transition point [the point of slope change in the extension curve, Fig. 2.8(b)]. Data were collected at 200 Hz (light gray dots), then median filtered (dark points) using a 0.1 sec time window to show dynamic switching between discrete extension states. The panels on the right of each time series plot show histograms of the raw data using 10 nm bins (gray shaded area); the y-axis is the same for left and right panels. The histograms were fit to a sum of multiple Gaussian distributions, where the dark line is the best-fit distribution and corresponds to the sum of the individual Gaussians shown in gray lines. The sum of two Gaussian distributions fit the data better at low catenation (Ca=-23), indicating nucleation of the first plectoneme domain, whereas, the sum of at least three Gaussians is required to fit the histograms at higher catenation (Ca=-24, -25), due to the appearance of multiple plectoneme domains. (b) Schematic diagram of three braid-extension states accessible via thermal fluctuations near the buckling transition. (c) Theoretically predicted extension histograms near the buckling transition, where the black dotted line shows the total extension distribution (P_{tot}); also plotted are the individual contributions from the straight braid (P_0, cyan solid line) and the buckled braid with one (P_1, red dot-dashed line) and two (P_2, blue dashed line) plectoneme domains [see Eq. (2.36)]. Contributions from the three (P_3) and the four-domain (P_4) plectonemes are also plotted, however, the negligible statistical weight of those states for the plotted range of catenation renders them almost invisible in the predicted histograms.

As the magnet is rotated, the extension of the braid under fixed force decreases when the catenation is increased, producing a characteristic bell-shaped curve [Fig. 2.8(b)]. The change in extension of the braid between catenation numbers 0 and 1 is related to the distance between the two tether points; closely-spaced DNA tethers produce a small jump, whereas, a larger separation between the DNA tethering points show a sharper initial jump [13, 46, 55, 56]. In Fig. 2.8(b), the jump is small relative to the length of the braided molecules ($\approx$ 1.6 μm), indicating that the two DNAs are tethered close together ($\approx$ 0.19 μm). After the first catenane is introduced, the extension of the braid decreases with increasing catenation due to formation and consequent increase in size of the helically-wrapped region of the braid. Since the individual DNA molecules cannot be supercoiled, the extension plots are symmetric for positive and negative catenations [Fig. 2.8(b)]; our model assumes this symmetry as there are no DNA-twist-energy terms in the free energy expressions.

2.3.B Multimodal distribution of measured braid extension

Appearance of a plectoneme domain requires nucleation of the braid end loop, which causes a discrete change in braid extension since the end loops are finite-sized structures. Figure 2.9(a) shows data for a time series of braid extension under 0.8 pN force and 100 mM NaCl salt concentration, at a fixed catenation near the buckling transition, *i.e.*, near the point of slope change in the extension curves [Fig. 2.8(b)]; and the histograms show the probability density of braid extension. Near the buckling transition point [Ca=-23 at 0.8 pN, see Fig. 2.8(b)], the probability distribution of braid extension is bimodal [Fig. 2.9(a)], where the higher and the lower extension peaks respectively correspond to the straight braid and the one-domain plectoneme braid (i.e., a plectoneme with one end loop).

In the vicinity of the buckling transition, the occupancy of the lower-extension state increases with increasing catenation, due to the appearance of the first buckled plectoneme domain. Simultaneously, the occupancy of the higher-extension state decreases as the purely-straight state of the braid disappears. The data also show appearance of multiple discrete-extension states

after the nucleation of the first domain [multiple peaks in the histograms for Ca=-24 and -25, see Fig. 2.9(a)], where the lowest-extension state corresponds to a two-domain plectoneme braid [plectoneme with two braid end loops, see Fig. 2.9(b)].

2.3.C Calculated probability distribution of braid extension

Using the partition function described in Eq. (2.29), we consider the probability of each state considered in the partition sum. The extension distribution in each of the summed-over state is Gaussian:

$$\mathcal{P}_{m,L_p}(z) = \frac{1}{\sqrt{2\pi}\Delta_{m,L_p}} \exp\left[-\frac{(z-\bar{z}_{m,L_p})^2}{2\Delta^2_{m,L_p}}\right] \tag{2.35}$$

where the mean and the variance are respectively given by: $\bar{z}_{m,L_p} = -\partial F/\partial f$, and $\beta\Delta^2_{m,L_p} = -\partial^2 F/(\partial f^2)$. Here, $F(L_p, m)$ is the total free energy of the plectoneme-coexistence state, which includes the force-coupled straight braid, superhelically-bent braid of size L_p, and m end loop(s) [Eq. (2.28)]. The total distribution of extension at a given catenation is obtained from summing all the corresponding Gaussian distributions [Eq. (2.35)] with their respective Boltzmann weights:

$$P_{tot}(z) = \mathcal{P}_{0,0}\frac{e^{-\beta F(0,0)}}{\mathcal{Z}} + \sum_m \sum_{L_p} \mathcal{P}_{m,L_p}\frac{e^{-\beta F}}{\mathcal{Z}} = P_0 + \sum_{m=1,2,\cdots} P_m \tag{2.36}$$

where $P_0(z)$ and $P_m(z)$ are the respective contributions to the total extension probability distribution $P_{tot}(z)$ from the straight and the m-domain plectoneme states.

Figure 2.9(c) shows the calculated probability distributions of extension near the buckling transition point for braids at 100 mM monovalent salt condition [Fig. 2.8(b)]. The total probability distributions P_{tot} [black dotted lines, Fig. 2.9(c)] are multimodal, similar to the experimentally observed histograms. The individual Gaussian contributions from the straight phase (P_0), and the plectoneme phase with one (P_1) and two (P_2) domains are also plotted in Fig. 2.9(c). Near the buckling point, increasing the catenation in the braid makes the purely-straight

braid (P_0) less favorable than the one-domain plectoneme-coexistence state (P_1). Further increase in catenation number leads to the nucleation of new plectoneme domains, which gives a strong asymmetric character to the extension distributions. In Fig. 2.9(c), the contributions corresponding to three (P_3) and four-domain (P_4) plectonemes are also plotted, however, they are almost invisible due to negligible statistical weights of the states.

2.3.D Bulky structure of braids favor nucleation of multiple buckled domains

Braids being bulky structures favor multiple small plectoneme domains over a single large one, where structural bulkiness is derived from bending stiffness and excluded diameter of the braids. Since braids have two wrapped double helices, the effective braid bending stiffness is twice that of a double helix. Also, due to the electrostatic interactions, braids have an excluded diameter which is at least twice of that of a single double helix. Larger excluded volume increases the lower bound on braid-plectoneme diameter, which destabilizes the superhelical state relative to the braid end loops [46]. Increased bulkiness makes the two-domain plectoneme structure fluctuation-accessible, and results in the appearance of a finite-probability state with extension lower than that of the one-domain plectoneme. The probability of occupancy of the one-domain plectoneme state increases past the onset of the buckling transition, and then decreases as the two-domain state becomes more probable. From the median-filtered time signal [Fig. 2.9(a)], we estimate the nucleation rate of a braid plectoneme state $\approx 10\ s^{-1}$, which is similar to that observed for a plectoneme domain in twisted single dsDNA [89].

2.3.E Buckling transition is more abrupt than consequent nucleation of domains

Figure 2.10 shows the comparison of theoretically predicted change in extension upon nucleation of the first and the second domain of braid plectoneme with experimental data. For both theoretical calculations and experimental measurements, we define the extension jump upon nucleation of the first plectoneme domain as the distance between the means of the extension

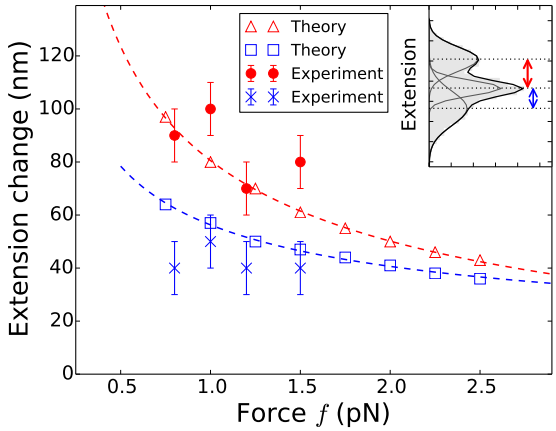

Figure 2.10: Comparison of theoretically-predicted change in extension upon nucleation of the first (red open triangles) and the second (blue open squares) plectoneme domain with experimental observations. Red points and blue cross marks represent the difference in extension between successive peaks (inset) in the experimental best-fit histograms, where the second jump (blue cross markers) is measured one unit catenation after the first jump (red points). The error bars represent binning error. The red and blue dashed lines show the expected $f^{-1/2}$ power law [Eq. (2.23)], best-fit to red open triangles and blue open squares, respectively.

distributions corresponding to the straight (P_0) and one-domain braid plectoneme (P_1), when both the states are almost equally-likely and most-probable. Similarly, the extension jump associated with nucleation of the second plectoneme domain is defined as the distance between the means of the extension distributions corresponding to one (P_1) and two-domain plectonemes (P_2), one catenation unit after the first jump is measured.

The predicted magnitude of extension jump upon nucleation of a plectoneme domain decreases with increasing force due to the decrease in size of the nucleated braid end loop [Eq. (2.23)], although this trend is not apparent in the experiments [Fig. 2.10]. However, we find both theoretically and experimentally that the extension change associated with nucleation of the second plectoneme domain is significantly smaller than that of the first one. This suggests that the nucleation energy cost of the first domain is larger than that of the second domain, or the nucleation of the first buckled domain is more abrupt than the consequent addition of buckled domains.

2.4 Summary

In this chapter we have presented a statistical-mechanical model for a pair of intertwined double-helix DNAs or a DNA braid [Fig. 2.1(a)]. We implemented a helical mean-field structure for the intertwined DNAs and treated thermal fluctuations as harmonic perturbations about the mean-helical structure. Our treatment of thermal fluctuations is consistent with the expected scaling for a worm-like chain in a confined tube [80], however, our approach is free of unknown scaling constants. We also presented experimental data that agrees well with theoretical predictions [46, 47].

Braids exhibit nonlinear torque due to twist stiffening Two intertwined double-helix DNAs are torsionally stressed due to their helical geometry. Note that DNA torque, originating from twisting of individual double-helices is absent in the current model and experiments, because DNAs are nicked or torsionally unconstrained. Braid torque is due to the wrapping geometry that consists of helical bends, giving rise to a torsional response. Braids with a higher catenation have a higher torque, due to the higher catenation density in the braid. However, the positive scaling of braid torque with catenation is nonlinear [Fig. 2.2(b)]. This is because braids are soft structures where the geometry may be adjusted, leading to a torque response that is weaker for smaller catenations but stronger when the catenation density in the braid is higher.

The slope of the torque-versus-catenation plot gives braid twist stiffness [Fig. 2.2(b)], which also has a positive scaling with catenation. Braids with higher catenation are more stiff to torsional perturbations due to a more rigid structure. The twist-stiffening behavior is due to the soft structure of braids that lack an inherent rigidity; on the other hand, dsDNA has a covalently bonded backbone that provides structural rigidity and a constant twist stiffness (see Chapter 3).

Torsionally stressed braid buckles abruptly into a braided plectoneme structure Torsional stress, originated from a high density of catenation, increases the free energy of the

braid to an extent that buckling becomes favorable. Buckling of the braid centerline gives a writhe contribution to the total catenation which does not couple to torque. Hence, increasing catenation in the buckled state results into an increased writhing of the braid due to a plectonemic structure, which keeps the torque from increasing. Buckling, however, costs additional bending energy, implying that buckling of the braid centerline will be favored only when the torsional stress is high enough that buckling decreases the total free energy of the braid.

Braid buckling is associated with the nucleation of a braid-plectoneme end loop [Fig. 2.1(a)], which posits an finite-energy barrier to buckling. The energy associated with the braid loop nucleation makes the buckling transition abrupt. The abruptness of the transition is reflected in a bimodal probability distribution of braid extension at the buckling transition, where the two modes correspond to the buckled and unbuckled states. Because the buckled state cannot be smaller than the braid end loop, the extension at the buckling point shows a discontinuity. Experimental braid extension shows dynamic switching between discrete extension states, and has a bimodal probability distribution, both of which occur due to the finite-sized braid loop necessary to nucleate a buckled domain.

Buckled braid features coexistence of multiple plectoneme domains Braids are bulky structures: the braid radii is at least twice the radii of dsDNA, and the bending stiffness of braids is also twice that of dsDNA. This bulkiness destabilizes the braid plectoneme state, because the energy of the plectoneme state scales positively with the plectoneme radii. Destabilization of the plectoneme state leads to proliferation of multiple coexistent domains, instead of increasing the length of the plectonemically buckled region.

Coexistence of multiple buckled domains in braid gives rise to a multimodal probability distribution of braid extension in the buckled state. Experimental distribution of extension indeed shows a multimodal distribution [Fig. 2.9]. Nucleation of the first buckled domain results in a larger extension discontinuity than that associated with the proliferation of subsequent domains, suggesting that the nucleation cost of the first buckled domain is the highest [Fig. 2.10].

Intertether distance characterizes braid mechanical response Braid geometry is affected by the intertether distance which also influences its mechanics. Braids with a larger intertether distance show a stronger twist stiffening response and buckles at a lower catenation. The critical buckling torque for braids is established at a lower catenation for larger intertether distances. Braids with asymmetric tethers, although does not affect the buckling response, have an asymmetric geometry, where the end corresponding to a larger intertether distance is longer. The fact that the mechanical response of braids can be controlled by altering the intertether distance may be biologically relevant.

Braid mechanics depends on the ionic strength of the solution Reduced ionic strength of the solution increases the electrostatic screening length. This has an effect of an increase in the effective DNA diameter due to stronger self-avoidance between DNA segments. The larger braid radii at lower salts destabilize the braided state and causes buckling at a lower catenation [Fig. 2.4]. The twist stiffening in braids with an increasing catenation occur at a faster rate at lower salts [Fig. 2.4(b)]. At higher salts, due to strong electrostatic screening, the braid diameter is lower. This not only stabilizes the braid resulting in buckling at a higher catenation, it also stabilizes the braid-plectoneme state leading to growth of a plectoneme domain rather than proliferation of multiple domains [Fig. 2.4(c)].

Braiding short DNAs It may be interesting to study short braided DNAs ($\approx$ 2 kb or 650 nm) because we find that for molecules of that short length, the successive addition of small buckled domains leads to a series of buckling transitions [Fig. 2.7], possibly observable as steps in the extension. This is directly related to less thermal fluctuations in smaller DNA molecules. The effect of variable curvature is more prominent in short braids, which may provide additional fluctuations that further mask discrete jumps in extension for short DNA braids.

Overall, the contrast between the mechanical properties of braided and supercoiled DNAs can be simply interpreted as a result of the structural bulkiness and linking-number-dependent elastic moduli in braids, and is well explained by the theoretical model [46]; whether these

potential differences directly influence cellular function or the way proteins interact and modify DNA remains to be determined.

Salt conc.	Debye length	Effective charge	Straight braid (Ca=10)		Buckled braid (Ca=60)			
(M)	λ_D (nm)	ν (nm^{-1})	$\langle R_s \rangle$ (nm)	$\langle P_s \rangle$ (nm)	$\langle R_s \rangle$ (nm)	$\langle P_s \rangle$ (nm)	$\langle \delta_p \rangle$ (°)	$\langle \alpha \rangle$ (°)
0.01	3.0	1.97	7.1	15.7	4.9	7.8	33.9	33.5
0.05	1.34	4.33	4.5	13.4	3.4	6.6	28.3	29.4
0.1	0.95	6.24	3.6	12.3	2.8	5.9	25.8	27.5
0.5	0.42	26.6	2.3	10.8	1.9	5.3	19.4	21.0

Table 2.1: Debye-Hückel parameters (the Debye length λ_D and the effective linear charge density ν of the double helix [76]) and the average values of the minimized free parameters for $L = 3.6$ μm, $d = 0.42L$ and $f = 2$ pN under various salt concentrations. Comparison of the braid parameters for the straight (Ca=10) and the buckled phase (Ca=60). R_s and $2\pi P_s$ are the radius and the pitch of the straight braid respectively, while δ_p and α are the braid helix angle and the superhelix angle in the plectoneme state respectively.

Chapter 3

Supercoiled double-helix DNA

A statistical-mechanical model for stretched twisted double-helix DNA is presented, where we view DNA as a semiflexible polymer chain or a worm-like chain featuring bending and twisting stiffness. Mechanical twisting of DNA perturbs the double-helix structure and generates torsion. DNA under high torsional stress buckles to helically writhe around itself, forming a plectoneme [Fig. 3.1(a)]. Helical self-wrapping or supercoiling of DNA is associated with a writhe linking number that does not contribute to torque, and as a result, additional DNA twisting in the buckled state increases the plectoneme length without further increasing DNA torsion. Our model describes coexistence of multiple plectoneme domains in long DNA molecules at physiological salt concentrations ($\approx$ 0.1 M Na^+). Each plectoneme domain contains one end loop that is a finite-sized structure and is responsible for a plectoneme-domain nucleation energy [Fig. 3.1(a)]. We find higher (lower) number of domains at lower (higher) ionic strengths of the solution and stretching forces, in accord with experimental observations.

The model is then used to study the effect of an *immobile point defect* on the DNA contour that allows a localized kink. DNA containing an inhomogeneity or a defect may be treated as an isotropic worm-like chain with a spatially-pinned defect. The thermal persistence length associated with the defect is smaller, which reduces the energy cost of a localized bend or a kink. The defect, however, suppresses diffusion of the bent structure which decreases its entropy. The

degree of the kink is controlled by the defect size, such that a larger defect further reduces the bending energy of the *defect-facilitated kinked end loop* [Fig. 3.1(b)]. We find that a defect can spatially pin a plectoneme domain via nucleation of a kinked end loop, in accord with experiments and simulations [48, 90, 91]. Our model is in accord with magnetic tweezer experiments [48] showing two buckling signatures: buckling and 'rebuckling' in supercoiled DNA with a base-unpaired region. Comparing with experiments, we find that under 1 pN force, a kinked end loop nucleated at a base-mismatched site reduces the bending energy by $\approx 0.7\ k_BT$ per unpaired base. We predict coexistence of three states at the buckling and rebuckling transitions that warrants new experiments. The findings of this chapter are published in Refs. [48, 49]

Significance Twisted DNA is abundant in the cell. While binding of various proteins perturbs the double-helix structure, resulting in DNA torsion, this torque is typically small enough not to induce DNA buckling. Large torsion may be generated by proteins via swiveling one DNA strand around another. Sufficiently twisted DNA undergoes buckling to form plectoneme structures. Bacterial DNA is maintained in a twisted condition, as a result, plectonemic or supercoiled DNA is a common occurrence. DNA supercoiling is associated with various cellular functions, such as genome organization, gene expression, and DNA recombination [53, 92–94]. *In vivo* manipulation of DNA linking number, carried out by topoisomerase enzymes [95], is essential for topological simplification of the entangled state of the genome, as well as manipulation of the compaction state of the genome.

The double helix is not a homogeneous polymer. Cellular DNA harbors inhomogeneities that allow a sharp localized bend, which may be associated with diverse biological phenomena, such as the appearance of a protein-induced kink on the DNA, or a single-stranded bulge, or a DNA hairpin [84, 96, 97]. Intrinsic mechanical inhomogeneity introduced by the base-pair sequence is typically small and may be ignored for a random occurrence of base pairs. However, DNA fragments containing certain periodic base pair arrays (e.g., a positioning sequence, such as the 601 sequence [98]) may have an intrinsic curvature in the double-helix backbone, which can alter its mechanics. Positioning sequences are thought to regulate cellular function via biasing

nucleosome positioning and altering genome access for transcription factors and other DNA-binding proteins [99–101]. Base-pair mismatches are another common source of inhomogeneities influencing DNA mechanics [48, 91]. Spatial pinning of plectoneme domains by a defect [Fig. 3.1(b)] may have important biological consequences. Protein-induced kinks can favor plectoneme branching and play a role in organization of the bacterial genome. Pinning of a plectoneme domain by a base-mismatch site will place the mismatch site at the tip of a plectoneme and may aid in its detection and repair [48].

The importance of theoretical models of DNA and DNA defects of inhomogeneities lies both in gaining a better understanding of biological mechanisms, and in predicting and explaining outcomes of single-molecule experiments [49, 60, 65, 66, 76, 102–104]. Single molecule experiments have been crucial in studying DNA mechanical response to linking number perturbation. While some experiments have directly visualized plectonemes using electron microscopy [105] and DNAs with fluorescent labels [89]; others have studied plectonemic buckling using tweezer techniques [35, 79, 83, 84, 89, 106, 107]. Previous theoretical work have greatly enhanced our understanding of DNA supercoiling, but there still are poorly understood aspects, such as whether DNA buckles to form a single or multiple plectonemic domains that we investigate in this chapter. Experiments [48, 91] and simulations [90] studying supercoiling DNA with unpaired bases have shown that a defect can spatially pin a plectoneme domain, however, the nature of buckling facilitated by the defect and the buckled states involved in the process are not known beforehand, which is also something our model sheds light on.

Outline Sec. 3.1 contains theory [Sec. 3.1.A] and numerical results [Sec. 3.1.B] for supercoiled defect-free DNA, where we also compare with available experimental data. We analyze the effects of salt concentration and length of the supercoiled DNA molecules on its statistical mechanics [Secs. 3.1.B.i and 3.1.B.vii]. Our results also explain the abrupt buckling transition and coexistence of multiple buckled domains in supercoiled DNA.

Sec. 3.2 describes how the model can take into account an immobile point defect on the DNA [Sec. 3.2.A], and explains the results of its numerical solution [Sec. 3.2.B]. Our results reproduce

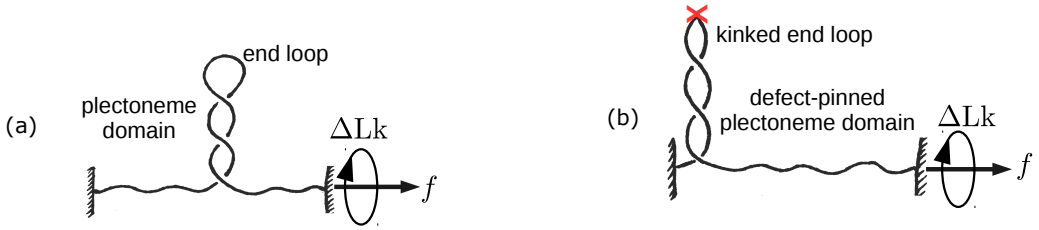

Figure 3.1: (a) Schematic of stretched defect-free double-helix DNA plectonemically buckled under torsional stress. A plectoneme domain contains an end loop that is associated with a nucleation cost of the domain. (b) DNA with a defect located on its contour (denoted by an 'X'). The defect allows a localized DNA kink that favors nucleation of a *defect-pinned* plectoneme domain characterized by an energy-saving kinked end loop. However, the immobile nature of the defect spatially pins the domain costing diffusion entropy.

the second buckling signature or rebuckling transition observed previously in magnetic tweezers experiments [48] [Sec. 3.2.B], and explains the free energy picture underlying buckling [Sec. 3.2.B.i] and rebuckling transitions [Sec. 3.2.B.vi]. We highlight the role of the size of the defect and predict coexistence of multiple states at the transitions [Sec. 3.2.B.xv]. In Sec. 3.2.C, we compare our theoretical results with experimental data of Ref. [48]. We explain the observed shift in transition points [Sec. 3.2.C.i] and quantitatively connect our theoretical defect size parameter with experiments [Sec. 3.2.C.iv]. The experimental trends for extension jump with varying defect sizes [Sec. 3.2.C.v], and the force and salt dependence of the rebuckling signal [Sec. 3.2.C.viii] are also in accord with the theoretical results. Finally, in Sec. 3.3, we conclude with a summary.

3.1 Supercoiled defect-free DNA

We consider DNA as a charged semiflexible polyelectrolyte with bending persistence length $A = 50$ nm and twist persistence length $C = 95$ nm. We define $\beta^{-1} \equiv k_{\rm B}T$ and use $T = 290$ K for all numerical purposes. In the following, we investigate the mechanical response of a double-helix DNA stretched under a constant force f and subject to twist such that the change in the DNA linking number is ΔLk. We present the model schematically in this section; mathematical details are in Appendices C.1 and C.2.

3.1.A Model

We partition total DNA length (L) and linking number (ΔLk) into a force-extended or unbuckled state (L_u, Lk$_u$), a plectonemically-buckled superhelical state (L_p, Lk$_p$), and m end loops each of size γ corresponding to m plectoneme domains.

$$\Delta\mathrm{Lk} = \mathrm{Lk}_u + \mathrm{Lk}_p + m \tag{3.1}$$

We consider the writhe linking number contribution from each end loop to be ≈ 1. There is also a constraint of fixed total DNA length: $L = L_u + L_p + m\gamma$.

The total free energy of a stretched-twisted DNA is written as follows.

$$\mathcal{F} = E_u + E_p - k_\mathrm{B}T \ln Z \tag{3.2}$$

where E_u corresponds to the force-extended or unbuckled part of the DNA containing contributions from DNA twist and force extension [Eq. (C.13)]. E_p is the mean-field energy corresponding to the plectonemically-buckled state [Eq. (C.11)]; and $-k_\mathrm{B}T \ln Z$ is the free energy contribution from thermal fluctuations of the DNA [Eq. (C.9)]. The total free energy is minimized with respect to partition of linking number to obtain the equilibrium linking numbers corresponding to the force-extended and plectoneme states.

3.1.A.i Buckled state

3.1.A.ii Plectoneme superhelix

The plectoneme state is characterized by superhelically-coiled DNA, and has no force-extension energy [Fig. 3.1(a)]. However, transverse fluctuations of the DNA within the superhelical structure are controlled by both the applied tension and electrostatic interactions (Appendix C.1). The plectoneme state costs bending energy, but at the same time reduces twist energy due to the writhe linking number contribution associated with the superhelical structure. Following White's theorem of partition of linking number into twist and writhe [82], the linking number

in the plectoneme state Lk_p is divided as follows.

$$\mathrm{Lk}_p = \mathrm{Tw}_p + L_p \frac{\sin 2\alpha}{4\pi r} \tag{3.3}$$

where the first right-hand-side term is the twist linking number contribution, and the second is the total writhe of a helical structure with opening angle α, helical radius r, and total plectoneme length L_p [82, 102].

3.1.A.iii Plectoneme end loop

The end loop is a finite-sized DNA structure where the double helix bends back in a plectoneme [Fig. 3.1(a)]. We compute the equilibrium size of the end loop: $\gamma = \sqrt{\rho A/(\beta f)}$, via separately minimizing the associated elastic energy expense [46, 65]:

$$\beta E_\gamma = \rho A/\gamma + \beta f \gamma \quad \Rightarrow \beta E_\gamma = \sqrt{\rho \beta A f} \tag{3.4}$$

The parameter ρ depends on the geometry of the loop, such that it is $2\pi^2$ for a circular loop, ≈ 14.0 for a "teardrop"-shaped loop with free ends [108], and ≈ 15.3 (exact) or ≈ 15.7 (simpler calculation) for an end-constrained teardrop loop [86, 87]. We use $\rho = 16$, however, we note that a small change ($\approx 10\%$) in the numerical value of ρ does not alter our conclusions. In the later part of this chapter, concerning defects, we analyze the effect of a relative variation in ρ [Eq. (3.8)].

3.1.A.iv Thermal fluctuations

The mean-field structure of the force-extended state is a twisted straight line, whereas, that of the plectoneme state is a regular helix made up of self-writhed twisted DNA. At a finite temperature, we treat thermal fluctuations as a small perturbation around the mean-field structure (Appendix C.1).

The total energy associated with transverse fluctuations of the DNA about its mean-field

shape is as follows [Eq. (C.14)].

$$-\ln Z = \frac{L_p}{2A}\left[\frac{3}{2}\sqrt{\mu} + \eta^{1/4}\cos\left(\frac{1}{2}\tan^{-1}\sqrt{\frac{4\eta}{\mu^2}-1}\right)\right] + (L-L_p)\sqrt{\beta f/A} \tag{3.5}$$

where $\mu \equiv \beta A f \cos\alpha$, is the effective tension in each of the two helically wrapped strands of the plectoneme; and $\eta \equiv (A^2/2)\partial_r^2\mathcal{U}$, is the effective electrostatic modulus of uniform radial deformations in the plectonemic superhelix (see Appendix C.1).

The first bracketed term in Eq. (3.5) corresponds to thermal fluctuations in the plectoneme structure [Eq. (C.9)]. Note that there are four degrees of freedom associated with transverse fluctuations in a plectoneme structure, two for each of the plectonemic strands. In a conveniently chosen reference frame [Eq. (C.3)], three of the degrees of freedom fluctuate independently under the external tension, as seen in the first term inside the brackets (the term proportional to $\sqrt{\mu}$). The second term within the brackets, dependent on the strength of the electrostatic repulsions via η, corresponds to electrostatically-coupled fluctuations where the two strands displace relative to each other. The second term in Eq. (3.5) corresponds to tension-controlled transverse fluctuations in the force-extended part of the DNA.

We note that this part of the computation substantially improves on the prior work where a scaling-like free energy cost from cylindrical confinement is used to account for fluctuations in the plectonemic superhelix [12, 66, 76, 80, 81]. Our approach, resulting in Eq. (3.5), proposes an explicit computation of thermal fluctuations, treating them as a perturbation about the mean-field helical geometry of the plectoneme. Eq. (3.5) is consistent with the previously assumed scaling: $-\ln Z \sim \sigma_r^{-2/3}$ (where σ_r is the fluctuation in plectoneme radii [Eq. (C.10)]), but does not depend on any unknown scaling constants.

Note that a similar treatment of thermal fluctuations in two helically-intertwined DNAs or braids can be found in Chapter 2. DNA braids have the same geometry as plectonemes, however experimentally studied braids [47, 55, 56] are force extended, while plectonemes are buckled structures that do not have force-extension energy [Appendix C.1].

3.1.A.v Extension and torque

DNA extension z can be obtained from the negative-force derivative of the total free energy [Eq. (3.2)]. Hence, DNA lengths in the force-extended and buckled states are respectively associated with positive and zero contributions to extension. Thermal fluctuations further reduce DNA extension, which is a sub-leading order effect.

Torque in the DNA τ is obtained via the linking-number derivative of the total free energy. The free energy is harmonic in twist linking number, however, conversion of twist to writhe in the plectoneme-coexistence state influences the torque response.

3.1.A.vi Partition function

We sum over states containing all possible lengths (L_p), and numbers of domains (m) of plectoneme to construct a canonical partition function $\mathcal{Z}$.

$$\mathcal{Z}(\Delta\mathrm{Lk}, f) = e^{-\beta\mathcal{F}(0,0)} + \sum_{m=1,2,\ldots} \sum_{L_p>0} e^{-\beta\mathcal{F}(L_p,m)} \tag{3.6}$$

where the coexistence-state energies $\mathcal{F}(L_p, m)$ [Eq. (3.2)] are obtained by free-energy minimization, which also ensures balance of torque in the summed-over states. Energy minimization for a coexistence state with fixed L_p and m determines the equilibrium plectoneme radius r, and opening angle α.

Equilibrium values of end-to-end extension (z), number of plectoneme domains (m), torque in the DNA (τ), and the total plectoneme length (L_p) at a fixed force and fixed linking number are obtained from the partition function [Eq. (C.15)].

3.1.A.vii Probability distributions

In a canonical ensemble of fixed force f, and fixed linking number $\Delta\mathrm{Lk}$, both z and τ undergo equilibrium fluctuations. The total probability distribution of $X \in \{z, \tau\}$ is obtained by adding

the contributions from various states in the partition sum:

$$P(X) = \mathcal{P}_0(X) + \sum_{m=1,2,\cdots} \mathcal{P}_m(X) \tag{3.7}$$

where $\mathcal{P}_0$ corresponds to the force-extended state, and $\mathcal{P}_m$, where $m \in \{1, 2, \cdots\}$, is the contribution from the buckled state featuring coexistence of m plectoneme domain(s). Buckled states corresponding to different plectoneme lengths in the partition sum are already taken into account in $\mathcal{P}_m$ [Eq. (C.16)].

3.1.B Results: defect-free DNA

In this section, we describe the numerical solutions of the model and compare our results with experimental data.

3.1.B.i Supercoiling at physiological salt

We begin with a discussion of rather short DNA molecules (≈2 kb), subject to twist under physiological salt conditions (≈ 0.15 M Na^+).

3.1.B.ii Slight twisting causes extension change due to chiral fluctuations

Untwisted DNA extension is 80-90% of its total contour length under 0.5-2 pN stretching force [Fig. 3.2(a)]. Higher forces suppress DNA excursions lateral to the force direction, resulting in a longer extension at zero linking number. Small twisting of the double helix results in a linear buildup of DNA torque [Fig. 3.2(b)]. Change in extension upon slight twisting of the double helix is small, due to twist-induced chiral fluctuations in DNA that leads to partial twist screening, which is taken into account via a renormalized twist stiffness for stretched-unbuckled DNA [Eq. (C.13)] [60].

Linear torque indicates a constant twist stiffness in DNA, which is a consequence of the rigidly-stacked double-helical structure. A softer structure, two-intertwined nicked DNAs or a "braid" has a linking number dependent twist stiffness [46], see Chapter 2.

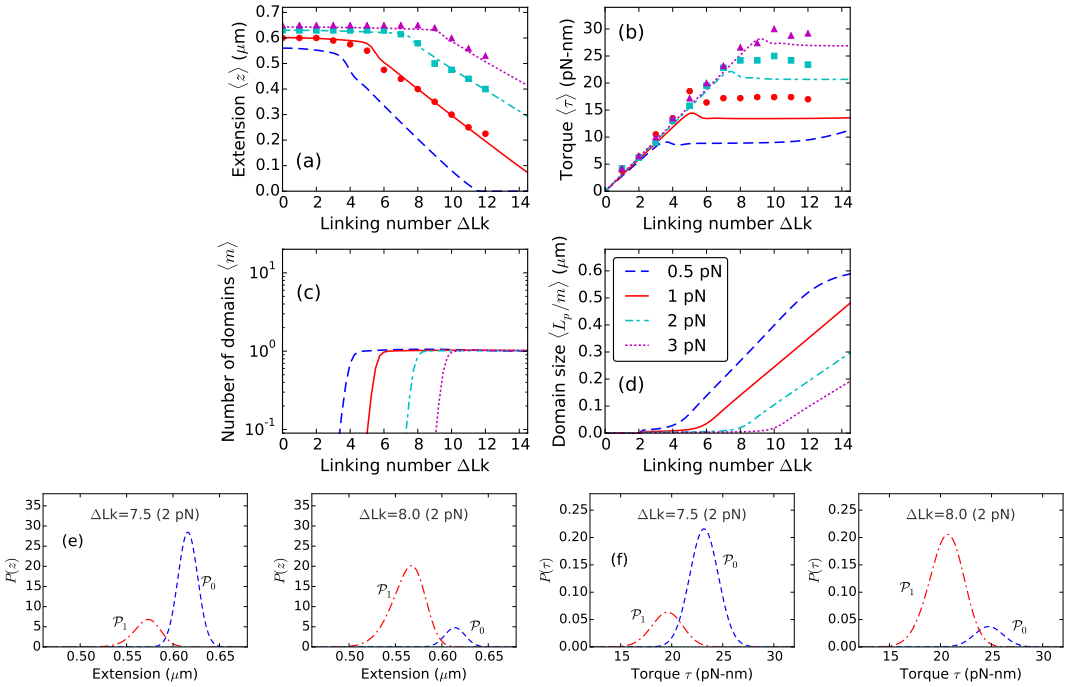

Figure 3.2: Theoretical curves for a 0.7 μm (≈2 kb) supercoiled DNA molecule, stretched under 0.5 (blue dashed lines), 1 (red solid lines), 2 (cyan dot-dashed lines), and 3 pN (magenta dotted lines) applied forces at 0.15 M Na^+. Experimental data, reproduced from Ref. [83], are plotted for 1 (red circles), 2 (cyan squares), and 3 pN (magenta triangles). (a) Extension versus linking number, shows a flat unbuckled regime at lower linking numbers. Extension decreases steeply at higher linking numbers corresponding to coexistence of a plectoneme state. The extension discontinuity connecting the two slopes corresponds to the buckling transition. (b) Torque increases linearly in the unbuckled state, and then saturates as a part of the DNA buckles to form plectoneme. Constant torque in the plectoneme coexistence state is due to the writhe contribution of the plectoneme geometry that screens DNA twist. The small overshoot in torque as well as the discontinuity in extension near the buckling transition point is related to the end loop-introduced nucleation cost of a plectoneme domain. (c) Equilibrium number of plectoneme domains grows to unity at the buckling transition point, showing the nucleation and coexistence of a plectoneme domain. (d) Average plectoneme domain size increases after the buckling point, indicating addition of superhelical turns to the buckled domain. Probability density of (e) extension and (f) torque near the buckling point is bimodal [ΔLk=7.5 and 8.0 under 2 pN force]. The modes of the distributions at higher extension and torque correspond to the unbuckled state ($\mathcal{P}_0$, blue dashed lines); whereas, the lower extension and torque modes correspond to the one-domain plectoneme state ($\mathcal{P}_1$, red dot-dashed lines). The average occupancy of the buckled state ($\mathcal{P}_1$) increases as the linking number is increased near the buckling point.

3.1.B.iii Torque due to higher twisting leads to DNA buckling

At higher linking numbers, the double helix buckles into a self-writhed plectoneme structure [12, 105]. Plectonemes with higher pitch-to-radius ratio have a substantial *writhe* linking number density [Eq. (3.3)], as a result, buckling avails conversion of twist into writhe. Post-buckling torque is nearly constant [Fig. 3.2(b)], suggesting that increasing the linking number in buckled DNA does not increase DNA twist but increases total writhe. Plectonemes save DNA twist energy but cost bending energy, hence buckling is favored only above a critical torque that corresponds to a *critical linking number*. Higher stretching forces stabilize the unbuckled state, resulting in an increase in the critical linking number [Fig. 3.2]. The mean-field plectoneme state does not contribute to end-to-end extension, resulting in a steeper decrease in extension in the coexistence regime [Fig. 3.2(a)].

3.1.B.iv Buckled DNA forms plectonemic structures

Figure 3.2(c) shows the appearance of a plectoneme domain at the buckling transition, which grows in size as the linking number is increased beyond the critical value [Fig. 3.2(d)]. The increase in the average size of the plectoneme domain is due to equilibrium DNA length being passed from the unbuckled into the buckled state, which increases the number of superhelical turns in the plectoneme domain.

3.1.B.v Buckling transition is abrupt due to nucleation of the end loop

The buckling transition is abrupt due to the finite-sized end loop that associates a *nucleation cost* to a plectoneme domain. The discontinuity in extension and the overshoot in torque at the buckling point [Fig. 3.2(a)-(b)] characterizes the abrupt nature of the transition. Near the buckling point, the unbuckled and the plectoneme states are thermally accessible, which implies that the probability of occupancy of either state is non-zero. Figure 3.2(e) shows a bimodal probability density of extension near the buckling point, where the unbuckled and the plectoneme states correspond to the higher ($\mathcal{P}_0$) and lower ($\mathcal{P}_1$) extension modes, respectively.

The discontinuity in extension at the buckling point (i.e., the non-zero distance between the two extension modes) is due to the fact that a buckled domain cannot be smaller than an end loop, which is $\sim O(1)$ DNA persistence length in size. As the linking number is increased near the buckling point, the probability of occupancy of the buckled state increases, and that of the unbuckled state decreases [Fig. 3.2(e)].

The average torques in the two fluctuation-accessible states ($\mathcal{P}_0$ and $\mathcal{P}_1$) at the buckling transition are different; this is due to the writhe associated with the nucleated buckled domain (plectoneme and the end loop). Figure 3.2(f) shows the bimodal torque distributions near the buckling point. Equilibrium fluctuations between the two states of different torques: $\mathcal{P}_0$ and $\mathcal{P}_1$, lead to an overshoot behavior seen in the ensemble-averaged DNA torque at the buckling transition [Fig. 3.2(b)]. A nonmonotonic mechanical torque, i.e., decreasing torque with increasing linking number indicates negative torsional stiffness, a signature of mechanical instability. However, the ensemble-averaged torque may show nonmonotonic behavior in an equilibrated system with monotonic mechanical torque. This is a consequence of nonmonotonic behavior of *torque fluctuations* near the buckling transition (see Sec. V of Ref. [65]). There is experimental evidence of an overshoot in DNA torque at the buckling transition [83, 107].

3.1.B.vi Long buckled DNA show coexistence of multiple plectoneme domains

Entropic stabilization of plectoneme domains, via one-dimensional diffusion along the DNA contour and exchange of DNA length among domains (i.e., fluctuations in relative size of the domains), increases logarithmically with DNA length [Eq. (C.11)] [65]. This leads to proliferation of multiple domains in supercoiled long DNA molecules (> 10 kb). Diffusion of plectonemes and coexistence of multiple domains have been observed in DNA visualization experiments [89, 91].

Figure 3.3 shows buckling behavior in long DNA molecules. The critical linking number is an extensive quantity that increases with DNA length [Figs. 3.2(a) and 3.3(a)]; however, the critical buckling torque is intensive and remains roughly the same for different length molecules [Figs. 3.2(b) and 3.3(b)].

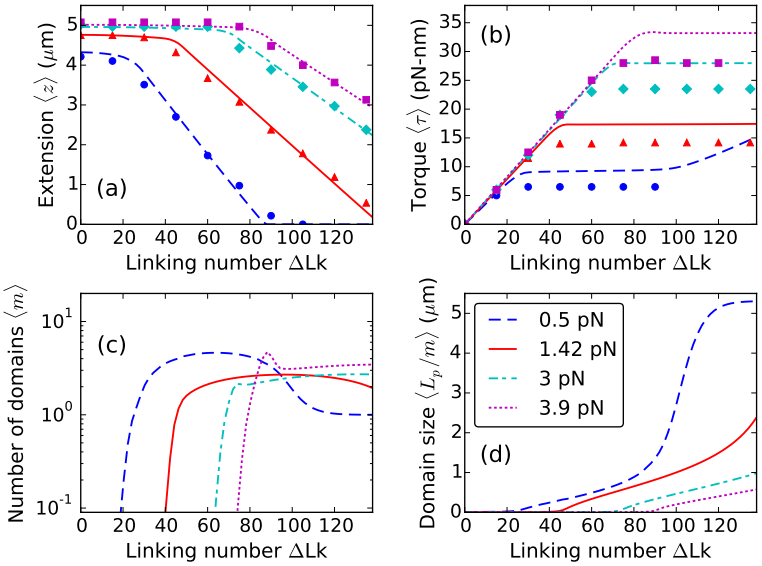

Figure 3.3: Theoretical curves for supercoiled 5.4 μm ($\approx$ 16 kb) DNA, stretched under 0.5 (blue dashed lines), 1.42 (red solid lines), 3 (cyan dot-dashed lines), and 3.9 pN (magenta dotted lines) forces at 0.1 M Na^+. Experimental data for 0.5 (blue circles), 1.42 (red triangles), 3 (cyan diamonds), and 3.9 pN (magenta squares) are reproduced from Ref. [79]. (a) Extension and (b) Torque plotted as a function of linking number show twisting behavior at lower linking numbers and plectoneme buckling at higher linking numbers. (c) Equilibrium number of plectoneme domains show proliferation of multiple plectoneme domains in the coexistence state. At higher forces, long molecules show a non-monotonic increase in the number of plectoneme domains at the buckling transition due to the large entropy associated with plectoneme diffusion. However, in the purely-plectoneme state (i.e., the zero extension state, refer to the 0.5 pN case, at linking numbers > 90), high stability of plectoneme superhelices and absence of diffusion entropy results in favoring a single plectoneme domain. Torque in the purely-plectoneme state increases because the DNA twist increases. (d) The steepness in the increase of the average domain size increases in the purely-plectoneme state due to coalescence of plectoneme domains.

The larger configuration entropy associated with long DNA molecules reduces the nucleation energy of a plectoneme domain, resulting in proliferation of multiple domains in the buckled state [Fig. 3.3(c)]. At lower forces, the nucleation energy cost is further reduced, resulting in an increased tendency to proliferate new plectoneme domains. However, in the *purely-plectoneme state* (i.e., the zero-extension state, where the entire DNA is in the plectoneme state, $L_u = 0$), a single plectoneme domain is favored due to reduced diffusion entropy of a domain. Figure 3.3(c) shows coalescence of multiple domains as the linking number is increased in the purely-

plectoneme state, which is the result of a highly stable plectoneme superhelix compared to an end loop.

Energy of the unbuckled and the plectoneme states vary linearly with force, whereas, that of the end loops vary as the square root of the applied stretching force. This leads to an increased probability of small plectoneme domains (i.e., end loop with a minimal amount of plectoneme superhelix) at the buckling transition under higher forces. As a result, twisted long DNA, under higher forces, buckles via nucleation of a few small loops that coalesce at a slightly higher linking number due to increased stability of plectonemic superhelices. This is seen as a small overshoot in the number of plectoneme domains at the buckling transition under higher forces [Fig. 3.3(c)].

Post-buckling torque is mostly constant in the plectoneme-coexistence state, however, increasing the linking number in the purely-plectoneme state causes an increase in the DNA twist, reflected in an increase in the torque [Fig. 3.3(b)].

Theoretical post-buckling torque appears to be an underestimation compared to experiments of Ref. [83] [Fig. 3.2(b)], where torque was inferred from angular fluctuations of an optically-trapped DNA-tethering particle. On the other hand, torque reported in Ref. [79], derived from the slope of the extension curves using Maxwell relations [109] is lower than the theoretical values [Fig. 3.3(b)]. While in-situ torque measurement is a remarkable step forward for experiments, equilibration might be an issue. Maxwell relations are a robust tool for torque estimation, however, the procedure employed by Ref. [79] assumes a constant torque in the plectoneme-coexistence state. Our model, although devoid of the above issues, assumes a regular plectoneme geometry and ignores any energy contribution from distortion of the helical plectoneme structure. Numerical values of the DNA effective charge [63, 76] as a function of salt concentration also affects the coexistence state torque. These small discrepancies call for future attention both from the theoretical and experimental perspectives.

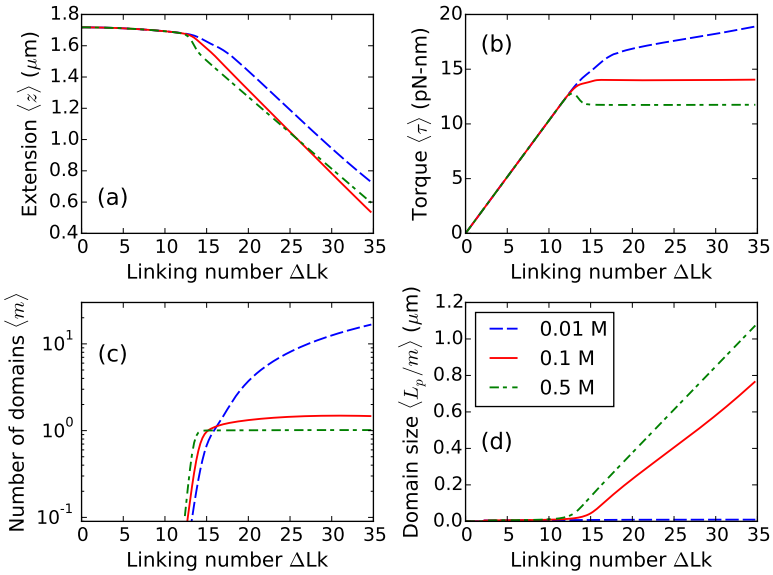

Figure 3.4: Effect of salt concentration on defect-free DNA. Supercoiled 2μm DNA at 1 pN stretching force under 0.01 (blue dashed line), 0.1 (red solid line), and 0.5 M Na^+ (green dot-dashed line). (a) Extension, and (b) Torque shows a more rounded buckling transition for lower salts due to a lower stability of plectoneme superhelices. Note that the torque increases in the buckled state for lower salts, due to an increase in DNA twist resulting from lower twist screening by plectoneme superhelices. (c) Number of plectoneme domains increase in the buckled state for lower salts, whereas, a single plectoneme domain is favored at higher salts. (d) Average size of a domain increases in the higher salt case. For lower salts, proliferation of multiple domains lead to a very small domain size in the buckled state.

3.1.B.vii Effect of salt concentration

An increased ionic concentration of the solution strengthens the electrostatic screening of the charges on the DNA backbone, which reduces the effective excluded diameter of DNA (measured in Debye-Hückel screening length λ_D). At lower salt concentrations, a larger screening length mimics stronger self-avoidance in DNA, which shifts the free energy balance in favor of looped structures over plectonemic superhelices. This effectively translates into a higher tendency to proliferate multiple domains of plectoneme as the salt concentration is lowered.

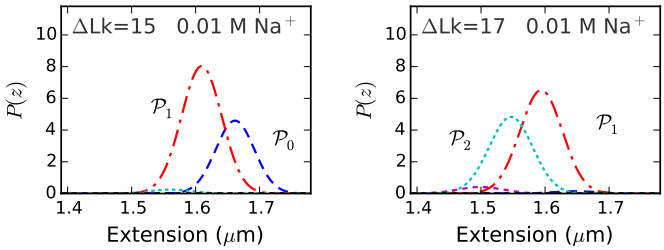

Figure 3.5: Probability density of extension at 0.01 M Na^+ under 1 pN force near the buckling transition for a 2 μm DNA ($\approx$ 6 kb) [Fig. 3.4]. The extension distribution is bimodal, however, the two modes, corresponding to unbuckled DNA ($\mathcal{P}_0$, blue dashed line) and one plectoneme domain ($\mathcal{P}_1$, red dot-dashed line), are less resolved at lower salts due to increased fluctuations. The probability distribution remains bimodal after the buckling point, due to appearance of multidomain plectoneme states (e.g., the two domain state $\mathcal{P}_2$, cyan dotted line at ΔLk=17). Decreased stability of the plectoneme superhelix at lower salts result in coexistence of multiple plectoneme domains. Increasing the linking number in the buckled state increases the probability of occupancy of a plectoneme state with a larger number of domains.

3.1.B.viii Buckling transition is less sharp at lower salt concentrations

The critical linking number increases with a decrease in the salt concentration (Fig. 3.4). Larger excluded diameter of DNA at lower salt concentrations increases the bending energy of the plectoneme state, leading to an increase in the critical linking number. The post-buckling state is that of many domains at lower salts ($\approx$ 0.01 M Na^+) (Fig. 3.4), which is again related to the increased energy of plectoneme superhelices that favors nucleation of many small plectonemes instead of a single large domain.

3.1.B.ix Buckled DNA at lower salts show coexistence of multiple domains

Torque in the buckled state is flat for higher salts and increases slowly with linking number for lower salt concentrations (Fig. 3.4). For lower salts, decreased stability of the plectonemic superhelix (due to higher DNA excluded diameter) causes a small increase in DNA twist in the plectoneme coexistence state. The extension distributions, at low salt concentrations ($\approx$ 0.01 M Na^+), are bimodal at the buckling transition and remain multimodal in the buckled state due to coexistence of multiple plectoneme domains (Fig. 3.5), which may be visible in tweezer experiments. However, increased fluctuations may decrease the resolution of the two peaks,

producing a broad-tailed single Gaussian shape.

DNA braids are structurally bulky, and as a result, mimic the low salt behavior of supercoiled DNAs. Braids show multimodal extension profiles corresponding to proliferation of buckled domains [46, 47], see Chapter 2.

We assume the unbuckled state to be decoupled from the electrostatics, because non-neighbor segments in unbuckled DNA are always distant. Nonetheless, increased repulsion between neighboring segments at lower salts (due to less DNA charge screening), is expected to induce additional stretching of the polymer. This maybe taken into account via a persistence length that gets longer with decreasing salt concentration. Experiments suggest a small change ($\approx 10\%$) in the persistence length over a wide range of salt concentrations (0.01-1 M Na^+) [78], which we ignore for simplicity. The ionic strength dependence of DNA torsional stiffness is also negligibly small [110].

3.1.B.x Abruptness of the buckling transition varies due to a differential stability between plectoneme superhelices and end loops

The discontinuous change in extension and overshoot in torque at the buckling point are measures of the abruptness of the transition, which decreases with decreasing salt concentration (Fig. 3.4). A more abrupt transition, i.e., two well-separated peaks in the extension profile, results from a larger size of the nucleated buckled domain. The nucleated domain at buckling consists of an end loop and a plectoneme comprising superhelical turns. When the superhelix is less stable, the case for lower salt concentrations (≈ 0.01 M Na^+), the nucleated domain is an end loop with minimal superhelix, which leads to a less abrupt extension change at buckling.

When the salt concentration is increased, plectoneme-superhelical turns are increasingly stabilized, and the amount of superhelically wrapped DNA in the nucleated domain also increases. This produces a larger extension change (i.e., strongly bimodal extension distribution) at buckling for higher ionic strengths.

3.2 DNA with a spatially-pinned point defect

In the following section, we describe an immobilized point defect on the DNA that can spatially-pin a kinked end loop. Consequently, nucleation of a spatially-pinned plectoneme domain may be favored at the defect site due to the higher bending energy of a teardrop end loop compared to a kinked loop. We introduce a defect size parameter ε that controls the relative stability of a kinked loop. We also predict a defect-size-dependent coexistence of three states at buckling and rebuckling transitions.

3.2.A Model for an immobile defect

We suppose that a defect acts as a soft-spot for bending deformations, such that the DNA can form a *kink* at the defect site [Fig. 3.1(b)]. We are motivated to describe defects arising from base-unpaired regions on the DNA [48], but the biological relevance of a defect-induced DNA kink is diverse, such as a protein-mediated DNA bend, or a single-stranded bulge on the DNA, or a damaged DNA base.

3.2.A.i Defect-pinned plectoneme domain

A plectoneme with its tip at the defect site can have a kinked end loop, thereby saving bending energy. However, the immobile nature of the defect restricts diffusion of a defect-pinned domain [Fig. 3.1(b)]. As a result, nucleation of a defect-pinned domain is expected to feature a competition between stabilization from lower bending energy of a kinked end loop and destabilization from spatial pinning.

3.2.A.ii Size of the defect ε

The defect may have a varied size that affects the degree of the defect-induced DNA kink; larger defects allow a sharper DNA kink, thus lowering the associated bending energy of a kinked end loop [87]. Following the scheme used for the defect-free DNA in Eq. (3.4), we use a defect-size dependent loop parameter: $(1-\varepsilon)^2\rho$, such that the energy of a kinked end loop ($E_{\gamma^\dagger}$) is a

defect-dependent fraction of that of the teardrop loop:

$$E_{\gamma^\dagger} = (1-\varepsilon)E_\gamma, \tag{3.8}$$

where $0 < \varepsilon < 1$ is the size of the defect, and $\beta E_\gamma = \sqrt{\rho\beta A f}$ is the energy of a teardrop-shaped loop [Eq. (3.4)]. Energy minimization gives the equilibrium size of a defect-kinked end loop:

$$\gamma^\dagger = (1-\varepsilon)\sqrt{\rho A/(\beta f)}. \tag{3.9}$$

Higher values of ε, corresponding to a larger defect, stabilizes the defect-pinned loop by allowing a sharper kink at the defect site. The experimental counterpart of ε is the number of adjacent base-pair mismatches on a supercoiled DNA [48]. A defect with a larger number of unpaired bases further reduces the bending energy of a defect-kinked loop, which corresponds to a larger value of ε.

The total free energy of a stretched twisted DNA with a defect of size ε is given as:

$$\mathcal{F}(L_p, m, m^\dagger) = E_u + E_p(\varepsilon) - k_\mathrm{B}T \ln Z, \tag{3.10}$$

where the free energy of the coexistence state now depends on the length of the plectoneme (L_p), the number of mobile plectoneme domains (m), and the number of defect-pinned plectoneme domains ($m^\dagger$) which can be either 0 or 1 (i.e., $m^\dagger \in \{0, 1\}$). The size of the defect, ε affects the plectoneme state energy by changing the bending contribution associated with a kinked loop [Eqs. (3.8), (C.17)]. The total free energy is minimized for a fixed total length:

$$L = L_u + L_p + m\gamma + m^\dagger\gamma^\dagger, \tag{3.11}$$

where γ and $\gamma^\dagger$ are the respective sizes of a teardrop and a kinked end loop. [Eqs. (3.4) and (3.9)]. The total linking number is also constrained, where both kinked and teardrop end loops contribute unit writhe linking number.

3.2.A.iii Critical size of the defect-pinned plectoneme

The size of a plectoneme domain is maximum at the end of the coexistence region, which is the purely-plectoneme state with no unbuckled DNA. However, a defect-pinned plectoneme can be made maximally big (corresponding to a *critical size*) at any point in the coexistence region by forcing its proximity to the tethering surface. For a defect located a distance L^* from one of the ends, the defect-pinned plectoneme has a critical size of $2L^*$ [48]. A defect-pinned domain must have the tip of the plectoneme at the defect site; this results in one of the ends of a critically big defect-pinned plectoneme domain coinciding with a surface-tether point of the DNA [Fig. 3.1(b)]. Consequently, at a coexistence point with total plectoneme length larger than $2L^*$ there must be at least one mobile plectoneme domain (i.e., $m \geq 1$).

3.2.A.iv Partition function

We sum over all possible sizes (L_p) and numbers of the plectoneme domains, both mobile (m) and pinned ($m^\dagger$), to construct a canonical partition function:

$$\mathcal{Z}^\dagger = e^{-\beta\mathcal{F}(0,0,0)} + \sum_{m^\dagger=0,1} \sum_{m=0,1\ldots} \sum_{L_p=\Delta}^{L} \left[e^{-\beta\mathcal{F}(L_p,m,m^\dagger)} \times (1-\delta_{m,0}\delta_{m^\dagger,0}) \left[1-\delta_{m,0}\Theta(L_p-2L^*)\right] \right] \tag{3.12}$$

The above partition function imposes a critical size on the defect-pinned plectoneme via the product of the Kronecker delta with the Theta function [Eq. (C.18)]. The product of the two Kronecker delta functions ensures the presence of at least one end loop in the buckled state of the DNA. We get equilibrium values of observables such as extension, torque, and number of plectoneme domains from the above partition function [Eqs. (C.15)].

3.2.A.v Extension distribution

The extension profile at a given linking number and force is also obtained from the partition function [Eq. (C.19)].

$$P(z) = \mathcal{P}_{00} + \sum_{m^\dagger,m}{}' \mathcal{P}_{m^\dagger m}, \tag{3.13}$$

where the primed sum corresponds to a restricted sum as defined in the above partition function [Eq. (3.12)]. Here, $\mathcal{P}_{00}$ is the contribution from the unbuckled state; and $\mathcal{P}_{m^\dagger m}$ is the contribution from a buckled state with $m^\dagger$ pinned and m mobile plectoneme domains. For instance, $\mathcal{P}_{01}$, $\mathcal{P}_{10}$, and $\mathcal{P}_{11}$ are the respective contributions from- the buckled state with one mobile plectoneme domain ($m^\dagger = 0$ and $m = 1$), the buckled state with a defect-pinned domain ($m^\dagger = 1$ and $m = 0$), and the two-domain plectoneme state containing one mobile and one defect-pinned domain ($m^\dagger = 1$ and $m = 1$).

3.2.B Results: DNA with a defect

Figure 3.6(a)-(b) shows extensions and torques, respectively, of a 6 kb DNA molecule with a defect of sizes $\varepsilon = 0.05, 0.15$, and 0.3 as a function of linking number. Small twisting of the molecule leads to a small extension change and a linear torque buildup, as seen for defect-free DNA (Fig. 3.2). An increase in the DNA linking number leads to buckling of the double helix. The second buckling signature in the extension and torque curves [Fig. 3.6] or the *rebuckling transition* [48] is related to the critical size associated with the defect-pinned plectoneme domain.

3.2.B.i Buckling transition

At the buckling transition, the plectoneme state becomes energetically favored because of its writhe linking number contribution that decreases DNA-twist energy and torque.

3.2.B.ii Defect size controls the post-buckling state

The plectoneme domain nucleated at the buckling transition can be defect-pinned ($\mathcal{P}_{10}$), which has a nucleation energy that decreases with increasing defect size, but the spatially-pinned nature of the domain costs diffusion entropy [Fig. 3.7(a)]. On the other hand, the nucleated

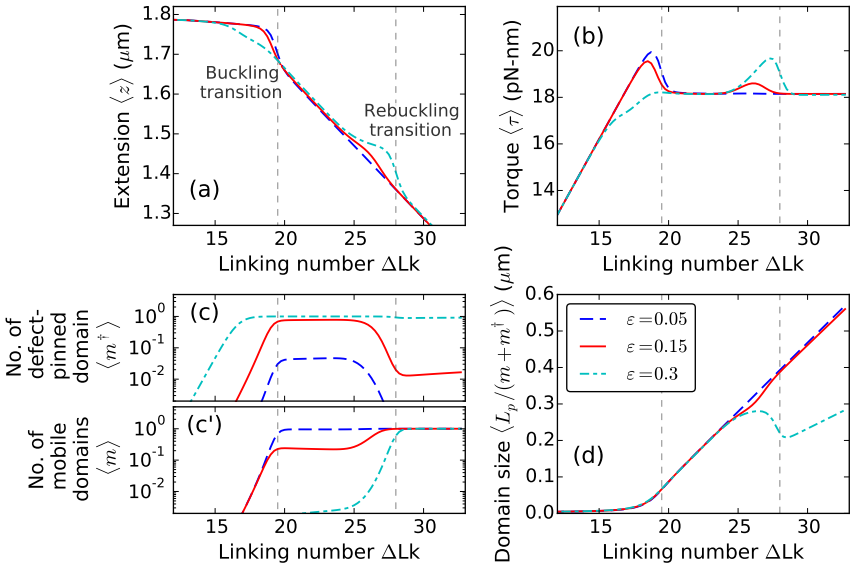

Figure 3.6: Supercoiling 2μm DNA ($\approx$ 6 kb) with a defect of size $\varepsilon = 0.05$ (blue dashed lines), 0.15 (red solid lines), and 0.3 (cyan dot-dashed lines) located L^*=150 nm ($\approx$ 450 bp) from the surface, under 2 pN stretching force and 0.5 M Na^+. The buckling transition is associated with nucleation of a plectoneme domain, whereas, the rebuckling transition is due to a maximum-size constraint on the defect-pinned plectoneme domain [48]. (a) Extension and (b) Torque versus linking number curves show, respectively, a sharp decrease and an overshoot at the buckling and rebuckling transition points. The magnitudes of torque overshoot and extension jump, associated with the nucleation cost at the transition, decrease with increasing defect size ε for the buckling transition; whereas, at the rebuckling point, they increase with increasing size of the defect. (c) Equilibrium number of pinned plectoneme domain shows an appearance of the defect-pinned plectoneme at the buckling point, however, probability of nucleating the defect-pinned domain is vanishingly small for *small defects* ($\varepsilon < 0.1$). Near the rebuckling point, the defect-pinned domain is stable only for *large defects* ($\varepsilon > 0.25$). (c$'$) Equilibrium number of mobile plectoneme domains shows that a mobile domain is favored at the buckling point only when the defect is small; for larger defects, a mobile domain does not appear before the rebuckling point. This suggests that the rebuckling transition does not occur for small defects. (d) Average size of a plectoneme domain shows an increase after the buckling point. Rebuckling transition occurs when the size of the defect-pinned domain is $2L^*$ or 0.3 μm. Near the rebuckling point, for larger defects, the average size of a domain shows an abrupt decrease due to nucleation of a mobile plectoneme domain. The vertical dashed lines correspond to buckling and rebuckling transitions [see Figs. 3.7 and 3.8], however, note that the critical linking numbers for the transitions are defect size dependent [Fig. 3.10].

plectoneme domain can be a mobile one ($\mathcal{P}_{01}$), which has a fixed nucleation cost associated with a teardrop loop and possesses extra stabilization from diffusion entropy. Figure 3.7(b) shows

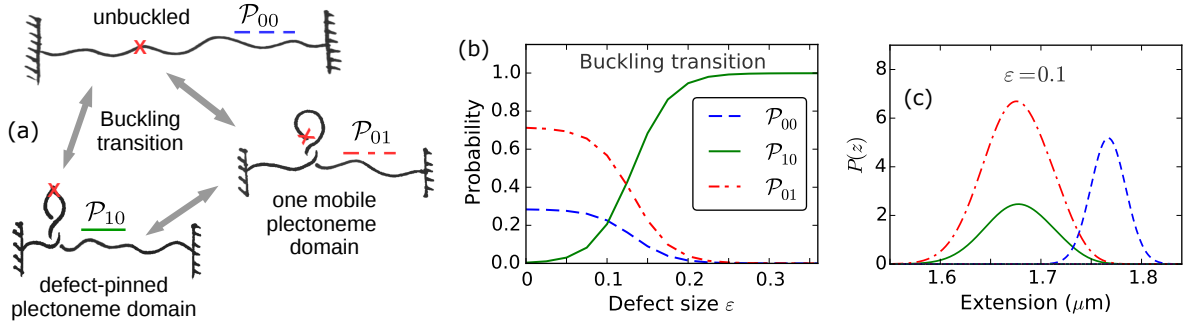

Figure 3.7: Buckling transition for various defect sizes. (a) Schematic of the three possible states at the buckling transition: unbuckled state ($\mathcal{P}_{00}$, blue dashed line), defect-pinned plectoneme ($\mathcal{P}_{10}$, green solid line), and mobile plectoneme domain ($\mathcal{P}_{01}$, red dot-dashed line). (b) Total probability of the three states at the buckling transition (ΔLk=19.5 under $f = 2$ pN and 0.5 M Na^+, see Fig. 3.6) as a function of the defect size ε. For larger defects ($\varepsilon > 0.1$), the defect-pinned domain ($\mathcal{P}_{10}$) is the favored post-buckling state, because of the lower bending energy of a kinked end loop associated with $\mathcal{P}_{10}$. While, for *small defects* ($\varepsilon < 0.1$) the bending energy saved from a kinked end loop is lower than the loss of diffusion entropy of the pinned state ($\mathcal{P}_{10}$), which makes the mobile domain ($\mathcal{P}_{01}$) the favored post-buckling state. Note the relatively higher probability of the unbuckled state for smaller defects. This is due to a shift of the buckling point towards lower linking numbers with increasing defect sizes (Fig. 3.10). (c) Probability density of DNA extension at the buckling transition shows the typical bimodal character observed for defect-free DNA (Fig. 3.2), however, the defect size controls the states populating the lower-extension mode of the distribution. This also suggests that measurement of the extension alone is insufficient to distinguish between the states involved at the buckling transition.

the probability of the three states: $\mathcal{P}_{00}$, $\mathcal{P}_{01}$, and $\mathcal{P}_{10}$, near the buckling transition as a function of the defect size. The defect-pinned domain ($\mathcal{P}_{10}$) is the favored post-buckling state for larger defects ($\varepsilon > 0.1$), whereas a mobile domain ($\mathcal{P}_{01}$) is the preferred post-buckling state for smaller defect sizes ($\varepsilon < 0.1$).

3.2.B.iii Small defects cannot reliably pin plectoneme domains

For small defects ($0 < \varepsilon < 0.1$), the loss of diffusion entropy for a defect-pinned domain is higher than the bending energy saved from a kinked end loop. This leads to relative stabilization of the mobile domain for small defect sizes. The probability of the mobile plectoneme domain decreases near the buckling transition with increasing defect size, and becomes negligible for

larger defects [Figs. 3.6(c) and 3.7(b)]. The probability of the defect-pinned plectoneme domain increases with the defect size, and the pinned domain becomes the most probable post-buckling state for larger defects. The total probability of an end loop or a plectoneme domain after the buckling transition is unity; the defect size controls the relative stability of the two possible types of end loops: mobile and kinked or defect-pinned, thus controlling the post-buckled state [Fig. 3.7(b)].

3.2.B.iv Larger defects cause buckling at a lower critical linking number

Buckling occurs at a lower linking number for higher values of ε. Figure 3.6(c) shows an increase in the probability of the defect-pinned plectoneme domain at the buckling point. The nucleation cost of a defect-pinned plectoneme domain decreases with increasing defect size, due to the lower bending energy associated with a kinked end loop of a larger defect [Eq. (3.8)]. This results in a decrease of the critical linking number for larger defects, as well as a smaller extension change and torque overshoot at the buckling transition [Fig. 3.6(a)-(b)].

However, small defects do not show a shift in the buckling transition, because the nucleation cost of the most probable post-buckling state ($\mathcal{P}_{01}$) is independent of the defect size.

3.2.B.v Plectoneme contribution to a nucleated buckled domain is higher for higher critical linking number

The average size of a plectoneme domain increases with increasing linking number due to nucleation and consequent growth of a plectoneme domain. Note that although the critical buckling point varies with the defect size, the plectoneme domain size near the buckling point does not depend on the defect [Fig. 3.6(d)]. DNA length contribution to a nucleated domain from plectoneme superhelical turns depends only on the linking number. As a result, the critical linking number at a transition determines the plectoneme contribution to the total size of the nucleated buckled domain, such that the contribution is larger for a higher critical linking number, i.e., a smaller defect. This also means that the extension change at the buckling transition is larger

for smaller defects [Fig. 3.6].

For small defects ($\varepsilon < 0.1$), the domain size increases after nucleation of the mobile end loop. However, for larger defects, the defect-facilitated kinked end loop is highly stable and becomes probable before plectoneme superhelices are favored [Fig. 3.6(a)-(c)]. As a result, for larger defects, there is a linking number interval at the buckling transition where the post-buckled state is just the kinked end loop with minimal superhelical turns. This interval shrinks as the defect size decreases. Supercoiling experiments using a 20 bp DNA hairpin as a defect have observed such a linking number interval [84].

3.2.B.vi Rebuckling transition

In Fig. 3.6, the defect is located 450 bp away from one of the ends of the 6 kb DNA molecule. As mentioned previously (Sec. 3.2.A.iii), the position of the defect imposes a critical size of 900 bp or 0.3 μm on the defect-pinned plectoneme domain. This is directly related to the fact that a defect-pinned domain must have the defect site at its tip, where an energy-saving kinked end loop is placed [Fig. 3.1(b)]. When a defect-pinned domain nucleated at buckling becomes critical in size, nucleation of a mobile plectoneme domain is required to store additional linking number as writhe, resulting in an increased probability of a mobile domain at the rebuckling transition [Fig. 3.6(c′)].

The nucleation of a mobile plectoneme domain may or may not be accompanied with a reduction in the equilibrium probability of the defect-pinned domain ($\mathcal{P}_{10}$). This leads to two possible post-rebuckling states: one mobile plectoneme state ($\mathcal{P}_{01}$), and the two-domain plectoneme state ($\mathcal{P}_{11}$) containing one mobile domain and a defect-pinned domain. Figure 3.8(a) shows a sketch of the three coexisting states at the rebuckling transition. The corresponding probability of occupancy of these states at the rebuckling point are also plotted for various defect sizes [Fig. 3.8(b)].

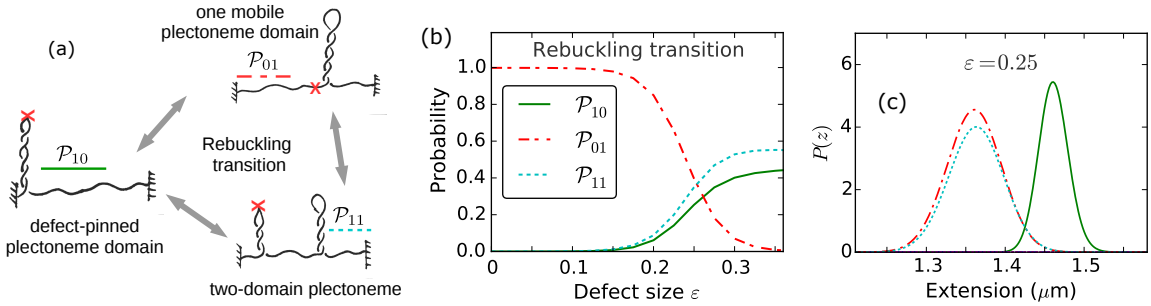

Figure 3.8: Rebuckling transition for various defect sizes. (a) Schematic of the three possible states at the rebuckling transition: the critically-big defect-pinned plectoneme domain ($\mathcal{P}_{10}$, green solid lines), one mobile plectoneme domain ($\mathcal{P}_{01}$, red dot-dashed lines), and two-domain plectoneme containing one defect-pinned and one mobile domains ($\mathcal{P}_{11}$, cyan dotted lines). (b) Total probability of the three states at the rebuckling point (ΔLk=28 under $f = 2$ pN and 0.5 M Na^+, see Fig. 3.6) as a function of the defect size ε. For *small defects* ($0 < \varepsilon < 0.1$), the post-buckling state is the mobile domain ($\mathcal{P}_{01}$) and not the defect-pinned domain ($\mathcal{P}_{10}$) (see Fig. 3.7). As a result, rebuckling is not observed for small defects and the $\mathcal{P}_{01}$ state continues to increase in size after buckling, same as the case for a defect-free DNA (Fig. 3.2). For *intermediate defects* ($0.1 < \varepsilon < 0.25$), the defect is large enough to bias nucleation of the defect-pinned domain ($\mathcal{P}_{10}$) at the buckling transition (Fig. 3.7); however, at the rebuckling point, one mobile domain ($\mathcal{P}_{01}$) is more stable than the two-domain state ($\mathcal{P}_{11}$). For *large defects* ($\varepsilon > 0.25$), the defect-pinned domain ($\mathcal{P}_{10}$) is highly stable, resulting in nucleation of a new mobile domain at the rebuckling point; this makes the two domain state ($\mathcal{P}_{11}$) favored after the rebuckling transition. Note the higher probability of $\mathcal{P}_{10}$ for larger defects, which is due to a shift of the rebuckling transition to higher linking numbers for larger defect sizes (Fig. 3.10). (c) Probability density of extension for $\varepsilon = 0.25$ at the rebuckling point. The bimodal extension profile is due to the finite nucleation energy associated with a teardrop loop of a mobile domain. The state populating the lower-extension mode of the distribution depends on the size of the defect, such that large and intermediate defects favor $\mathcal{P}_{11}$ and $\mathcal{P}_{01}$ states, respectively. Small defects show a unimodal extension profile after buckling transition, and do not exhibit rebuckling.

3.2.B.vii Nucleation cost at rebuckling depends on the defect size

The nucleation energy of the post-rebuckling state depends on the relative stability of the defect-kinked end loop with respect to a teardrop loop, such that the nucleation cost of the mobile plectoneme state ($\mathcal{P}_{01}$) increases with the defect size. Note that the total nucleation cost of the $\mathcal{P}_{01}$ state at the rebuckling point is associated with, first, the energy cost of a teardrop loop which is defect-size independent, and second, the cost of *unpinning* a defect-pinned domain (i.e., a decrease in the equilibrium probability of the pinned domain) which increases for larger

defects. While, the nucleation cost for the two-domain state ($\mathcal{P}_{11}$) depends only on the energy of the teardrop loop of the added mobile domain and does not change with the defect size. The discontinuity in extension and overshoot in torque, seen at the rebuckling point [Fig. 3.6(a)-(b)], is due to the finite nucleation cost associated with the post-rebuckling state.

3.2.B.viii Rebuckling transition is absent for small defects

Small defects ($0 < \varepsilon < 0.1$) do not exhibit rebuckling transition because the defect-pinned domain is not the most probable post-buckling state [Fig. 3.6(a)-(c)]. For small defects, a mobile plectoneme ($\mathcal{P}_{01}$) is nucleated at the buckling transition which increases in size with increasing linking number in the buckled state. Since a mobile plectoneme domain ($\mathcal{P}_{01}$) is the most probable buckled state for small defects, the probability corresponding to $\mathcal{P}_{01}$ is the highest for small defect sizes in Fig. 3.8(b). For small defects, the defect-pinned domain is the second most probable post-buckling state, and the probability goes to zero when the total plectoneme size is larger than its critical size [Fig. 3.8(c)].

3.2.B.ix Defect-pinned domain is unpinned at rebuckling transition for intermediate defects

For intermediate defects ($0.1 < \varepsilon < 0.25$), probability of the defect-pinned domain decreases at rebuckling. This means that the critically-big defect-pinned domain is "unpinned" or transformed into a mobile domain by displacing it along the DNA contour, thereby replacing the kinked end loop with a teardrop end loop. Hence, for intermediate defects, the pre-rebuckling state is the defect-pinned domain ($\mathcal{P}_{10}$) and the post-rebuckling state is a mobile plectoneme domain ($\mathcal{P}_{01}$) [Fig. 3.8(b)].

3.2.B.x Large defects show coexistence of two plectoneme domains after rebuckling

For large defect sizes ($\varepsilon > 0.25$), higher stability of the defect-pinned domain resists decrease in its equilibrium probability, and results in addition of a mobile plectoneme domain [Fig. 3.6(c′)].

Thus, for large defects, the favored post-rebuckling state is the two-domain plectoneme state ($\mathcal{P}_{11}$); whereas, the pre-rebuckling state is the defect-pinned domain ($\mathcal{P}_{10}$), which is essential for rebuckling to occur [Fig. 3.8].

3.2.B.xi Critical linking number for rebuckling is higher for larger defects

For intermediate defects, the rebuckling transition occurs at a higher linking number for larger defect sizes, which is related to the increased stability of the defect-pinned domain [Fig. 3.6]. This is because the nucleation cost at rebuckling increases with increasing defect size resulting in an increase in the associated critical linking number. However, for large defects, the nucleation cost does not depend on the defect size resulting in a critical linking number that does not change with the defect size. The shift in the critical linking number with defect size is also seen experimentally [48], and is quantitatively analyzed later in this article [Fig. 3.10].

3.2.B.xii Rebuckling is absent at low stretching forces

At lower forces, the nucleation cost of a plectoneme domain decreases, resulting in an increased tendency to proliferate multiple domains. The energy difference between a mobile and a kinked end loop also decreases with decreasing force, leading to a coexistence of the defect-pinned and mobile plectoneme domains in the post-buckling state. For intermediate defects, below ≈ 1 pN the mobile domain is the most probable post-buckling state, whereas, the defect-pinned domain is the second most probable. Hence, intermediate defects do not show rebuckling transition when the stretching force is less than ≈ 1 pN.

3.2.B.xiii Rebuckling is absent at low salt concentrations

Lower salts promote proliferation of multiple plectoneme domains due to decreased stability of plectoneme superhelices that contain close proximity of DNA segments (Fig. 3.4). This causes an increase in the probability of the two-domain state at the rebuckling transition as the ionic strength of the solution is decreased. Hence, at salt concentrations lower than ≈ 0.1 M

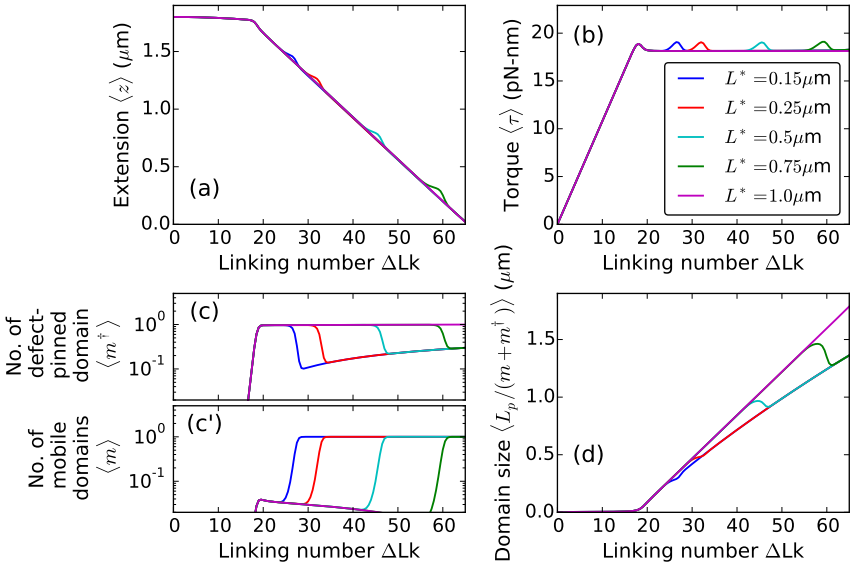

Figure 3.9: Displacement of the defect site. Twist response of 2 μm DNA under 2 pN force at 0.5 M Na^+ with an intermediate defect ($\varepsilon = 0.2$) located $L^* = 0.15$ (blue), 0.25 (red), 0.5 (cyan), 0.75 (green), and 1 μm (magenta) from one of the DNA ends. The location of the defect site controls the critical size of the pinned plectoneme domain ($2L^*$) nucleated at the buckling point. This results in an increase of the critical linking number for the rebuckling transition for defects located farther away from the end, seen as a shift in extension and torque bumps. For a centrally-located defect ($L^* = 1$ μm, magenta lines) the rebuckling transition does not occur because the critical size of the defect-pinned domain is equal to the total size of the DNA. Note that unpinning of the defect-pinned domain occurs at the rebuckling transition, as expected for intermediate defects.

Na^+, intermediate defects preferentially nucleate the two-domain state ($\mathcal{P}_{11}$) at the rebuckling transition.

3.2.B.xiv Displacing the defect shifts the rebuckling point

For a given force and length of the molecule, the critical linking number associated with rebuckling depends on the location of the defect. A defect located farther away from the closest end of the DNA (i.e., larger L^*) increases the critical size of the defect-pinned domain, resulting in an increase in the rebuckling critical linking number (Fig. 3.9). Note that for a defect located at the middle of the DNA, the defect-pinned domain is critically big only in the purely plectoneme state (no unbuckled DNA or zero extension), and the rebuckling transition does not occur, as

has also been shown experimentally [48].

3.2.B.xv Comments on experimental detection of three-state coexistence

The extension distributions at the buckling and rebuckling transitions appear bimodal, even when there are three coexisting states [Figs. 3.7(c) and 3.8(c)]. The lower-extension peak at the transition is the sum of contributions from the two possible post-transition states ($\mathcal{P}_{01}$ and $\mathcal{P}_{10}$ for buckling, and $\mathcal{P}_{01}$ and $\mathcal{P}_{11}$ for rebuckling). This means that experimentally measuring the extension profile at the buckling or rebuckling transitions, as done in magnetic tweezer experiments [48], does not inform on the identity of the post-transition state.

Measurements of the lifetime of the lower and higher extension modes are also unlikely to shed light on the possibility of multiple states contributing to the lower-extension peak. The lifetimes of the higher and lower extension states are simple exponential distributions. In case of a three-state transition where one of the states is *hidden* (i.e., the transition out of one state is not the same as the transition in to the other and vice-versa), the lifetimes follow a Gamma distribution (polynomial increase for small times and exponential decay for long times). However, if two of the three states are grouped together, like the predicted scenario, the transition out of one state is the same as the transition in to the other and vice-versa, resulting in an exponential distribution of the lifetimes. Such grouping of the two states in a three-state transition simply blinds the observer to $\sim 1/3$ of the transition events and the overall kinetics appears to be that of a two-state transition.

However, DNA-visualization experiments, where the DNA backbone is labeled with a fluorescent dye [89, 91], may be able to distinguish between a defect-pinned domain, a mobile domain, and a two-domain plectoneme state. Our model predicts the most probable state for a given linking number and force, but does not inform on the kinetic pathways at the transition. Precise control of the DNA linking number in visualization experiments may also be able to report on kinetically-favored states at buckling and rebuckling transitions.

3.2.C Comparison with experimental data

In this section we compare our results with magnetic tweezer experiments on supercoiled DNA with a base-unpaired region [48]. We obtain a quantitative relation between the number of adjacent base mismatches n, and the theoretical defect size ε. Experiments and theory show good agreement for various thermodynamic trends.

3.2.C.i Critical linking number shift

3.2.C.ii Buckling occurs at a lower linking number for larger defects

The critical linking number associated with buckling shows a general decreasing trend with increasing defect size, because of the lower nucleation cost for larger defects [gray arrows in Fig. 3.10(a)-(a′)]. We define the critical linking number as the point where the higher and lower extension peaks have equal weights (i.e., $\Delta F = 0$ in Fig. 3.10). Experiments show a similar shift in the buckling point to a lower linking number as the number of adjacent base mismatches n is increased [48] [Fig. 3.10(a′)]. For small defects, the nucleation cost at buckling transition is less sensitive to the defect size because of the lower probability of the defect-pinned domain [Fig. 3.7(b)], resulting in a very small to no shift in the buckling point [Fig. 3.10(a)].

3.2.C.iii Rebuckling occurs at a higher linking number for larger defects

The rebuckling point shifts to higher linking numbers with increasing defect size for intermediate defects. As previously mentioned, rebuckling does not occur for small defects [In Fig. 3.10(b), the blue dashed curve corresponding to a small defect does not intersect $\Delta F = 0$]. For large defects, the critical linking number for rebuckling transition does not depend on the defect size. The nucleation cost of the post-transition state (two-domain plectoneme, $\mathcal{P}_{11}$) for large defects is associated with a mobile end loop and does not depend on the defect size. Experiments show the expected trend for the rebuckling point shift [Fig. 3.10(b′)]. Note that $n = 4$ and 16 bp cases show a very small shift in the rebuckling point, the expected behavior for large defects.

We compare theoretical results for 0.5 M Na^+ with experimental observations at 1 M NaCl

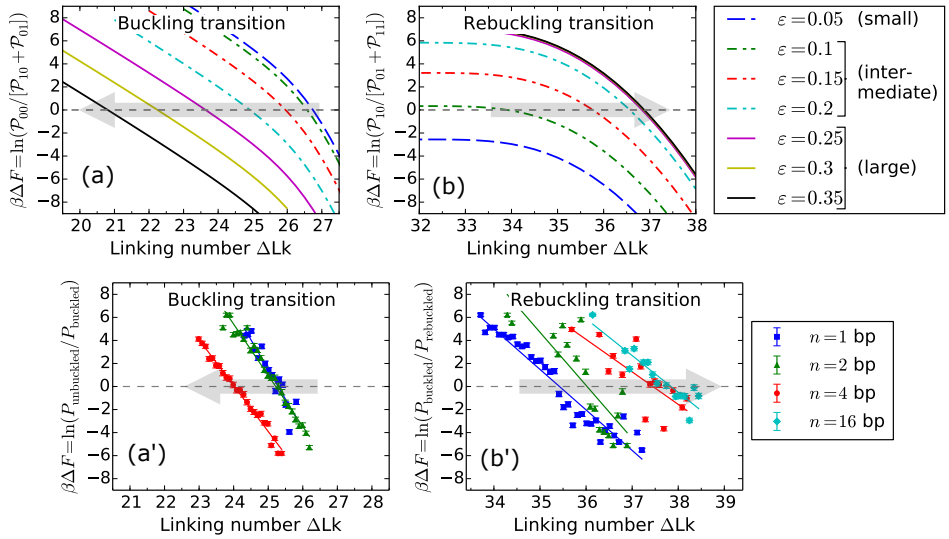

Figure 3.10: Comparison of theoretical and experimental shifts in the critical linking numbers associated with buckling and rebuckling transitions as a function of the defect size for $f = 3.6$ pN. The size of the defect is defined theoretically via the parameter ε, and experimentally, as n, the number of adjacent base pair mismatches on the DNA [48]. Theoretically, the defects are categorized into small, intermediate, and large (Fig. 3.8) depending on the numerical value of ε, as shown in the figure legend. The free energy difference, ΔF between the lower and higher extension states at the buckling or rebuckling transitions is obtained from the logarithm of the ratio of area-under-curve of extension histograms (as shown on the y-axis labels). For a specified defect size, the critical linking number corresponds to $\Delta F = 0$. The gray shaded arrows show the direction of increasing defect sizes. (a)Buckling transition: Theoretical plot shows a decrease in the critical linking number with an increase in ε for intermediate and large defects, whereas, the buckling point does not shift for small defects. Large and intermediate defects nucleate a defect-pinned plectoneme domain that has a lower nucleation energy, this causes a decrease in the critical linking number. Small defects nucleate a mobile domain at the buckling transition, as a result the critical linking number is independent of the defect size. (a′) A similar shift of the critical buckling point to lower linking numbers with increasing n is observed experimentally (see Fig. S4(b) of Ref. [48]), where the solid lines show the best-fit linear regression for various n. (b) Rebuckling transition: Theoretical plot showing the change in the critical linking number as a function of ε. Small defects do not show rebuckling transition: the blue dashed line does not intersect $\Delta F = 0$. Intermediate defects show an increase in the associated critical linking number with increasing ε. Rebuckling for intermediate defects involves a decrease (increase) in the probability of the defect-pinned domain (one mobile domain), and consequently, a higher stability of the defect-pinned domain delays the rebuckling transition. Large defects show rebuckling but not a shift in the rebuckling point with ε. For large defects, a mobile domain is added to the defect-pinned domain at the rebuckling transition, making the rebuckling critical linking number independent of ε. (b′) Experimental plot of ΔF near the rebuckling transition for various n (solid curves are the best-fit straight lines) agrees with the theoretical trend of the shift in the rebuckling critical linking number.

[48] [Figs. 3.10 and 3.11]. The increased abruptness of the buckling and rebuckling transitions at higher salt concentrations, related to the higher stability of plectoneme superhelices, makes the high salt scenario suitable for experimental studies. However, at salt concentrations higher than 0.5 M, the Debye-Hückel approximation becomes questionable at best.

3.2.C.iv Relation between ε and n

The experimental (n) and theoretical (ε) defect sizes are expected to have a direct monotonic relationship for $n \geq 1$, and $\varepsilon = 0$ for $n = 0$. We look for a simple linear variation for $n \geq 1$: $\varepsilon = a + bn$, ignoring higher order terms that only contribute for very large defect sizes.

The fact that rebuckling is observed experimentally for $n = 1$ suggests that it is not a small defect. The rebuckling point shifts in the experimental plot with increasing defect size [Fig. 3.10(b′)]; this indicates that at least $n = 1$ and 2 must be intermediate defects. Comparing the critical linking number shifts we see that an increase in n by 1 roughly corresponds to an increase of 0.05 in ε, implying $b \approx 0.05$.

The smallest defect sizes corresponding to rebuckling transition are $\varepsilon \approx 0.1$ and $n = 1$; this suggests $a + b \approx 0.1$. Hence, we find $n = 1, 2$, and 4 respectively correspond to $\varepsilon = 0.1, 0.15$, and 0.25.

The bending energy saved by increasing the defect size by $\Delta\varepsilon$ is $\beta\Delta E = \Delta\varepsilon\sqrt{\rho\beta A f}$ [Eq. (3.8)]. This suggests that increasing the size of a DNA lesion by one base pair ($\Delta\varepsilon = 0.05$) saves $\approx 0.7\ k_{\mathrm{B}}T$ of bending energy by allowing a sharper kink at the lesion site under physiological forces (1 pN). Our estimation is based on intermediate defects, and may be an underestimation for larger defect sizes because, first, we assumed a linear relationship between n and ε; and second, we ignored any defect-induced perturbation in DNA twist.

One may ask, given that a larger defect stabilizes the kinked end loop, when is it favorable to break the intact base pairs adjacent to a base-unpaired region, i.e., save bending energy by increasing the defect size at the cost of base-pairing energy? The energy saved for an increase in the defect size by 1 bp ($\Delta\varepsilon = 0.05$) when equated with the average base-pairing energy

$\approx 2.5\ k_{\mathrm{B}}T$ [111], we find a critical force ≈ 12.5 pN which compares with DNA unzipping force ≈ 12 pN [112]. This means that the base-unpaired region of a supercoiled DNA will show an equilibrium increase in size when stretched under forces larger than ≈ 12.5 pN. This value, based on our estimation of the defect size, provides a reference point. Experimental significance of this force is not clear because we ignore transitions in the double-helical structure of DNA, which are known to occur at torques > 40 pN-nm, which corresponds to buckling torque under < 6 pN stretching force [36].

3.2.C.v Abruptness of the transitions

The extension change at a transition is a measure of the abruptness of the transition. Transitions with more resolved peaks in the bimodal extension profile are more abrupt.

3.2.C.vi Buckling transition

At the buckling transition, the total size of the nucleated domain (lower extension state) has contributions from the end loop (kinked or mobile) and the plectoneme superhelical structure. Larger nucleation cost increases the critical buckling linking number which increases the amount of plectonemic turns in the nucleated domain. As the defect size gets bigger (intermediate and large defects), the nucleation cost decreases, which reduces the plectonemic contribution to the nucleated domain, resulting in a decrease of the extension change [Fig. 3.11]. Note that the size of a kinked end loop at 3.6 pN is $30 - 20$ nm for defect sizes $0 < \varepsilon < 0.35$ [Fig. 3.11, and Eq. (3.9)], and it goes to zero as $\varepsilon \to 1$.

Hence, the steep decrease in the extension change at the buckling transition for intermediate defects is predominantly due to a decrease in the plectoneme contribution to the nucleated domain [Fig. 3.11]. A large defect nucleates a kinked end loop with minimal plectoneme. This makes the extension change depend solely on the size of the kinked end loop, and the small change in the kinked end loop length with increasing defect size results in the shallow decrease in extension change for larger defects [Fig. 3.11]. Small defects nucleate a mobile domain,

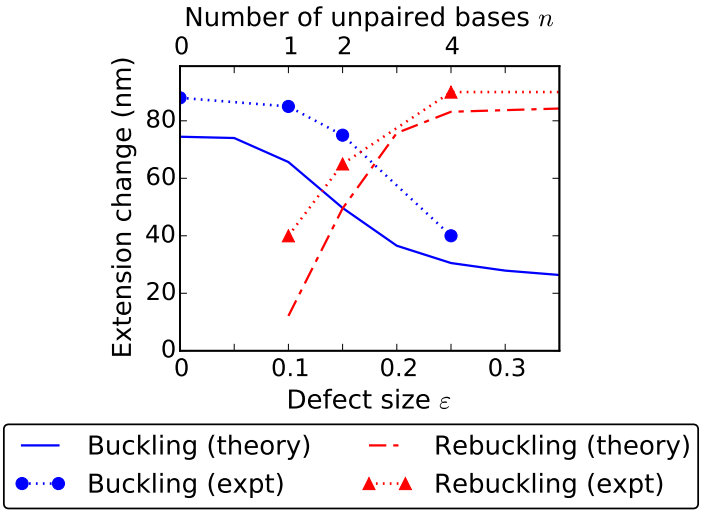

Figure 3.11: Change in extension at the buckling and rebuckling transitions as a function of the defect size for 3.6 pN force. The extension change at a transition depends on the size of the nucleated domain, which has contributions from the end loop and the plectoneme superhelical state. In case of buckling (blue solid line), the critical linking number decreases with increasing defect size causing a decrease in the plectoneme contribution as well as the size of the end loop, reducing the extension change at the transition. However, for small defects, nucleation of a mobile domain causes a saturation in the extension change for buckling transition and an absence of rebuckling transition, hence no associated extension change (red dot-dashed line). For intermediate defects, rebuckling extension change increases with the defect size due to an increase in the associated critical linking number resulting in a higher superhelix contribution. Whereas, for large defects, the extension change is constant because of a fixed superhelix contribution corresponding to a fixed critical linking number (Fig. 3.10). Experimental data for the extension change (see Fig. 2(c) in Ref. [48]) as function of the number of adjacent unpaired bases n (top x-axis) compares well with theory. The experimental error bars, omitted in the plot, are smaller than the size of the point markers.

resulting in an extension change that does not depend on the defect size [Fig. 3.11].

3.2.C.vii Rebuckling transition

Extension change at the rebuckling transition increases with increasing defect size for intermediate defects. Rebuckling occurs at a higher linking number for larger defects due to increased nucleation cost, resulting in an increased plectoneme contribution to the nucleated domain. This means that at the rebuckling point, total plectoneme length in one mobile plectoneme domain ($\mathcal{P}_{01}$, the favored post-rebuckling state for intermediate defects) is larger than that for the

defect-pinned plectoneme domain ($\mathcal{P}_{10}$, the pre-rebuckling state). The plectoneme contribution to the post-rebuckling state also increases with the defect size, resulting in an increase of the rebuckling extension change [Fig. 3.11].

Large defects nucleate the two-domain plectoneme at the rebuckling point, which has a defect-independent nucleation cost, and as a result the plectoneme contribution to the nucleated domain does not depend on the defect size. This produces a fixed extension change for rebuckling in large defects [Fig. 3.11]. Small defects do not show rebuckling, hence there is no extension change associated with small defects at rebuckling.

3.2.C.viii Detection of rebuckling signal in force-salt landscape

Experimental detection of the rebuckling transition relies on the resolution of the corresponding bimodal extension distribution [48]. As the force and ionic strength are lowered, the rebuckling signal diminishes due to overlapping peaks in the extension profile that gives the distribution an overall unimodal character. Consequently, rebuckling is experimentally observed mainly in the high salt and high force regime [48].

Pearson chi-squared test

We implement a chi-squared analysis that compares the theoretical extension distribution to a single Gaussian that best fits the distribution. We fit the theoretical total-extension histograms [$P(z)$, Eq. (3.13)] near the rebuckling point to a Gaussian distribution, where the mean and the variance are obtained via least-squared method. We then use a Pearson chi-squared test [113] to find the *p-value* of the fit; the p-values corresponding to various salts and forces at the rebuckling transition are plotted in Fig. 3.12.

Low p-values indicate that our null hypothesis of the chi-squared test: *extension histograms are Gaussian*, is less likely, which arises from the distinctly bimodal nature of the extension profile. As a result, low p-values correspond to the regions in the force-salt landscape where rebuckling produces an extension profile with two resolvable peaks, and is more likely to be

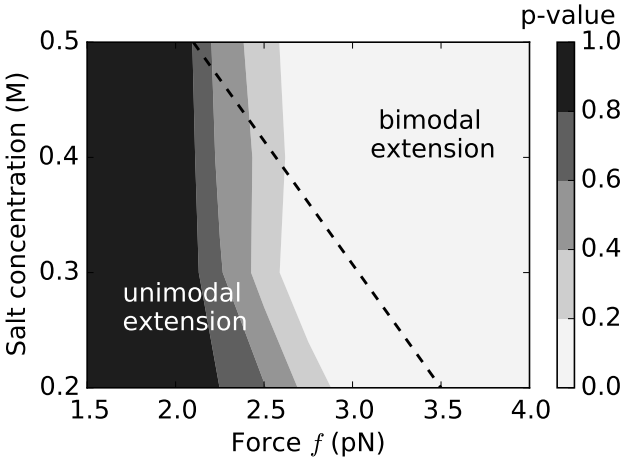

Figure 3.12: Contour plot of p-values near rebuckling transition ($\varepsilon = 0.15$) for various points on the force-salt landscape. We fit the theoretical extension distributions near the rebuckling point to a single Gaussian distribution, and using the chi-squared test calculate the p-value, which serves as a goodness-of-fit statistic. Lower p-values (lighter shade) indicate that the extension histograms near the rebuckling point are characteristically bimodal and are not well fitted by single Gaussian distributions; whereas, higher p-values (darker shade) indicate the rebuckling extension histograms are well approximated as Gaussian distributions. Since the experimental signal associated with the rebuckling transition is the bimodal character of DNA extension, we find that the rebuckling transition is more likely to be observed experimentally when the p-value is low, i.e., higher forces and higher salts. The dotted black line shows the experimentally observed boundary for the appearance of the rebuckling signal associated with a defect containing two adjacent base-pair mismatches ($n = 2$ bp) (see Fig. 3 of Ref. [48]). The bimodal rebuckling signal was reliably observed in experiments for salts and forces on the right-hand side of the dotted line.

observed experimentally. On the other hand, higher p-values suggest that the total-extension profile is well fit by a single Gaussian distribution (unimodal), and experimental detection of rebuckling is less probable.

The dotted line in Fig. 3.12 shows the experimentally-observed boundary between the disappearance of the rebuckling signal on the left-hand side of the dotted line, i.e., at lower forces and salts; and the appearance of the signal on the right-hand side of the dotted line, i.e., for higher forces and salt concentrations [48]. The theoretical picture correlates well with the observed statistics of the rebuckling signal, where bimodal extension profile is expected for higher salts and higher forces (low p-value) [lighter shade in Fig. 3.12]. We note that the experimental

boundary [dotted line in Fig. 3.12] is associated with a statistical disappearance of the experimental rebuckling signal [48]; and theory suggests that this disappearance is not due to an absence of the rebuckling transition, but thermal fluctuations overshadowing the experimental signal at lower forces and salts.

3.3 Summary

3.3.A Defect-free DNA

Supercoiled defect-free DNA shows a linear torque buildup and small extension change upon small linking number change in the DNA (Fig. 3.2). Larger torques for higher linking numbers drive coexistence of a buckled plectoneme state that is favored due to its substantial writhe linking number contribution associated with the geometry (Fig. 3.1). DNA torque is nearly constant in the plectoneme-coexistence state due to plectonemic writhe. Torque, however, increases in the purely-plectoneme state due to an increase in DNA twist with increasing linking number (Fig. 3.3).

Abrupt plectoneme-buckling transition The buckling transition marks the onset of a plectoneme-coexistence state. Nucleation of a plectoneme domain is abrupt due to the finite-energy cost associated with a plectoneme end loop. The nucleation cost is also related to the discontinuous change in extension and overshoot in torque at the buckling transition (Fig. 3.2).

Multiple plectoneme domains in long DNA molecules Longer DNA molecules have a larger configuration entropy associated with a plectoneme domain that reduces the nucleation cost of the domain. This results in an increase in the equilibrium number of plectoneme domains coexisting in buckled DNA (Fig. 3.3). However, in the purely-plectoneme state of the DNA, a single plectoneme domain is favored and the post-buckling torque increases (Fig. 3.3).

Highly stable plectoneme at higher salts Increased salt concentration of the solution causes an increased screening of the DNA charge that leads to highly stable plectoneme super-

helices. Whereas, at lower salts, DNA loops are favored over plectoneme superhelices (Fig. 3.4). This reduces the DNA-length contribution of plectoneme superhelices to the nucleated buckled domain at the buckling transition for lower salts, resulting in a lower extension discontinuity or more rounded buckling transition.

The extension distribution is bimodal in the buckled state for low salt concentrations due to proliferation of multiple domains in the plectoneme-coexistence state (Fig. 3.5), which may be possible to observe in magnetic tweezer experiments.

Plectoneme tails A plectoneme tail is a finite-sized structure connecting the plectoneme domain to the unbuckled part of the DNA. The energy cost of a tail region is related to the constraint of a continuous change of DNA curvature from the plectoneme to the unbuckled DNA, and compares with the energy cost of constraining the ends of a DNA loop ($\approx$ 10% of the bending energy of a loop, i.e., $\approx 2\ k_BT$ under 2 pN force, see Sec. IVA of Ref. [87]). The plectoneme tail causes a small increase in the nucleation cost of the plectoneme domain, which we have ignored in our model for simplicity.

However, the tail region may induce a preference for the spatial location of the plectoneme domain. By placing the plectoneme domain at one of the ends of the DNA, so that its tail coincides with a DNA-tether point, the energy associated with the tail can be halved. This is because one of the ends of plectoneme is the tether point and the constraint of a continuous DNA curvature is released for that end. Lowering the plectoneme tail energy by $\approx 1\ k_BT$ results in $\approx$ 3-fold increase in the probability of localizing the domain at one of the DNA ends. There is experimental evidence for preferential localization of a plectoneme domain near the DNA end [91].

3.3.B DNA with a defect

We analyzed the effect of an immobile point defect, e.g., a short base-unpaired region [48], on the mechanical response of DNA. We hypothesized that a defect allows a kink on the DNA at the defect site, and the location of the defect imposes a critical size on the plectoneme domain

that nucleates at the defect site. A plectoneme domain with the tip of its end loop placed at the defect site (i.e., the defect-pinned plectoneme domain), is energetically favored due to a lower-energy kinked end loop. However, the requirement of putting the end loop at the defect site causes the other end of the plectoneme (i.e., the plectoneme tail) to coincide with a DNA tether point for a certain size of the defect-pinned plectoneme domain, which is its critical size. From simple geometry, a critically-big defect-pinned domain cannot store more superhelical turns [Fig. 3.1(b)].

In previously reported experiments [48], we used one or more adjacent base pair mismatches at a specified location on the DNA as a point defect. Twisting a base-pair mismatched DNA showed rebuckling transition near the point where the defect-pinned plectoneme domain reaches its critical size [Fig. 3.6].

Defect size A defect of size ε can nucleate a kinked end loop that has a bending energy equal to (1-ε) times the bending energy corresponding to a teardrop loop [Eq. 3.8]. The defect size ε is related to n the number of adjacent base-pair mismatches on the DNA, such that a larger value of n indicates higher ε. We categorized the defect size into: small ($0 < \varepsilon < 0.1$), intermediate ($0.1 < \varepsilon < 0.25$), and large ($0.25 < \varepsilon < 1$).

Buckling transition Twisting the double helix containing a defect buckles when it is energetically favorable to convert twist into writhe. For intermediate and large defects, the buckling transition occurs via nucleation of a kinked end loop, which makes the defect-pinned plectoneme domain favored at the buckling transition (Fig. 3.7). For small defects, a mobile plectoneme domain is nucleated at the buckling transition, the DNA kink at the defect reduces bending energy but the defect-pinned domain lacks stabilization from diffusion entropy. This leads to a higher net nucleation cost for the defect-pinned domain featuring small defects than a mobile domain.

The critical linking number associated with the buckling transition decreases with increasing defect size for intermediate and large defects. This is because of the lower nucleation cost

associated with a more sharply kinked end loop of a larger defect size (Fig. 3.6). However, the critical linking number does not change for small defects, because a mobile plectoneme domain is nucleated at the buckling point.

Rebuckling transition As the linking number is increased after the buckling transition, the defect-pinned domain grows in size and becomes critically big. Further increase in DNA linking number leads to nucleation of a mobile plectoneme domain, required to store additional plectonemic-superhelical turns.

For large defects, the high stability of the kinked end loop leads to addition of a mobile domain such that the post-rebuckling state is that of a two-domain plectoneme featuring one mobile and one defect-pinned domains (Fig. 3.8). For intermediate defects, the defect-pinned domain unpins itself to nucleate a mobile plectoneme domain. Since only one domain is favored due to high stability of plectoneme superhelices (higher salt concentration scenario), the critically-big defect-pinned domain decreases in probability as the mobile domain becomes more probable after the rebuckling point. For small defects, the rebuckling transition does not occur as the defect-pinned domain is not the most probable after buckling transition (Fig. 3.8).

For intermediate defects, the critical linking number associated with rebuckling transition increases with increasing defect sizes. This is due to an increased stability of the defect-pinned domain for larger defect sizes (Fig. 3.6). However, for large defects, the critical linking number does not change because the energy difference between the pre- (critically-big defect-pinned domain) and post-rebuckling (two-domain plectoneme) states does not depend on the defect size.

Lowering the salt concentration of the solution results in an increased proliferation of buckled domains. This increases the probability of the two-domain state for lower salts. Consequently, at low salt concentrations (< 0.1 M Na^+) unpinning of the defect-pinned domain is not favored for intermediate defects.

An unpinned mobile domain may have a preference for placing its tail at one of the DNA ends or the defect site. This is due to the energy saved from an unconstrained plectoneme tail

at the DNA ends or the defect site, which is expected to increase the spatial plectoneme density at those regions and may be visible in fluorescent-DNA experiments [89, 91].

Number of unpaired DNA bases n, and defect size ε We compare theoretical and experimental shifts in the critical linking numbers at buckling and rebuckling transitions for various defect sizes (Fig. 3.10). Experimentally, the defect size is controlled via n the number of adjacent unpaired bases on the DNA [48], which has a direct monotonic variation with the theoretical defect size ε. We find that $n = 1, 2$, and 4 roughly correspond to $\varepsilon = 0.1, 0.15$, and 0.25, respectively (Fig. 3.10). This indicates that, for intermediate defects, increasing the size of a base-unpaired region by 1 bp saves $\approx 0.7\ k_BT$ of bending energy under 1 pN stretching force.

We predict $n = 1$ and 2 to be intermediate defects, i.e., the defect-pinned domain unpins itself at the rebuckling transition, as opposed to nucleating a two-domain plectoneme under higher salt concentrations (≈ 0.5 M Na^+) [Fig. 3.8(b)]. The extension histograms at the rebuckling transition are bimodal even though there are three states: both the mobile and the two-domain plectoneme state contribute to the lower-extension mode [Fig. 3.8(c)]. As a result, measuring the extension, as done in magnetic tweezer experiments [48] does not inform whether the post-rebuckling state is a mobile domain or a two-domain state. However, DNA visualization using fluorescence imaging experiments may elucidate the existence of these states.

We assume that only the bending degree of freedom of DNA (and not the twist degree of freedom) is affected by the presence of the defect. This assumption is a simplifying one. Certain defects, such as the one produced by a base-unpaired region on the DNA [48], may absorb twist at the defect site. A more sophisticated model may consider the defect to have a lower twist modulus than the double helix, which is essential when treating the base-unpaired region as a coexisting state [114]. However, when the number of adjacent unpaired bases is only a few compared to thousands of paired bases, twist absorption at the defect is expected to be a small effect and may be ignored.

Extension change at the transitions The change in extension at a transition is non-zero because the nucleated plectoneme domain cannot be smaller than the end loop. Moreover, the relative stability of the end loop compared to plectoneme superhelices determine the contribution of the plectoneme-superhelix state to the nucleated domain; such that when superhelices are highly stable, the plectoneme content of the nucleated domain is larger and results in a larger extension discontinuity. The extension change decreases with increasing defect size at the buckling transition because of a decrease in the plectoneme contribution to the nucleated domain as well as a decrease in the size of the kinked end loop for larger defects (Fig. 3.11). The extension change at the rebuckling transition increases with an increase in the defect size for intermediate defects, and saturates for large defects.

Disappearance of experimental rebuckling signal at lower force and salt The extension distributions at the rebuckling transition are bimodal with well-resolved peaks at higher salts and forces (Fig. 3.12). However, for lower salts and forces, the two peaks overlap giving the extension distribution a unimodal character that obscures experimental detection of rebuckling transition. Lower stability of plectonemic superhelices at lower salt concentrations produces more rounded transitions and masks the rebuckling signal [48].

Biological significance of a defect The parameter ε directly controls the amount of bending energy saved when the tip of a plectoneme end loop is placed at the defect location. The biological relevance of a defect is diverse. Adjacent base-pair mismatches on the double-helix backbone can introduce a spatially-pinned defect [48, 91]; alternately, local structural rearrangement of the double helix by a protein, thereby allowing easy local bending of the DNA backbone [96, 97], might also act as an immobile defect site. Double helices with a single-stranded bulge are a common substrate in single molecule study of various DNA-binding proteins [88]; DNA bulges can also be treated as a defect. The defect size parameter ε may be used as a common scale to compare relative perturbations introduced by defects of varied origin.

Spatial pinning of a plectoneme domain by a defect may have relevance in double-helix base-

pair repair mechanism in the cell [48]. A common mechanism to locate targets in the cell, such as locating a DNA base-pair mismatch region by the repair machinery, is that of a diffusive search [115–117]. Preferential positioning of the defect site containing mismatched base pairs at the tip of a plectoneme may facilitate easier access to the lesion. Moreover, DNA kinks are known to stabilize binding of the enzymes associated with the repair process [118].

Defect induced by DNA sequence Spatial inhomogeneity in a DNA polymer is not only from varied intra-base-pair interactions, but stacking interactions between adjacent base pairs can enhance the inhomogeneity locally for certain positioning sequences (or, nucleosome-positioning sequences) [98–101]. Pinning of a plectoneme domain by certain sequences have recently been demonstrated experimentally [119]. Occurrence of such positioning sequences in a DNA with otherwise random base pairs may be modeled as a spatially-pinned defect. Local stiffness change from one base pair or a weak sequence-induced defect may be expected to be small, i.e., buckling is not necessarily favored at the defect site and rebuckling is not observed. On the other hand, some positioning sequences may generate an intermediate or large defect, thereby favoring nucleation of a buckled domain at the defect site and exhibiting rebuckling transition. Our prediction that rebuckling does not occur for small defects ($\varepsilon < 0.1$) may be used in classifying various positioning sequences.

The possibility of a sequence-induced defect makes the relationship between ε and n (Fig. 3.10) less exact. Placing a base-pair mismatch of size n inside a positioning sequence is expected to generate a larger defect than placing it at a random location. Enhancement of defect-facilitated buckling for a positioning sequence with unpaired bases may be relevant for its *in vivo* mismatch repair. Similarly, occurrence of a sequence-induced defect near one of the DNA ends may favor plectoneme positioning at that end over the other [91, 119].

An alternate model, where the defect is associated with an increased bending energy cost may also be useful. *In vivo*, such a defect will disfavor in-situ nucleosome assembly, thus regulating genome access for DNA-binding proteins. Studying supercoiled DNA with multiple defects is an interesting future prospect. Such studies may show pinning of a plectoneme domain at a defect

site after an unpinning from a larger defect. Furthermore, studying the role of defects on the mechanics of chromatin fibers may also be an interesting future possibility.

To conclude, we have presented a theoretical model for DNA buckling and the analyzed the consequences of introducing a defect. We provided an explicit theoretical treatment of thermal fluctuations in plectonemic DNA that may be relevant in modeling fluctuations of geometrically constrained polymers. We also classified defects depending on their size which leads to various possible states corresponding to different defect sizes that can be probed experimentally.

Chapter 4

Chromosome structure and topology

We model chromosomes as a succession of polymer loops of chromatin [1], where active extrusion of loops allows manipulation of the chromosome structure. This results in a cylindrical morphology of chromosomes where chromatin loops emanate from a flexible chromatin backbone [Fig. 4.1(b)]. Extrusion of loops to increase the loop size leads to a smaller interloop distance, i.e., a higher linear density of loops on the backbone, and drives increased compaction of chromosomes [20, 22]. Compact chromosomes have a high density of chromatin monomers or nucleosomes at the loop bases, i.e., near the chromosome backbone, leading to an osmotic pressure gradient that stretches the loops radially. The high osmotic pressure stiffens the chromosomes and generates a higher thermal persistence length associated with bending of the backbone. High nucleosome density along the backbone forms a cylindrical core composed of densely packed chromatin [Fig. 4.1(c)]. The core imparts mechanical rigidity to chromosomes and reproduces the elastic moduli of compact mitotic chromosomes [43, 120–124].

Cellular chromosomes, by virtue of their confinement, strongly overlap with each other, resulting in an entangled conformation [18, 19]. We find that lengthwise compaction of chromosomes, via loop extrusion, can resolve their entangled topology and drive segregation, given Type-II DNA topoisomerases (TopoII) that allow topology fluctuations. TopoII act locally and

[1] Chromatin is a string of nucleosomes or DNA-protein complexes that forms the basic unit of eukaryotic chromosomes.

by themselves cannot disentangle chromosomes, however, by mediating topology fluctuations, they ensure entanglements do not hinder active compaction and segregation. The level of inter-chromosome entanglement in a confined volume containing multiple loop-extruded chromosomes shows a non-monotonic variation with the loop size, featuring an entanglement-minimizing *optimal* loop length that scales positively with the genome size. Inter-chromosomal entanglements scale negatively with the degree of chromosome compaction, indicating that optimal loops also maximize compaction. Applying our results to chromosomes during interphase, we find that the organization of the genome into chromosomal domains or "Topologically Associating Domains" (TADs) may suppress inter-chromosome entanglements. The model described in this chapter provides a microscopic view of the structural evolution of chromosomes during the cell cycle, that is consistent with existing observations and makes testable predictions. The contents of this chapter are being prepared for publication.

Significance Cellular chromosomes are long biopolymers residing in confined spaces. For eukaryotes, the nucleus constitutes the confining compartment, while bacterial chromosomes are confined by the cell wall. The high degree of confinement promotes topological entanglement between the chromosomes. Simplification of their entangled topology is essential for proper segregation of chromosomes, as occurs during mitosis. Active topology manipulation by TopoII is essential but not sufficient to disentangle the chromosomes, since individual TopoII cannot sense global chromosome topology. However, chromosome disentanglement may be achieved by driving compaction in presence of topology fluctuations. Chromosomes are compacted into cylindrical rod-like structures during cell division. The increased stiffness of compact chromosomes, derived from the high osmotic pressure inside their loop-extruded structures, drives chromosome disentanglement and segregation. This is because denser chromosomes repel each other more strongly upon overlap due to a higher osmotic pressure in the overlap volume. Although the idea that compaction via loop extrusion can drive chromosome segregation has been proposed [18, 19, 21], and since been simulated to show its viability [125], there lacks a quantitative analysis of its implications, which this chapter provides.

Our model provides insight into the structure of mitotic chromosomes: compact chromosomes have a densely packed chromatin core along the backbone, which recapitulates the mechanical properties of mitotic chromosomes observed experimentally [43, 120–124]. The core behaves as an elastic solid made up of closely packed nucleosomes, and is responsible for the structural rigidity of mitotic chromosomes. Our view of the chromosome core is similar to the dense spherical "core" in star polymers, as analyzed by Daoud and Cotton in 1982 [126], where the polymer inside the core is fully stretched and maximally dense. Although, non-histone proteins are essential to generate a compaction pressure leading to formation of the core, the mechanical rigidity of chromosomes, in our case, comes from the densely packed chromatin, as opposed to the elasticity of a protein scaffold [127].

Outline The layout of this chapter is as follows. The necessary details of the model are given in Sec. 4.1, which are subdivided into the structural aspects of loop-extruded chromosomes [Sec. 4.1.A] and the implications of structure on inter-chromosomal entanglements in confinement [Sec. 4.1.B]. We discuss the results of our model in Sec. 4.2, where we compare our results with existing experimental observations. Finally, we conclude in Sec. 4.3 with discussion and future scope of the study.

4.1 Model

In this section, we develop the model through introducing the relevant parameters and concepts. The scaling calculation has been briefly sketched, and the intermediate steps can be found in Appendix D.

A fundamental length scale in for a chromatin chain of N monomers is given by its end-to-end distance or radius of gyration R_F that scales positively with the number of monomers: $R_F(N) \sim N^{3/5}$ [Fig. 4.1(a)]. The end-to-end distance also scales positively with w, the excluded volume associated with the nucleosome monomers: $R_F \sim w^{1/5}$, see Eq. (1.8).

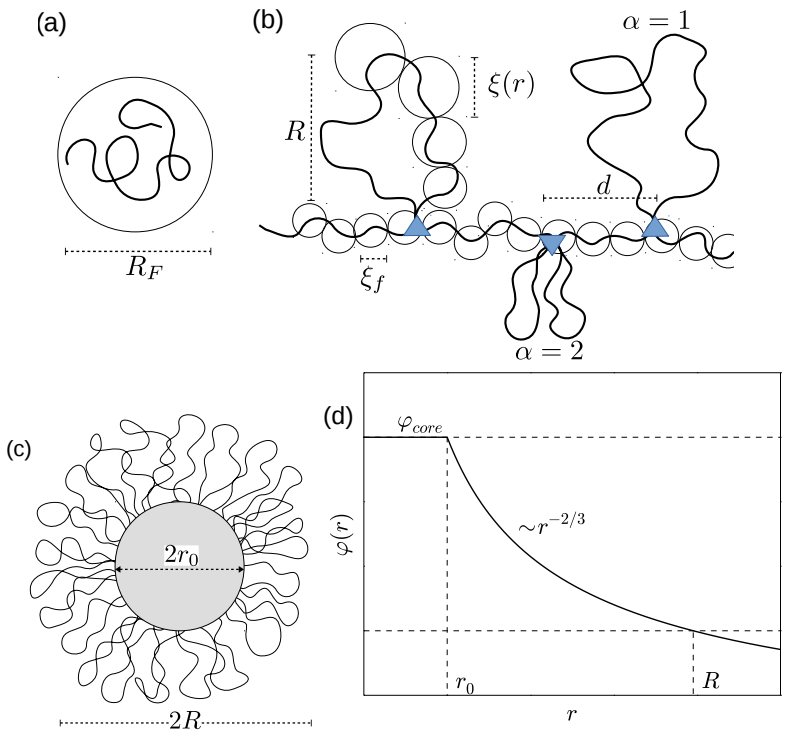

Figure 4.1: (a) Flory radius of a self-avoiding polymer [Eq. (1.8)]. (b) Sketch of a loop-extruded chromosome showing chromatin loops on a backbone, where the circles represent the correlation length inside the chromosome. Out of the three schematic loops, the middle one is divalent ($\alpha = 2$) and the other two are monovalent ($\alpha = 1$). (c) Cross-sectional view of a compact chromosome. The circular cross-section corresponding to the gray shaded area is the densely-packed-chromatin core of diameter $2r_0$. (d) Radial profile of monomer volume fraction $\varphi(r)$ [Eq. (4.1)]. The volume fraction is the highest inside the core [$\varphi_{core} \approx 1$, Eq. (4.4)], and decays radially outwards.

4.1.A Cylindrical loop-extruded chromosomes

Extrusion of chromatin loops [20, 22, 125, 128] by active elements that form the loop anchoring complexes leads to a cylindrical configuration of the chromosomes, where loops emanate radially from a stretched backbone [Fig. 4.1(b)].

4.1.A.i Loop size n

The number of monomers (nucleosomes) in a chromatin loop n, can be obtained from the size of a loop in base pairs (bp) divided by $\approx$ 200 bp corresponding to one nucleosome monomer

(e.g., a 10 kb loop has $n = 50$ monomers). The loop size is a parameter actively controlled by the loop-extruding complexes.

4.1.A.ii Interloop distance d

The linear distance between the loop anchors, or the interloop distance d [Fig. 4.1(b)], is also an actively manipulated length scale that parameterizes the extruded state of a chromosome. Extrusion of chromatin loops in a finite-sized genome causes an increase in the number of loop monomers n, and a corresponding decrease in the interloop distance d. Loop extrusion generates compaction via driving the interloop distance to be smaller than the equilibrium size of the loops: $d < R_F(n)$, which also forces overlap among adjacent loops.

4.1.A.iii Loop valency α

A loop of size n may be divided into α smaller loops (each of size n/α), all of them grafted at the same site [Fig. 4.1(b)]. Multivalent loops are more compact and have a higher local volume fraction. The valency of a loop α, is another active parameter controlling the chromosome structure. Loop extruding enzymes can divide a big loops into multiple smaller ones [20, 22, 125, 129], which increases their loop valencies and drives compaction.

4.1.A.iv Radially decaying monomer volume fraction

Lower interloop distances attained via processive extrusion of chromatin loops lead to a radial gradient of monomer concentration, such that the volume fraction is the highest at the base of the loops and decays radially outwards. This is a well-established feature of polymers grafted on curved surfaces [126, 130–133]. The radial profile of the monomer volume fraction is given as follows [Eq. (D.1)].

$$\varphi(r) = \frac{\alpha^{2/3}}{w^{1/3}} \left(\frac{a^2}{rd}\right)^{2/3} \tag{4.1}$$

The volume fraction is a dimensionless, bounded quantity: $0 \leq \varphi \leq 1$, where a region with maximally packed nucleosomes correspond to $\varphi \approx 1$.

The correlation length or the mesh size inside the extruded chromosome has a radially increasing profile: $\xi(r) = \sqrt{rd/\alpha}$, because the correlation length, which encodes the length scale over which polymer correlations are screened, is higher for lower concentration of monomers [26].

4.1.A.v Loop extension R

The osmotic pressure gradient generated due to the radial decay of monomer volume fraction stretches the loops radially. The radial extension of the loops is obtained from integrating φ over the total volume accessible to a loop.

$$\int_0^R \mathrm{d}r\ \varphi(r) rd = na^3 \quad \Rightarrow \frac{R}{a} = \frac{w^{1/4} n^{3/4}}{\alpha^{1/2}} \left(\frac{a}{d}\right)^{1/4} \tag{4.2}$$

Larger loops (higher n) have a higher extension, however, increasing the loop valency (higher α) causes a decrease in the radial extension of the loop. Decreasing the interloop spacing (lower d) generates a steeper osmotic pressure gradient which results in a higher loop extension.

The scaling dependence: $R \sim n^{3/4}$ is identical to a 2D self-avoiding walk, which is the consequence of confinement of a loop in an effective 2D slit generated by the neighboring loops [134].

4.1.A.vi Loop overlap stretches backbone

Overlap between adjacent loops, due to a small interloop spacing ($d < R_F(n)$), generates a tension along the backbone, because the osmotic pressure tends to push the loops away from each other. The tension f, may be calculated from the axial gradient of the loop free energy F, via: $f = -\partial F/\partial d$ [Eqs. (D.4), (D.5)]. The same tension can be derived from the net osmotic pressure per unit cross-sectional area of the chromosomes.

The correlation length along the backbone is set by the force: $\xi_f = k_B T/f$, and as a result the backbone polymer behaves as a string of tension-induced blobs, each of size ξ_f (Pincus polymer [135]) [Fig. 4.1(b)]. The interloop distance d, is equal to the end-to-end distance of the

stretched backbone [Eq. (D.6)]:

$$d = maw^{1/3}\left(\frac{fa}{k_BT}\right)^{2/3} = aw^{1/5}\alpha^{2/5}m^{12/25}n^{3/25} \tag{4.3}$$

where m is the number of monomers in a segment of backbone between two adjacent loops. The above relation suggests that the polymer degrees of freedom of the backbone chromatin will be completely stretched for $f > k_BT/a \approx 1$ pN. However, this transition to a fully stretched backbone ($d \approx ma$), occurs only in the limit of large loops and small backbone: $m < n^{3/13}$ [obtained using $d > ma$ in Eq. (4.3)].

Fully stretched backbone: tandem arrangement of monovalent loops. Loop extrusion drives an increase in the loop size leading to a fully stretched backbone ($n \gg m$), where the correlation length is minimum: $\xi_f = a$, and the end-to-end extension is linear with the number of monomers: $d = ma$. This configuration corresponds to a tandem arrangement of monovalent loops ($\alpha = 1$, $d \approx a$). However, note that for this configuration, the monomer volume fraction near the backbone approaches the maximum [$\varphi_m = ma^3/(\xi_f^2 d) \approx 1$, Eq. (D.6)], indicating a highly compact chromatin near the loop bases.

4.1.A.vii Dense backbone core

High nucleosome density near the loop bases lead to a cylindrical dense-chromatin core of radius r_0 along the backbone [Fig. 4.1(c)]. Tandemly arranged monovalent loops establish a minimal core ($r_0 = a$). Further loop compaction, e.g., via increasing the loop valency, causes radial spreading of the high density region, i.e., core thickening. This is because the nucleosome density in already maximum in the core volume.

The number of monomers per unit length d along the core is $m + n_c$; where n_c is the number of loop monomers residing in the core, such that $n_c < n$, and m is the number of monomers in the fully stretched backbone ($m \ll n$). The volume fraction in the core approaches unity.

$$\varphi_{core} = (m + n_c)a^3/(r_0^2 d) \approx 1 \tag{4.4}$$

The correlation length inside the core is at its minimum (one monomer unit), and the edge of the core corresponds to the radial distance beyond which the correlation length is higher or the volume fraction is lower [Fig. 4.1(c)-(d)]. This means that every monomer on the surface of the core belongs to a loop grafting site, which yields a geometric relation between the loop valency α, and the core radius r_0, as follows.

$$r_0 = \alpha a^2/d \tag{4.5}$$

The above equation suggests the thickness of the dense core region can be increased by increasing the valency of the loops (α), or by decreasing the interloop distance (d). Note that in order to build a dense core ($r_0 > a$), the following criterion is necessary: $\alpha a > d$.

Loops with a maximum valency:"saturated loops". The hardcore limit on the valency of a loop is the number of the monomers in the loop. However, when the interloop distance is small ($d \approx a$), the core spans the entire chromosome width before the valency reaches this hard limit. Using Eqs. (4.4) and (4.5), we get the maximum value of valency α_{max}, when $n_c = n$.

$$\alpha_{max} = \sqrt{nd/a} \qquad \Rightarrow \max\{r_0\} = a\sqrt{na/d} \tag{4.6}$$

The maximum valency of a loop is higher for a larger loop and for a larger interloop distance. The loops with a maximum valency are saturated, as the volume fraction in the core is already maximum; we refer to chromosomes with a maximum core as being "saturated", which corresponds to a limit of compaction achieved by core formation.

Saturated chromosomes are like solid objects made up of nucleosomes, with a Young's modulus $\approx k_BT/a^3 \approx$ kPa. This internal pressure, however, needs to be actively compensated via a compaction pressure that keeps the chromosomes from relaxing.

4.1.A.viii Persistence length ρ

The cylindrical chromosomes, with loops emanating from a backbone that runs along the axis, have a bending rigidity because bending of the backbone increases the local volume fraction on

the concave side of the bend [45, 132]. The persistence length of the chromosomes ρ, may be calculated from the change in free energy associated with a perturbation in the volume fraction [Eq. (D.7)] [45], or directly using the formula [132]: $k_BT\rho = R^2 d(\partial^2 F/\partial d^2)$, both yield the following scaling relation.

$$\rho = a w^{5/8} \alpha^{1/4} n^{15/8} \, (a/d)^{17/8} \tag{4.7}$$

The persistence length is higher, i.e., chromosomes become stiffer when the loop size increases, because larger loops have stronger overlap with their neighbors leading to a stiffer response to thermal bending. Similarly, chromosomes become more rigid with increasing loop valency and decreasing interloop distance. Loop extrusion drives an increase in n and α and a corresponding decrease in d, all of which contribute to an effective increase in the bending persistence length.

Contribution from the core. The core behaves as a solid with elastic modulus $\approx k_BT/a^3$, and the corresponding persistence length depends on core thickness: $\rho_{core} = r_0^4/a^3$. The core makes the chromosome stiffer, and the net persistence length of chromosomes may be obtained by adding the above contribution from the core to that of the loops [Eq. (4.7)]. However, in the limit of saturated chromosomes [Eq. (4.6)], both the contributions have an identical scaling: $\rho = a(na/d)^2$. Hence, it is sufficient to use the contribution from overlapping loops for chromosome stiffness, and we will employ Eq. (4.7) for persistence length in our calculations.

4.1.A.ix Transformation of chromatin polymer to chromosome

Loop extrusion transforms N units of chromatin monomers or nucleosomes of diameter a, into N' chromosome monomers that are cylindrical units of height ρ and radii R. The coarse-graining of chromatin units to chromosomal units may be depicted as:

$$N(a, w, L) \rightarrow N'(\rho, R, L') \tag{4.8}$$

where $L = Na$, is the total contour length of chromatin, and $L' = N'\rho$ is that of the chromosome. The excluded volume per chromosome monomer: $\rho^2 R \sim w^{3/2}$ [Eqs. (4.2), (4.7)], has

a positive scaling with the microscopic excluded-volume parameter w, suggesting that a lower self-avoidance between nucleosomes also correspond to a lower self-avoidance between chromosomes.

Chromosome polymer is shorter. The contour length of the chromosome L', has a leading order contribution from the cylindrical array of loops, which is the number of loops times the interloop distance, as shown in the first term on the right hand side in the following equation.

$$L' = Nd/(n+m) + \alpha_{end}^{1/5} R_F(n) \tag{4.9}$$

The second term is the contribution from the two ends of a chromosome, which resemble a star polymer, see Ref. [126] for a derivation of the scaling relation. Each end has loops emanating from a hemispherical core of radius r_0, such that the total number of loops, both ends combined, is given by $\alpha_{end} = (\alpha a/d)^2$; this condition is equivalent to Eq. (4.5) for a spherical geometry.

The end-to-end distance of the chromosome polymer is given as: $R_F(N') = \rho^{2/5}(\rho^2 R)^{1/5} N'^{3/5}$, where $\rho^2 R$ is the excluded volume per chromosome monomer [Eq. (1.8)]. The transformation of chromosome end-to-end extension under loop extrusion is given as follows.

$$\frac{R_F(N')}{R_F(N)} = \frac{(d/a)^{1/8}}{w^{1/40}\alpha^{1/20} n^{3/40}} \left(1 + \frac{w^{1/5}\alpha^{2/5} n^{8/5}}{(d/a)^{7/5} N} + \cdots\right) \tag{4.10}$$

The leading order term is the contribution from the cylindrical part of the chromosomes, and the second-order contribution (second term inside the brackets, note the $1/N$ which keeps the term small as long as the loops are not comparable to the genome size) is the contribution from ends. We have implemented the limit $n \gg m$ in the above equation.

Note that loop extrusion causes an increase in n and α, and a decrease in d, which drives a decrease in the leading order term of the end-to-end distance, thus driving compaction. However, the same mechanism causes an increase in the contribution to extension from the ends [second term in Eq. (4.10)], suggesting there is an optimal loop size that minimizes chromosome end-to-end extension.

4.1.B Chromosome entanglements

Entanglements from blob collisions. Every nearby crossing or collision between two polymer segments contributes to entanglement; because type-II DNA topoisomerases (TopoII) can locally change the topology at a nearby crossing [18, 19]. TopoII enzymes can facilitate passing of a double-helix DNA segment through another via introducing a transient cut in the DNA, which makes the total number of nearby collisions a linear scale to measure overall entanglement between chromosomes. Locally acting TopoII enzymes, however, cannot simplify the global topology of chromosomes. These enzymes facilitate topology fluctuations, allowing the active renormalization from nucleosomes to chromosome monomers [Eq. (4.8)] that eventually drives global resolution of chromosome topology, followed by their segregation.

Collisions between segments inside a blob are rare, because the polymers exhibit strong self-avoidance. However, collisions between blobs are abundant, since the chromatin solution is a system of closely-packed blobs. The number of inter-blob collisions scales linearly with entanglement, as every inter-blob collision contributes ~ 1 to the overall entanglement of the polymers.

Note that the average catenation $\langle \mathrm{Ca} \rangle$, is not a good identifier of polymer entanglement [19]. This is because a highly entangled system may have $\langle \mathrm{Ca} \rangle \approx 0$ where the total number of positive and negative crossings are the same and cancel each other. However, catenation-squared $\langle \mathrm{Ca}^2 \rangle$, scales linearly with the total number of nearby crossings between polymer segments [19], and is zero only for unentangled polymers.

Average volume fraction in confinement ϕ. The average volume fraction of chromatin inside a nucleus of diameter D is given as follows.

$$\phi = N(a/D)^3 \tag{4.11}$$

Transformation of chromatin to chromosome is accompanied by a strong inhomogeneity in the local volume fraction, because the volume fraction is higher at the core of the chromosome and

decays radially. However, the average volume fraction of the genome inside the nucleus remains the same (Table 4.1).

4.1.B.i Concentration blobs in a confined solution of chromosomes

The correlation length in confinement ξ_c, scales with g_c, the number of chromosome monomers in confinement: $\xi_c = \rho^{2/5}(\rho^2 R)^{1/5} g_c^{3/5}$ [Eq. (1.8)].

The number density of chromosome monomers inside a blob is the same as the average density in confinement, which furnishes the following relation for the correlation length.

$$\frac{N'}{D^3} = \frac{g_c}{\xi_c^3} \qquad \Rightarrow \xi_c = a\phi^{-3/4} \left(\frac{R_F(N)}{R_F(N')} \right)^{5/4} \tag{4.12}$$

In the unextruded state ($N = N'$), the blob size is $\approx a\phi^{-3/4}$, as expected for a semidilute polymer solution [26]. However, via loop extrusion, the chromosome polymer becomes shorter ($R_F(N) > R_F(N')$) [Eq. (4.10)], which increases the blob size in confinement. Hence, renormalization of chromatin to stiffer and more compact chromosomes result in larger confinement blobs.

Number of concentration blobs. The total number of blobs in confinement N_{blobs}, depends not only on the confinement volume fraction ϕ and the total number of microscopic monomers N, but also on the transformation of the end-to-end distance under loop extrusion.

$$N_{blobs} = N'/g_c = N\phi^{5/4} \left(\frac{R_F(N')}{R_F(N)} \right)^{15/4} \tag{4.13}$$

The number of blobs is the maximum in the unextruded state ($N = N'$); however, when the chromosomes become cylindrical brush polymers ($N' < N$), a larger blob size results in a fewer number of blobs .

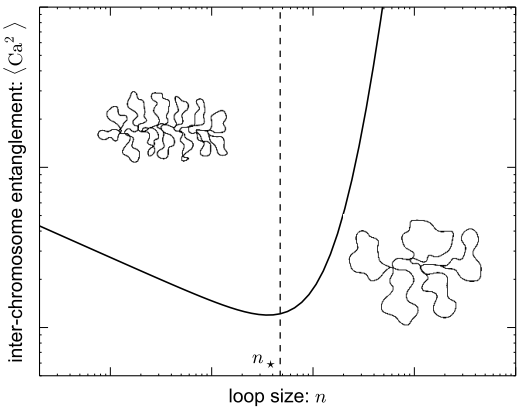

Figure 4.2: Inter-chromosomal entanglements per chromosome $\langle \mathrm{Ca}^2 \rangle$, plotted as a function of the loop size for a finite genome size N [Eq. (4.14)]. Entanglement initially decreases as loops get bigger, since the end-to-end distance decreases. However, larger loops also lead to a larger contribution from the ends of the chromosomes that increases the net end-to-end distance of chromosomes [Eq. (4.9)]. This results in an optimal loop size: $n_\star \sim N^{5/8}$, that minimizes inter-chromosome entanglements [Eq. (4.15)]. Note that the number of loops is higher for a finite genome with smaller loop sizes and vice-versa. The schematic pictures of finite-size chromosomes in small and large loop regime are shown.

4.1.B.ii Inter-chromosome entanglement in confinement

Net entanglement in the nucleus is obtained from the number of blobs scaled with the *entanglement number* N_e. Entanglement is not defined between two isolated monomers, two chains need to be of a certain number of units, $N_e \approx 100$, to be entangled [19, 136, 137]. The inter-chromosomal entanglement per chromosome

$$\langle \mathrm{Ca}^2 \rangle = \frac{N_{blobs}}{\mathrm{k} N_e} \tag{4.14}$$

scales inversely with the number of polymers in confinement or the *karyotype number* k. The average length of the chromosomes N/k, is smaller in a nuclei with a higher k, and consequently, the degree of inter-chromosomal entanglement is lower.

4.1.B.iii Optimal loops minimize inter-chromosome entanglement

Larger loops shorten the cylindrical part of the chromosomes, however, the contribution from the ends is higher [Eq. (4.10)]. Hence, for a given genome size and karyotype, there is an optimal loop size $n_\star$, that minimizes both the end-to-end distance of chromosomes and their inter-chromosomal entanglements (Fig. 4.2).

$$\left(\frac{\partial\langle \mathrm{Ca}^2\rangle}{\partial n}\right)_{n_\star} = 0 \quad \Rightarrow n_\star = \frac{(d/a)^{7/8}}{w^{1/8}\alpha^{1/4}}\left(\frac{N}{\mathrm{k}}\right)^{5/8} \tag{4.15}$$

The optimal loop size is larger for larger genomes, but decreases for chromosomes with a thicker core (higher α). Chromosomes with a lower interloop distance (d) are more compact and require smaller loops to minimize entanglement.

4.1.B.iv Chromosome self-entanglement

The number of blob collisions inside a chromosome is a linear scale to measure self-entanglement. However, contrary to the confinement-induced blobs used to compute inter-chromosome entanglements [Eq. (4.12)], the blobs contributing to self-entanglement: $\xi = aw^{-1/4}\langle\varphi\rangle^{-3/4}$, are associated with the average volume fraction of loops: $\langle\varphi\rangle = na^3/(R^2 d)$.

Self-entanglement of a chromosome is thus obtained using the number of nucleosomes g, per blob of size $\xi = R_F(g)$ [Eq. (1.8)], as follows:

$$\langle \mathrm{Ca}^2\rangle_{self} = \frac{N}{\mathrm{k}N_e g} = \frac{w^{13/64}\alpha^{45/32}}{N_e(d/a)^{75/64}}\left(\frac{N}{\mathrm{k}}\right)^{39/64} \tag{4.16}$$

where we have used the optimal loop size corresponding to minimum inter-chromosome entanglements [Eq. (4.15)]. Note that self-entanglement is higher for a more compact state, i.e., lower interloop distance and higher loop valency.

4.2 Results

4.2.A Chromosome structure

4.2.A.i Loop extrusion controls chromosome structure

Active manipulation of the three structural parameters: loop size n, interloop distance d, and loop valency α, allows control over the compaction state of the chromosomes [Fig. 4.1(b)]. Loop extrusion provides a mechanism for directed activity at the loop bases, such that the loop-extruding complexes, that establish the loop-anchoring point, may also drive a processive increase in the number of loop monomers [20, 22, 125, 128]. This processivity leads to a more compact chromosome configuration, where large loops are separated by a small interloop distance: $d < R_F(n)$, resulting in overlap between adjacent loops. The level of compaction may be increased by division of bigger loops into smaller ones, i.e., via increasing loop valency α, which may also be driven by loop extrusion [22, 129]. The mechanism of loop extrusion may be used to attain a highly compact chromosome state corresponding to a fully stretched backbone ($d \approx a$), however, even higher compaction may be achieved by driving $d < a$ [Eq. (4.10)]. It has been recently proposed that mitotic chromosomes may possess a helical backbone [129], which may achieve $d < a$, however, the microscopic mechanism of establishing such a structure is not clear at the moment.

Directed activity at the loop bases is essential to transition from a less-compact chromosome state to a highly compact one, as occurs during the cell cycle. Proteins of the family Structural Maintenance of Chromosome (SMC) are pivotal to chromosome structure [138, 139], and recently have been strongly implicated as the enzymes that drive loop extrusion via ATP hydrolysis [140–142]. In order to extrude loops, these enzymes need to generate forces in the range $k_BT/a \approx 1$ pN, that corresponds to stretching the entropic degrees of freedom in the genome [Eq. (4.3)].

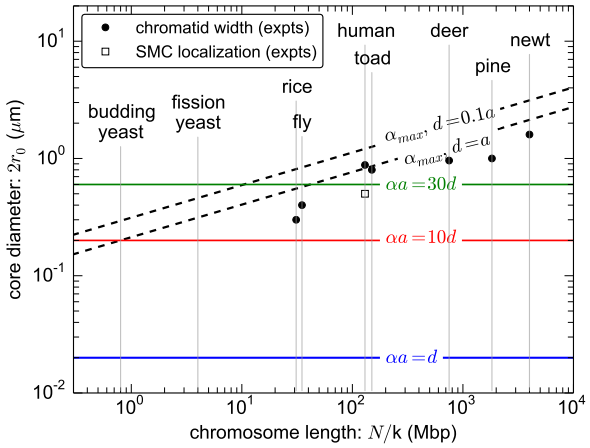

Figure 4.3: Core diameter $2r_0$, is larger for chromosomes with loops of higher valency α and/or lower interloop distance d (green and red solid lines, respectively corresponding to $\alpha = 30,\ d = a$ and $\alpha = 10,\ d = a$). A minimal core corresponds to tandem monovalent loops (blue line: $\alpha = 1,\ d = a$). The dashed lines correspond to a maximum core size: $\max\{r_0\} \sim (N/\mathrm{k})^{5/18}$ [Eqs. (4.6), (4.15)] for two interloop distances: $d = a$ and $d = 0.1a$; the core width is maximized for chromosomes with optimal loops of maximum valency, i.e., saturated chromosomes. The dashed lines show that larger chromosomes can afford a thicker core. Black circles represent experimental data for chromatid width, which is an upper bound for the core diameter (Table 4.3). Open square represents experimental observation of SMC co-localization width during prometaphase [143]. SMC proteins possibly play an important role in generating the compaction pressure required to maintain a core, thus a higher width associated with co-localization of SMCs on the chromatid axes correlates with a thicker core.

4.2.A.ii Compact chromosomes have a dense cylindrical core

Loop extrusion increases n and decreases d, which leads to a nucleosome density that is highest along the cylindrical backbone and decays radially outwards [Fig. 4.1(c)-(d)]. High volume fraction of nucleosomes near the loop base builds a cylindrical core along the chromatid axes, that is composed of closely-packed chromatin. The volume of the core is completely filled with nucleosomes, such that there is no entropic freedom of chromatin within the core. The osmotic pressure inside the core: $k_BT/a^3 \approx 1$ kPa, needs to be counteracted by an *active-compaction pressure* to keep the core from relaxing under the osmotic stress. Proteins such as SMCs and TopoII are known to co-localize along the chromatid axes during mitosis [143, 144], and possibly play a role in building the core via generating this compaction pressure. SMCs may generate compaction pressure by loop extrusion, however, the contribution of TopoII to the osmotic

pressure of compaction is poorly understood [145]. Higher cross-linking of chromatin inside the core, possibly by both non/histone-like proteins, can also help maintain a higher compaction pressure. There is experimental evidence that mitotic nucleosomes tend to aggregate more than interphase ones, due to post-translational modifications to their histones [146]. This hints at cell-cycle-dependent chemical modification of chromatin as a possible mechanism to stabilize the compact core.

Higher loop valency results in a thicker core. Increasing the valency of chromosome loops is a mechanism to build the core, since the core radius scales linearly with loop valency [Eq. (4.5)]. The core radius is minimum (equal to one nucleosome diameter a) for a tandem arrangement of monovalent loops ($\alpha = 1,\ d = a$) (Fig. 4.3), and becomes thicker for a higher loop valency ($\alpha > 1$, Fig. 4.3). This is the result of a higher nucleosome density near the base of a loop with a higher valency. Since the nucleosome density is already maximum in the core, increasing loop valency cannot further increase the nucleosome density in the core, but has to increase the volume of the core, leading to a thicker cross-section. The upper limit of loop valency corresponds to the core spanning the entire chromosome width, which we call "saturated chromosomes". The maximum valency scales positively with loop size: $\alpha_{max} \sim \sqrt{n}$ [Eq. (4.6)]. The dashed lines in Fig. 4.3 corresponds to saturated chromosomes with different interloop distances. Saturated chromosomes correspond to the maximally compact state achievable by loop division.

Simulation studies have shown that division of larger loops into smaller ones via loop extrusion can drive compaction [22, 125, 129]; such a mechanism can also build a core via increasing the loop valency. It may be interesting to compute the radial variation of monomer density in the simulated structures, which may be able to see a regime of shallower radial decay of monomer concentration corresponding to the chromosome core. The nucleosome concentration profile predicts a radially decreasing fluorescence intensity in chromatin-stained mitotic chromosomes, which appears to be qualitatively consistent [139, 143, 144]. An experimental quantification of the radial decay could be interesting, however note that the fluorescence intensity may not scale linearly with nucleosome density. The radial dimension of SMC co-localization along the back-

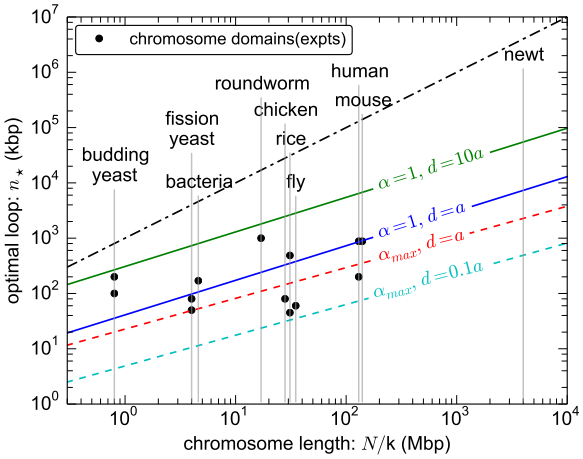

Figure 4.4: Optimal loop size $n_\star$, versus average size of chromosomes for various loop valencies α and interloop distances d are shown. Optimal loops minimize chromosome end-to-end distance and inter-chromosomal entanglements, such that the optimal loops are bigger for larger chromosomes. Note that the loop size $n_\star$ is smaller for a thicker core, i.e., higher α and lower d, because a thicker core drives a stronger decrease in chromosome end-to-end distance and entanglement [Eqs. (4.10), (4.14)]. The power law behavior of $n_\star$ for saturated chromosomes (α_{max}): $n_\star \sim (N/\mathrm{k})^{5/9}$ (dashed lines), is slightly weaker than the case of chromosomes with a smaller core: $n_\star \sim (N/\mathrm{k})^{5/8}$ (solid lines). The dot-dashed curve depicts a loop size equal to the chromosome size, hence loop sizes larger than this are not accessible. The gray lines indicate the chromosome sizes of various organisms, and the black circles are the typical TAD (Topologically Associating Domains) or chromosome domain sizes from experiments studying physical contacts within chromosomes (Table 4.2).

bone is a correlative measure of the core width [143, 144]. A higher SMC-width indicates that the compaction pressure is maintained over a larger cylindrical volume, resulting in a thicker core (Fig. 4.3).

4.2.B Coupling of chromosome structure and topology

4.2.B.i Lower inter-chromosome entanglement in a higher karyotype

Nuclear chromatin is highly entangled, because the volume fraction of nucleosomes in nuclear confinement ϕ, is more than two orders of magnitude higher than the overlap threshold $\phi^\star$ (Table 4.1). The end-to-end distance of a chromosome is maximum in the unextruded state, which maximizes entanglements in unextruded chromatin [$N = N'$, plotted in Fig. 4.5(a)]. However,

entanglements per chromosome is lower for nuclei with a higher karyotype: $\langle Ca^2 \rangle \sim 1/k$, because the average length of a chromosome is shorter leading to a lower number of contacts with other chromosomes [Eq. (4.14)].

Dilemma of Yeast. An interesting example to study the effect of karyotype on chromosome entanglements is the contrast between budding and fission yeasts: both have the same genome size, but budding yeast has a higher karyotype, and consequently, a smaller average chromosome length compared to fission yeast (Table 4.1). Inter-chromosome entanglement, $\langle Ca^2 \rangle \sim \phi^{5/4}/k$ [Eq. (4.14)], is expected to be lower for budding yeast ($k = 32$) than fission yeast ($k = 6$), because of a higher karyotype number. However, fission yeast nuclei are larger (lower ϕ, Table 4.1), leading to an entanglement value similar to that of budding yeast [Fig. 4.5(a)]. Proper segregation of chromosomes is likely a strong selective pressure in evolution, since chromosome mis-segregation can be a lethal stress [147]. This suggests that an increase in the size of fission yeast nuclei could be an evolutionary response to its lower karyotype number.

4.2.B.ii Optimized loop extrusion drives chromosome segregation

Increasing the loop size decreases inter-chromosome entanglements as long as the loops are smaller than $n_\star$ [Eq. (4.15)], beyond this critical value, entanglement increases with the loop size (Fig. 4.2). This feature of loop-extruded chromosomes comes from their finite size, i.e., the two rounded ends capping the cylindrical volume. Longer loops decrease the axial contour length of the cylindrical part of the chromosomes, however, longer loops also make the ends larger, contributing positively to the net end-to-end extension of the chromosomes [Eq. (4.10)]. Hence, there is an optimal loop size that minimizes the end-to-end distance of the chromosome polymer. Inter-chromosomal entanglements scale positively with the end-to-end extension [Eq. (4.14)], as a result, optimal loops also minimize entanglements [Fig. 4.2, Eq. (4.15)].

Chromosome disentanglement may be driven via an optimized loop extrusion process that maintains a chromosomal-loop length comparable to the optimal value: $n_\star$ (Fig. 4.4). This optimization may be achieved by titrating the concentration of loop-extruding enzymes with

respect to the average chromosome size. Note that the optimal loops are independent of the chromatin volume fraction in the nucleus ϕ. This suggests that a variation in the nuclear volume, although changes the overall level of inter-chromosomal entanglement [Fig. 4.5(d)], does not affect the optimal loop length required to minimize entanglements.

Optimal loops are larger for larger genomes. The optimal loop size, plotted as a function of the average chromosome length (Fig. 4.4) has a positive slope indicating that a larger loop is required to minimize entanglements between larger chromosomes: $n_\star \sim (N/\mathrm{k})^{5/8}$ [Eq. (4.15)]. Note that the optimal loop size is lower for chromosomes with loops of higher valency and lower interloop distance. This is because formation of a core via increasing the loop valency and decreasing the interloop distance leads to a higher loop extension which increases entanglement.

Entanglement removal for chromosomes with a tandem arrangement of optimal monovalent loops. Inter-chromosomal contacts, leading to entanglements, reduce upon extruding the linear genome into loops, because the end-to-end distance of the extruded chromosome is smaller. Inter-chromosome entanglement scales positively with the chromosome length for optimal monovalent loops: $\langle \mathrm{Ca}^2 \rangle \sim (N/\mathrm{k})^{211/256}$ [Eq. (D.13)]. Entanglement for chromosomes with a minimum core ($\alpha = 1$, and $d = a$), which also corresponds to the configuration of tandem monovalent loops that are optimal, is plotted in Fig. 4.5(b). Note the dashed line (Fig. 4.5) that corresponds to $\langle \mathrm{Ca}^2 \rangle = 1$. For points in the phase space below the dashed line (e.g., yeast), inter-chromosomal contacts are rarer and the chromosomes are less entangled. Loop extrusion drives the dashed line towards higher volume fraction and higher chromosome lengths [i.e., towards the upper right corner of the contour plots in Fig. 4.5(b)-(c)], depicting a decreasing inter-chromosomal entanglement in higher eukaryotes. This is a consequence of the transformation of chromatin into a loop-extruded polymer brush that has larger and fewer monomer units [Eq. (4.8)].

Compaction via core formation drives further disentanglement of chromosomes. Formation of a chromosome core, via increasing the valency of loops and/or decreasing the interloop distance, decreases the end-to-end distance of the chromosomes [Eq. (4.10)]. Loop valency greater than unity and smaller than the maximum allowed value: $1 < \alpha < \alpha_{max}$, leads to a semi-saturated

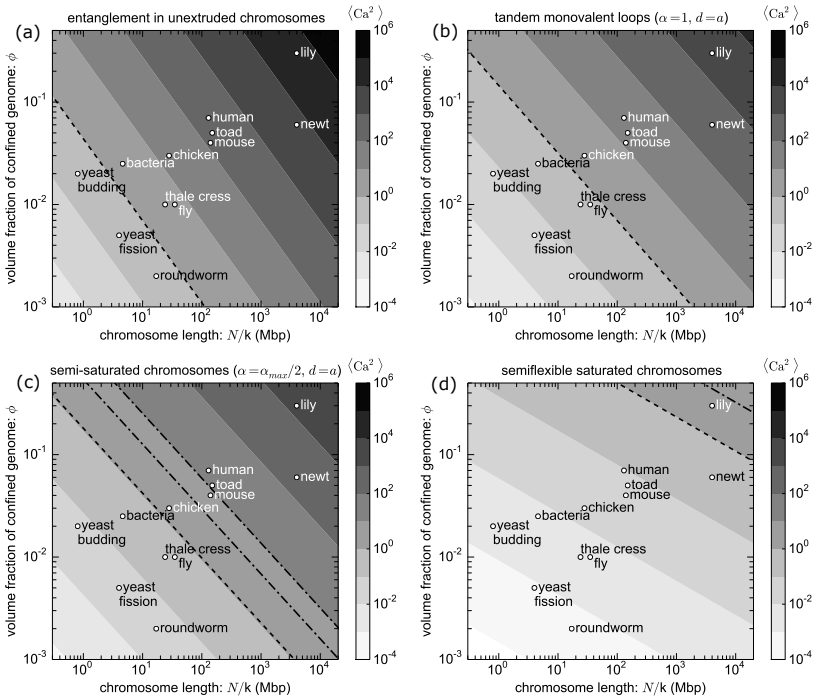

Figure 4.5: Entanglement removal by loop extrusion in steady state. Inter-chromosome entanglement $\langle \mathrm{Ca}^2 \rangle$ is shown as a contour plot for various chromatin volume fractions ϕ, and chromosome sizes N/k, where the lighter (darker) end of the spectrum corresponds to lower (higher) entanglement. Various organisms are placed in the $(\phi, N/\mathrm{k})$-phase space, see Table 4.1, where the dashed line corresponds to $\langle \mathrm{Ca}^2 \rangle = 1$. Hence, organisms lying below the dashed line have essentially disentangled chromosomes. (a) Entanglement in the unextruded state. Lower eukaryotes, e.g., fungi (yeast) and nematodes (roundworm), have a low entanglement in unextruded chromatin. This results from either a large nuclei or a small genome (Table 4.1). (b) Inter-chromosomal entanglements are lower in the loop-extruded state with a minimum core ($\alpha = 1,\ d = a$), than that in the unextruded state. This suggests that a minimum core is enough to segregate the chromosomes in lower eukaryotes and bacteria. Loop-extruded chromosomes with a minimum core possibly corresponds to prophase chromosomes in the eukaryotic cell cycle. (c) Semi-saturated chromosomes with a loop valency half of the maximum value ($\alpha = \alpha_{max}/2,\ d = a$) show further removal of entanglements via forming a core; although, the entanglement level is not much smaller than that for the minimal core [compare with (b)]. Inter-chromosome entanglement can be reduced further via nuclear envelope breakdown (NEB) due to an increase in the volume accessible to chromosomes, i.e., a decrease in the confinement volume fraction. The dot-dashed lines show $\langle \mathrm{Ca}^2 \rangle = 1$ corresponding to 3 and 6-fold increase in the volume, where the dot-dashed line on the further right corresponds to the 6-fold case. (d) Saturated chromosomes, such that chromosomal loops feature maximum valency, are stiff or semiflexible polymers where the end-to-end is linear with its contour length [Eq. (4.17)]. Saturated semiflexible chromosomes, possibly corresponding to eukaryotic metaphase chromosomes, show a strong removal of entanglements due to their stiff rod-like behavior.

state of chromosomes, where inter-chromosomal entanglements are only a little lower than the monovalent configuration, because of a weak power law: $\langle \mathrm{Ca}^2 \rangle \sim \alpha^{-15/128}$ [Eq. (D.13)]. However, nuclear envelope breakdown (NEB) causes a decrease in the confinement volume fraction which decreases inter-chromosomal entanglements in semi-saturated chromosomes [Fig. 4.5(c)], that possibly correspond to eukaryotic prophase chromosomes.

4.2.B.iii Semiflexibility of saturated chromosomes drives disentanglement

The maximum valency of an optimal loop: $\alpha_{max} \sim \sqrt{n_\star}$, corresponds to a case where the core spans the entire chromosome. These saturated chromosomes are compact cylindrical structures that are semiflexible or rod-like stiff polymers, such that their end-to-end extension scales linearly with the contour length. Hence, using Eq. (4.9), we get the following for semiflexible saturated chromosomes.

$$\left(\frac{R_F(N')}{R_F(N)}\right)_{sat} = N^{2/5}\frac{d}{na} + \left(\frac{\alpha a}{d}\right)^{2/5}\left(\frac{n}{N}\right)^{3/5} \tag{4.17}$$

The inter-chromosome entanglement, that scales positively with end-to-end extension: $\langle \mathrm{Ca}^2 \rangle \sim R_F(N')^{15/4}$ [Eqs. (4.13), (4.14)], obtained from the above equation, is also minimized by $n_\star$ [Eq. (4.15)]. However, in the semiflexible limit, the scaling of entanglement with chromosome length is substantially weaker: $\langle \mathrm{Ca}^2 \rangle \sim (N/\mathrm{k})^{5/12}$, resulting in a strong drive towards disentanglement for saturated chromosomes [Fig. 4.5(d)]. The persistence length of saturated chromosomes are comparable to their contour length, resulting in a semiflexible-rod-like statistics, that decreases entanglement.

4.2.B.iv Open mitosis aids faithful chromosome segregation

Higher eukaryotes, like mammals, have *open mitosis*, i.e., their nuclear envelope completely dissolves during mitosis and reemerges after the chromosomes have segregated [148–150]. Contrarily, some lower eukaryotes, like fungi (*S. cerevisiae* and *S. pombe*) and molds (*D. discoidieum*), have *closed mitosis*, where the nuclear envelope remains intact during mitosis, or in some cases, partially dissolves (*semi-open mitosis*) [148]. The cause of this divergence in the evolution of

early eukaryotes remains a puzzle, however, high fidelity of chromosome segregation may be an important determinant.

Nuclear envelope breakdown (NEB) helps chromosome disentanglement. Entanglement between semi-saturated chromosomes inside the nucleus is not negligible in higher eukaryotes [e.g., see human or newt in Fig. 4.5(c)]. However, a decrease in the volume fraction due to NEB decreases entanglement [dot-dashed lines in Fig. 4.5(c)], enabling higher fidelity of chromosome segregation. Driving compaction to establish a saturated state will lead to disentanglement of chromosomes even inside the nucleus [Fig. 4.5(d)], nonetheless, disentanglement of smaller chromosomes in organisms, such as fly or thale cress, may be attained in the semi-saturated state via NEB [Fig. 4.5(c)]. Larger genomes, such as newt or lily, however, requires high level of compaction to drive chromosome disentanglement [Fig. 4.5(d)].

Lower eukaryotes need not dissolve their nuclear envelope to disentangle chromosomes. For smaller genomes (e.g., yeast, see Table 4.1 and Fig. 4.5), the chromosomes are less entangled and can be segregated faithfully inside the nucleus. Figure 4.5(a) suggests that budding and fission yeasts can afford a closed form of mitosis, since entanglements per chromosome are low ($\langle \mathrm{Ca}^2 \rangle < 1$), and interestingly, these organisms do exhibit closed mitosis [149, 150]. This hints at faithful chromosome segregation as a metric of evolutionary selection influencing whether mitosis features NEB.

There are, of course, other selective pressures contributing to deciding the fate of the nuclear envelope during mitosis, one of them being the location of the spindle-pole bodies (SPBs) that organize the microtubules at the spindle poles [148, 149]. Exo-nuclear positioning of SPBs will necessitate NEB. Roundworms (*C. elegans*) can afford closed mitosis (inter-chromosomal entanglement is low, see Fig. 4.5); however their SPBs reside outside the nucleus, making NEB essential for mitosis. In *C. elegans*, NEB occurs only in late mitosis (anaphase), and the rupture in the envelope is partial (semi-open) and localized near the spindle poles [151]. Interestingly, fungi, like yeast, also have their SPBs outside the nucleus, which are 'fenestrated' into their nuclear envelope [148, 149, 152, 153], thus enabling intra-nuclear chromosome segregation .

Flies (*D. melanogaster*), also a case with low inter-chromosomal entanglements in the nuclei, shows partial NEB in late mitosis during early developmental stages [150, 151, 154]. Nuclear envelope dis/assembly depends on, among others, the nuclear lamin proteins and the nuclear pore complexes, which may also play crucial roles in determining whether an organism has open mitosis [149].

4.2.B.v Chromosome self-entanglement is higher in the compact state

Self-entanglement scales linearly with the number of chromosome self-contacts, which has a positive scaling with the average nucleosome volume fraction inside the chromosomal structure: $\langle \mathrm{Ca}^2 \rangle_{self} \sim \langle \varphi \rangle^{5/4}$ [Eqs. (4.13), (4.14), (4.16)]. As a result, a higher compaction state, associated with a higher average nucleosome concentration in its cylindrical volume, has a higher self-entanglement. The level of self-entanglement is higher for chromosomes with a thicker core (higher α) and a larger size (higher N/k) [Eq. (4.16)]. Higher self-entanglement is expected in more compact chromosomes, because the chromatin mesh size is lower, leading to a larger number of self contacts.

4.2.B.vi Synchronized core formation drives sister-chromatid segregation

Higher chromosome self-entanglement indicates that the sister chromosomes are more entangled, which is detrimental to their segregation. This necessitates resolution of sister-chromatid entanglements before a high degree of compaction is established via formation of the core. The compaction machinery that drives a thicker core must ensure that the sister chromosomes are not "fused" at the core, leading to their mis-segregation. This may be achieved by modulating the rate of compaction, so that the topology-resolution machinery (TopoII) is able to resolve sister-chromosome entanglements in a timely manner.

Processive loop extrusion initially establishes a minimal core for each sister chromosome, associated with a tandem arrangement of monovalent loops ($\alpha = 1,\ d = a$). However, the two cores may be heavily intertwined at this stage, and further compaction by increasing loop valency

may fuse the cores into a single unit of densely packed chromatin. Note that the two intertwined cores repel each other due to a high osmotic pressure in the region of their overlap [45, 125]. This repulsive force can drive physical segregation of the two chromatids in presence of topology fluctuations mediated by TopoII. However, segregation may be hindered if core thickening is driven faster than the disentanglement dynamics of the sisters chromatids.

The net repulsive force between the two sister-chromosome cores is obtained from the total osmotic pressure per unit cross-sectional area of overlap: $f_{rep} = \Pi R L'/\mathrm{k}$, where $\Pi = k_B T/\langle\xi\rangle^3$ is the osmotic pressure derived from the average correlation length inside chromosomes: $\langle\xi\rangle = a w^{-1/4}\langle\varphi\rangle^{-3/4}$, which depends on the average volume fraction $\langle\varphi\rangle = n a^3/(R^2 d)$. The net repulsive force between the backbone cores per sister-chromatid pair is given as follows:

$$f_{rep} = \frac{k_B T}{a}\frac{\alpha^{67/32} w^{3/64}}{(d/a)^{101/64}}\left(\frac{N}{\mathrm{k}}\right)^{9/64} \tag{4.18}$$

where we have used the optimal loop criterion [Eq. (4.15)]. For human chromosomes $N/\mathrm{k} \approx 130$ Mbp, with a minimal core ($\alpha = 1,\ d = a$), the repulsive force is ≈ 2 pN. However, the repulsive force has a strong positive scaling with loop valency, indicating large repulsion between thick cores, derived from a higher osmotic stress.

This mechanism is sufficient to drive sister chromatid segregation, however, a concerted action of the topology manipulating machinery (TopoII) and the compaction machinery is crucial. Although the repulsive force increases with core thickness, disentanglement dynamics of intertwined sister-chromatid cores becomes exceedingly slow as the cores become thick. This is a direct consequence of the high polymer viscosity in concentrated solutions, suggesting that the time scale for thick cores to pass through each other (facilitated by TopoII activity) can be prohibitively large. However, a compaction synchronized with the rate of topology fluctuations by TopoII, that allows enough time for the two cores to segregate before they undergo compaction and thickening, can lead to individualized chromatid cores. Once the cores are disentangled, compaction drives an increase in the repulsive force leading to their gradual entropic segregation. There is experimental evidence of a cross-talk between TopoII and the SMC protein Condensin,

coordinating sister chromatid segregation [155].

4.2.B.vii Chromosome domains in interphase reduce inter-chromosome entanglements

Experiments studying the three-dimensional conformation of the genome have observed a high probability for certain non-neighboring genomic loci to be in physical contact [156, 157]. Such contacts within a chromosome are abundant and establish chromatin loops. The genomic distances between the contact points show a peaked distribution, indicating a typical loop size for organisms; these loops are termed, "Topologically Associating Domains" (TADs), or self-interacting domains, or chromosome interaction domains [156–162].

The black circles in Fig. 4.4 represent the typical TADs observed in experiments, which indicate that the TAD sizes compare with that of the optimal loops. Larger genomes have larger optimal loops, and the TADs for organisms with larger chromosomes are also larger (Fig. 4.4); this strongly suggests a possible role of TADs in minimizing inter-chromosomal entanglements. TADs are prominent during interphase, and the sequence-specificity of the TAD loops are known to play an important role during gene transcription and DNA replication; however, TADs also ensure that the interphase genome is maintained in a less entangled state.

Note that the TAD organization is gradually lost during mitosis [129, 163], this necessarily implicates a loss of the sequence specificity at the base of the loops, not the loop organization itself. Entry into mitosis triggers compaction of the chromosomes, which can be achieved by forming a core that involves dividing the loops and bringing the loop bases closer. Formation of the core generates many new genomic contacts, however the statistics of these contacts need not correlate with the genomic sequences as gene transcription ceases during mitosis. A reappearance of the sequence-specific TAD structure in the interphase of the next cell cycle necessitates a disassembly of the core region; and indeed, interphase chromosomes do not appear as rigid cylindrical rods. Extrusion of loops needs to follow core disassembly to reestablish the TAD organization [128].

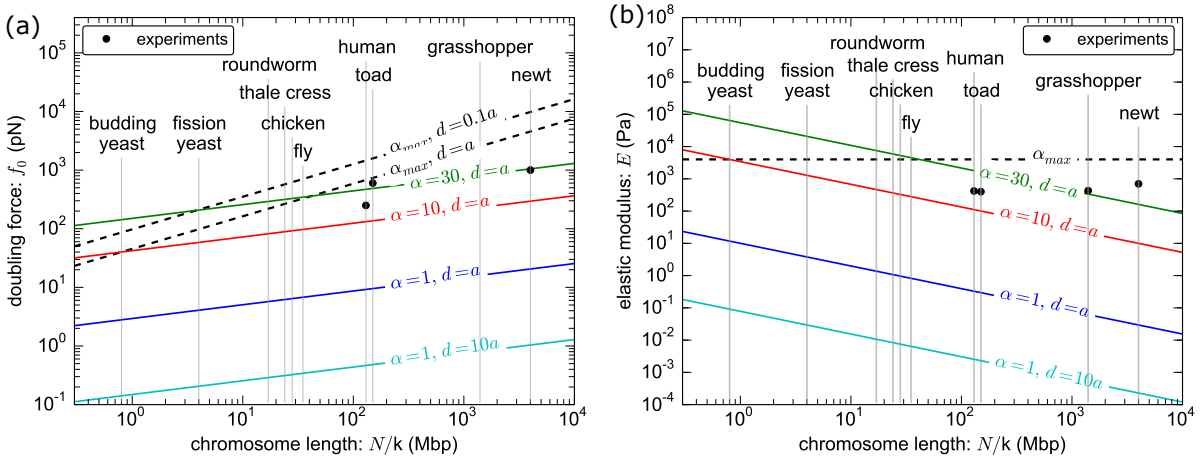

Figure 4.6: (a) Chromosome doubling force f_0 [Eq. (4.20)], defined as the interpolated force associated with doubling the length of the chromosomes, plotted as a function of chromosome length for optimal loops. The doubling force is higher for chromosomes with a thicker core, i.e., higher α and lower d (solid lines). The dashed lines, corresponding to chromosomes with a maximum core or saturated chromosomes (α_{max}), show the upper limit of the doubling force. (b) Elastic moduli E [Eq. (4.19)], derived from the average osmotic pressure inside the chromosomes, is higher for thicker cores (solid lines), and is maximum for a saturated chromosome $\approx k_BT/a^3 \approx$ 4 kPa (dashed line). The negative power law of E with the genome size for a constant α and d is a result of a lower average monomer density inside larger chromosomes with a thinner core. Equivalently, it states that larger chromosomes require a thicker core to maintain their elastic moduli. The black circles correspond to micromechanical experiments on metaphase chromosomes [43, 120–124], see Table 4.3.

4.2.C Chromosome elastic rigidity

Core imparts mechanical rigidity to chromosome The elastic moduli of mammalian metaphase chromosomes is ∼kPa, and requires ≈ 0.1 nN forces for appreciable stretching [43, 122, 123]. Loop-extruded chromosomes have a cylindrical rod-like morphology, however, without a core the they lack the structural rigidity observed in metaphase chromosomes.

Elastic modulus of chromosomes E. The internal osmotic pressure, derived from the average correlation length: $\langle\xi\rangle = aw^{-1/4}\langle\varphi\rangle^{-3/4}$, serves as an estimate of elastic moduli of the chromosomes [26].

$$E = \frac{k_BT}{\langle\xi\rangle^3} = \frac{k_BT}{a^3}\frac{\alpha^{81/32}}{w^{15/64}(d/a)^{135/64}}\left(\frac{N}{\mathrm{k}}\right)^{-45/64} \tag{4.19}$$

The above expression corresponds to optimal loop sizes, and is plotted in Fig. 4.6. The elastic

moduli is higher for chromosomes with a thicker core, since the average volume fraction inside the chromosome is higher. The negative power law dependence of the elastic moduli on chromosome sizes indicates that a thicker core is required in larger chromosomes to maintain its elastic moduli. Note that the maximum elastic moduli corresponds to $\langle\varphi\rangle \approx 1$, i.e., saturated chromosomes: $E_{sat} = k_B T/a^3 \approx 4$ kPa. However, in case of a core smaller than the maximum size, chromosomes are highly heterogeneous and the elastic modulus is not well defined.

Chromosomes with a thicker core have a higher doubling force. The doubling force of chromosomes is an intrinsic force constant reflecting the extrapolated force associated with doubling the length of the chromosomes [120, 122]. We define the doubling force as $f_0 = d(\partial^2 F/\partial d^2)$, where F is the free energy per loop [Eq. (D.4)], furnishing the following for optimal loops.

$$f_0 = \frac{k_B T}{a} \frac{\alpha^{37/32} w^{5/64}}{(d/a)^{83/64}} \left(\frac{N}{\mathrm{k}}\right)^{15/64} \tag{4.20}$$

Figure 4.6 shows the doubling force as a function of average chromosome length for optimal loops, where the black circles represent experimental observations.

Chromosomes with a minimum core (tandem monovalent loops, $\alpha = 1$, $d = a$) have a low doubling force ($\approx$ 10 pN), whereas, the chromosomes with a thicker core (higher α) are more rigid to stretching perturbations, since $f_0 \sim \alpha^{37/32}$ (Fig. 4.6). The origin of this elastic response is the densely packed chromatin inside the core, where the osmotic pressure is the highest. The curve corresponding to an optimal saturated chromosome (dashed lines in Fig. 4.6) suggests that organisms with a larger genome can afford a thicker core, $f_0 \sim (N/\mathrm{k})^{5/9}(a/d)^{1/3}$, which may explain the higher stretching forces in newts ($\approx$nN) compared to that in humans ($\approx$0.1 nN) [120, 123]. Note that the high internal osmotic pressure in the core volume must be actively maintained, and the candidate proteins for this activity are SMC complexes and various chromatin crosslinking proteins including histones. Non-specific proteolysis of mitotic chromosomes leads to a smaller elasticity to stretching perturbations [164]. This is consistent with our results, because digestion of proteins in the chromosomes will lead to a lower compaction pressure yielding a lower doubling force. The doubling force is also expected to be lower for a

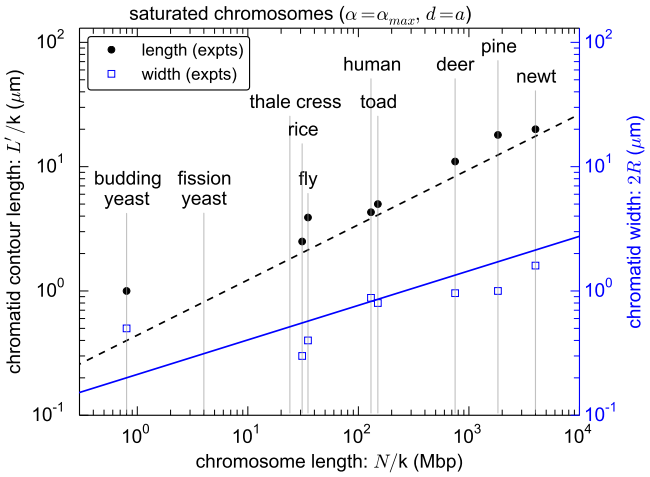

Figure 4.7: Contour length of chromatids (L'/k) and their widths ($2R$) are plotted [Eqs. (D.18), (D.19)] on the left and right y-axes, respectively. The dashed black line shows the contour lengths, whereas, the solid blue line shows the widths of optimal saturated chromosomes (α_{max}, $d = a$). Larger chromosomes have longer contour lengths and widths, however, chromatid length has a steeper scaling with genome size than width, indicating a higher length-to-width aspect ratio for larger chromosomes. The black circles and blue squares are experimental observations [165–174] for the contour length and width, respectively, see Table 4.3.

less compact core (Fig. 4.6).

Bending rigidity of chromosomes. The bending rigidity of loop-extruded chromosomes is given by, $B = k_B T \rho$, where ρ is the persistence length [Eq. (4.7)]. For the genomes of higher eukaryotes, i.e., chromosome lengths $\approx 10^2 - 10^3$ Mbp (Table 4.1), $B \approx 10^{-21} - 10^{-22}$ N·m^2 for saturated chromosomes. Experimental values of bending rigidity, reported for mammalian chromosomes [43, 122, 123], are in good agreement.

Longer chromosomes have a higher aspect ratio. The contour lengths (L'/k) and widths ($2R$) of chromosomes with optimal loops are higher for organisms with larger chromosomes. Figure 4.7 plots the lengths and widths of optimal saturated chromosomes, where the positive scaling of the contour length with total chromosome length, $L'_{sat} \sim (N/\mathrm{k})^{4/9}$, is steeper than that for the width $R_{sat} \sim (N/\mathrm{k})^{5/18}$ [Eqs. (D.18), (D.19)]. This indicates that the aspect ratio of the cylindrical chromosomes is larger for longer chromosomes, however, the scaling is weak: $L'_{sat}/R_{sat} \sim (N/\mathrm{k})^{1/6}$.

4.2.D Effect of solvent quality

Good solvent for chromatin is essential for chromosome segregation Good solvent for chromatin ($w > 0$) means the nucleosomal monomers have a preference to be in contact with the solvent than other nucleosomes. The solvent quality may deteriorate leading to a weak self-adhesion among nucleosomes that decreases the intensity of their self-avoidance, which corresponds to $0 < w < 1$. In plotting our results, we have always used $w = 1$; in what follows we discuss the scaling with the excluded-volume parameter and study the effect of deteriorating solvent quality.

Inter-chromosomal entanglements scale negatively with the excluded-volume parameter w: $\langle \mathrm{Ca}^2 \rangle \sim w^{-3/32}$ [Eqs. (4.10), (4.14)]. This suggests that entanglements increase when there is higher self-adhesion among nucleosomes, which can be understood as the effect of random sticking of nucleosomes belonging to different chromosomes. Moreover, the repulsion between sister-chromatid arms is lower for a lower excluded volume: $f_{rep} \sim w^{3/64}$ [Eq. (4.18)], which emphasizes the requirement of low nucleosome self-adhesion. Strong self-adhesion in chromatin promotes attraction between the sister chromatids, as well as different chromosomes, leading to a highly entangled state that disfavors their segregation [45].

Surfactant activity helps chromosome segregation. Surfactants can drive further compaction of the chromosomes via reducing the excluded volume of chromosome monomers in a manner independent of the microscopic excluded volume. Surfactant activity may be studied using a surfactant parameter w_s, such that $R_F(N') \sim w_s^{1/5}$, which is independent of nucleosome self-adhesion parameter w [Eq. (1.8)]. Higher surfactant activity reduces w_s, which leads to a smaller entanglement between chromosomes: $\langle \mathrm{Ca}^2 \rangle \sim R_F(N')^{15/4} \sim w_s^{3/4}$ [Eq. (4.10)]. Upon disassembly of the nuclear envelope, mitotic chromosomes get coated with a host of proteins, like Ki-67, that form a perichromosomal layer [175–177]. Such a coating may act as a biological surfactant that reduces the excluded volume of the chromosome monomers. Ki-67, a widely used cell proliferation marker, has been proposed to act as a surfactant by building a charged layer on the chromosomal periphery [175–177].

Strong self-adhesion among nucleosomes, however, can be turned on after the chromosomes are coated with surfactants. Once there is an energy barrier to sticking of inter-chromosomal nucleosomes, such as the charge barrier effectuated by Ki-67 [175–177], a strong self-adhesive interaction can help further compact the chromosomes. A higher tendency of mitotic nucleosomes to aggregate, compared to interphase ones [146] is consistent with the idea of cell-cycle dependent modulation of chromatin solvent quality, as a mechanism to achieve compaction while not hampering segregation.

4.2.E Segregation of bacterial chromosomes

Bacteria, unlike eukaryotes, do not have nucleosomes, instead the DNA inside bacterial cells are coated with a variety of nucleoid associated proteins, like HU and IHF [178]. Bacteria, however, like eukaryotes, contain SMC proteins [such as MukBEF [179]] and TopoII, which are essential for compaction and segregation of chromosomes in all kingdoms of life. Bacterial chromosomes can be treated as a flexible self-avoiding polymer with cylindrical monomer units of length $a \approx 50$ nm, corresponding to the persistence length of naked DNA [12]; and width $b \approx 4$ nm, which roughly corresponds to the thickness of a protein-bound DNA segment. All our results discussed above for chromatin loops in eukaryotes are directly applicable to bacterial chromosomes, i.e., extrusion of DNA loops lead to a polymer-brush morphology that drives disentanglement of chromosomes.

The entanglement-minimizing loops in bacteria have the same scaling relation as Eq. (4.15). However, there is an extra factor associated with the aspect ratio of the cylindrical DNA monomers: $n_\star \sim (a/b)^{1/4}$, and $\langle \mathrm{Ca}^2 \rangle \sim (a/b)^{3/20}$ [Eqs. (D.23), (D.24)]. This suggests that increasing DNA volume by covering the surface of the double helix with proteins (higher b), causes a decrease in the optimal loop size and a decrease in entanglement, although the effect is small ($a/b \approx 10$). Figure 4.4 shows the comparison of optimal loop size with experimentally observed chromosome interacting domain in bacteria (*C. crescentus*) [159], suggesting that chromosome domains in bacteria, similar to the eukaryotes, help maintain a low entanglement

level.

Bacteria do not have nuclei enclosing their genome, however the volume fraction of the genome inside the cell is quite high compared to the nuclear volume fraction of chromatin in lower eukaryotes (e.g., yeast and roundworms) (see Table 4.1). Figure 4.5(a) shows that bacterial chromosomes are entangled in the unextruded state. However, a small degree of compaction achieved via formation of a minimal core ($\alpha = 1,\ d = a$) is enough to disentangle bacterial chromosomes [Fig. 4.5(c)]. Similar to sister chromatid segregation in eukaryotes, the compaction is needed to be driven slow enough to allow for disentanglement dynamics of sister chromosomes. This is a more serious issue for bacteria, because the volume fraction of the genome in the cell is moderately high, leading to higher polymer viscosity.

Bacterial DNA is maintained in a supercoiled condition, which favors plectoneme formation [12]. Branching of plectonemes may provide a way to manipulate the size of DNA loops and their valencies. Protein-mediated kinks may also play an important role in facilitating plectoneme branching [49], thus driving the structure parameters. Bacterial cell volumes are typically asymmetric, this generates an entropic-segregation pressure gradient along the longer axis of the cell. The segregation pressure may lead to equilibrium chromosome segregation at long enough time scales [180], however, extrusion of DNA loops to form a cylindrical brush structure enhances the osmotic pressure gradient providing an active mechanism to drive segregation.

4.3 Summary

We presented a model for chromosomes, where the linear genome is arranged as chromatin loops emanating from a chromatin backbone (similar to a bottle-brush polymer) [Fig. 4.1(b)]. We analyzed a trend towards self-segregation of chromosomes driven by an active manipulation of the chromosome structure parameters: loop size n, interloop spacing d, and valency of the loops α, which denotes the number of loops emanating from one grafting point on the backbone. Active extrusion of chromatin loops [20, 22, 125] can provide an efficient mechanism to manipulate these parameters in our model that drive compaction and segregation of chromosomes.

Loop extrusion Lengthwise compaction of the genome can drive chromosome segregation [18, 19]. Loop extrusion, an active mode of lengthwise compaction, has been very successful in interpreting experiments studying the three-dimensional conformation of cellular chromosomes [128, 129, 163]. Proteins of the family Structural Maintenance of Chromosomes (SMC), like Condensin and Cohesin, play an integral role in maintaining the structure of compact metaphase chromosomes; and more recently, Condensins have been strongly implicated as the loop-extruding agents [140, 141]. Loop extrusion, by actively increasing the loop size and simultaneously decreasing the interloop distance [20, 22, 125], can control the compaction state of the chromosomes. Loop extrusion can also divide larger loops into smaller ones [20, 22, 125, 129], thus increasing the valency of loops, which is again a mechanism to control the compaction state of the chromosomes.

Optimal loops minimize inter-chromosome entanglement Entanglements between chromosomes are minimized for an optimal loop size that increases with the average size of chromosomes (Fig. 4.4). Entanglements are maximum in an equilibrium solution of chromatin confined in the nucleus. Extruding loops to form a polymer brush structure decreases entanglement initially, but once a threshold loop size is reached the entanglement starts to increase for larger loops (Fig. 4.2). Larger loops make the chromosome thicker which decreases end-to-end distance of a finite-sized genome, however, larger loops also increase the contribution to the end-to-end distance from the chromosome ends, leading to an optimal loop size that minimizes end-to-end extension and inter-chromosome entanglement [Eq. (4.15)]. To establish an optimal loop size, loop extrusion has to be optimized which can be done by titrating the concentration of loop-extruding agents.

Loop extrusion to build a core The optimal loop size, however, is not a constant. It increases with the genome size, and decreases with increasing loop valency [Eq. (4.15)]. The valency of the loops is an internal parameter that controls the level of compaction. Higher valency, larger number of loops emanating from one anchor along the backbone, indicates a

higher density of nucleosomes at the loop base, and leads to formation of a cylindrical core along the backbone [Figs. 4.1(c), 4.3]. The core gives the loop-extruded chromosomes the compactness and structural rigidity observed in metaphase chromosomes [Fig. 4.6]. Thus, extrusion of the existing chromatin loops to increase their valency can compact the chromosomes via building a core.

TADs reduce chromosome entanglements in interphase genome Sequence-specific physical contacts are a unique feature of the interphase genome, that establishes loops with well-defined genomic sequences at their bases, called "Topologically Associating Domains" (TADs), or self-interacting or chromosome-interacting domains [156, 157, 159]. The optimal loop sizes that minimize inter-chromosomal entanglements compare with the TAD sizes observed experimentally in interphase [156, 158, 181]. This highlights a potential role of TADs in maintaining a low entanglement level inside the nuclei, while also regulating gene transcription. The loss of TADs during mitosis [129, 163] could be a result of the core formation, which proceeds via increasing the valency of the loops irrespective of the genomic sequences at the loop bases.

Chromosome compartmentalization during interphase features a separation of the less compact, early replicating, active or "A"-compartment (euchromatin) from the more-compact, late-replicating inactive or "B"-compartment (heterochromatin) [182]. The inactive compartments, typically tethered to the nuclear lamina, may be thought of as densely-packed chromatin cores, however, the geometry of the core and the proteins responsible for maintaining the compaction pressure in the compartments are different. Chromatin cross-linking and tethering proteins, like HP1α [183], are thought to be the primary organizers of the interphase compartments.

Active topology fluctuations are essential for chromosome segregation Topology fluctuations, mediated by type-II DNA topoisomerases (Topo II), are an essential ingredient for chromosome segregation. Topo II is a locally acting enzyme that is agnostic of the global topology of chromosomes, which means, TopoII alone is incapable of segregating the chromosomes. However, by maintaining a fluctuating topology, Topo II ensures that local entanglements do

not hinder the compaction machinery (SMCs), which is able to drive a global segregation of chromosomes by lengthwise compaction.

The net rate of topology fluctuations, governed by the cellular concentrations of TopoII, can affect the time scales of compaction. Topology fluctuations by TopoII may be incorporated as an effective polymer viscosity governing the time scales of chromosome dynamics, which is left for future considerations.

Sister-chromatid segregation requires interplay between compaction and topology-manipulation activities A higher-compaction state for the loop-extruded chromosomes also have a higher self-entanglement, which hinders segregation of sister chromatids. Entanglements between the sister chromatid backbones should be resolved before the compaction machinery drives a core formation via increasing loop valency; otherwise, both the sister chromosomes may develop a heavily intertwined common core, making disentanglement dynamics prohibitively slow. Proper segregation of the sister chromatids require a modulated compaction rate, such that the sister core are duly disentangled before they become thick. Osmotic pressure from core overlap generates a repulsive force between the cores [Eq. (4.18)], however, the dynamics of the core is slow due to high nucleosome density, which again hints at the necessary interplay between TopoII and SMCs (compaction machinery) in order to disentangle the sister chromosomes.

Comparison with other models of mitotic chromosomes The view of compact metaphase chromosomes as layers of stacked chromatin loops dates back to the early electron microscopy images that led to the "radial loop" models [184, 185]. Note, that the core in our case is built entirely out of closely packed chromatin, which reproduces the mechanical properties of metaphase chromosomes (Fig. 4.6) [43, 120, 122, 123]. This is different from the scenario of a protein scaffold along the chromosome axes supporting the radial loops and providing the mechanical rigidity of chromosomes [127, 184]. However, the core in our case is actively built by packing chromatin, hence a higher density of proteins associated with the activity (like SMCs) is expected along the core [138, 143, 144]. Another view of mitotic chromosomes is that of

hierarchically folded chromatin with an "axial glue" comprised of chromatin-crosslinking proteins [186]. The saturated state of the chromosomes in our model has qualitative similarities with this view, however, the hierarchical folding in our case is obtained by extruding loops that leads to a higher axial density of chromatin. Mitotic chromosomes as self-assembled stacks of thin plates containing irregularly folded chromatin [174, 187] may provide structural rigidity but how such self-assembly of nucleosomes will lead to segregation is not clear. Polymer models of chromosomes obtained from simulating loop extrusion have the 3D conformation of cylindrical chromosomes [125, 129], and it would be interesting to study aspects of the simulated structure like the radial density profile of monomers.

We have modeled chromosomes as chromatin loops arranged on a flexible chromatin backbone. In presence of topology fluctuations, active extrusion of the loops can control the compaction state of the genome and drive segregation of the chromosomes. A high density of loops on the backbone leads to a dense chromatin core along the cylindrical axis that provides chromosomes with structural rigidity and rod-like morphology observed during mitotis. We also find that inter-chromosome entanglement is minimized for an optimal loop size that is higher (lower) in a less (more) compact genome, and scales positively with the genome length. Entanglement minimization by loop-organization of the genome hints at the potential role of chromosomal domains or TADs in maintaining a lower degree of inter-chromosomal entanglement during interphase.

Organism	Genome (Mbp)	Karyotype k	Chromosome length N/k (Mbp)	Confinement volume D^3 (μm^3)	Volume fraction ϕ	Overlap threshold $\phi^\star$
Lily (*L. longiflorum*)	97000	24	4000	1347 [188]	0.3	10^{-7}
Newt (*N. viridescens*)	95000	24	4000	4174 [189]	0.06	10^{-7}
Human (*H. sapiens*)	6000	46	130	300 [189, 190]	0.07	10^{-6}
Mouse (*M. musculus*)	5600	40	140	400 [189, 191, 192]	0.04	10^{-6}
Toad (*X. laevis*)	5400	46	150	307 [189]	0.05	10^{-6}
Chicken (*G. gallus*)	2200	78	28	210 [189]	0.03	10^{-6}
Fly (*D. melanogaster*)	280	8	35	78 [189]	0.01	10^{-5}
Thale cress (*A. thaliana*)	240	10	24	70 [193, 194]	0.01	10^{-5}
Roundworm (*C. elegans*)	200	12	17	200 [195, 196]	0.003	10^{-5}
Fission yeast (*S. pombe*)	25	6	4	12 [197–199]	0.005	10^{-4}
Budding yeast (*S. cerevisiae*)	25	32	0.8	3 [189, 200, 201]	0.02	10^{-4}
Bacteria (*E. coli*)	9.2	2	4.6	2 [202]	0.025	10^{-4}

Table 4.1: Genome size in diploid nuclei, in Mega-base pair units (Mbp=10^6 bp), and the corresponding karyotype (number of chromosomes) for various model organisms [188, 189, 203]. Average chromosome length N/k, is obtained by dividing the total genome size by the karyotype number k. Note that the number of monomers per chromosome N/k, when written in Mbp units represent the net genomic length of each chromosome. Nuclear volumes of various organisms are used to compute the average volume fraction of chromatin inside nuclear confinement, ϕ. The critical volume fraction above which there is inter-chromosome overlap, $\phi^\star = Na^3/R_F(N)^3 = N^{-4/5}$, is much smaller than the chromatin volume fraction, indicating strong inter-chromosome overlap between chromosomes inside the nucleus. Note, bacterial chromosomes are made up of cylindrical segments of length $a \approx 50$ nm and width $b \approx 4$ nm (corresponding to protein-bound DNA), such that the volume fraction is computed as $\phi = Nab^2/D^3$.

Organism	TAD size (kbp)
Human	200 [158], 880 [156]
Mouse	880 [156]
Chicken	80 [129]
Fly	60 [156, 161]
Rice	45[160], 487[162]
Roundworm	1000 [204]
Bacteria	170 [159]
Fission yeast	50 [205, 206], 80 [207, 208]
Budding yeast	100[209], 200 [181]

Table 4.2: Topologically Associating Domain or TAD sizes, as observed in chromosome conformation capture experiments, plotted in Fig. 4.4. The genome size and karyotype numbers of these organisms are given in Tables 4.1 and 4.3. Note, the bacterial TAD size is for *C. crescentus* [159] which has a genome size similar to *E. coli*.

Organism	Genome size (Gbp)	Karyotype k	Chromatid length: L'/k (μm)	Chromatid width: $2R$ (μm)	Stretching modulus: f_0 (nN)	Elastic modulus: E (kPa)
Newt (*N. viridescens*)	95 [189]	24 [189]	20 [123]	1.6 [123]	1 [123]	0.7 [123]
Pine (*P. taeda*)	44 [203]	24 [203]	18 [174]	1 [174]	–	–
Grasshopper (*C. parallelus*)	26 [169]	18 [169]	–	–	–	0.43 [124]
Deer (*M. muntjak*)	6 [170, 171]	8 [170, 171]	11 [171]	0.96 [171]	–	–
Human (*H. sapiens*)	6 [203]	46 [203]	4.3 [171]	0.88 [171]	0.25 [120]	0.42[120]
Toad (*X.laevis*)	5.4 [203]	46 [203]	5 [123]	0.8 [123]	0.6 [122]	0.4[123]
Rice (*O. sativa*)	0.75 [203]	24 [203]	2.5 [174]	0.3 [174]	–	–
Fly (*D. melanogaster*)	0.28 [203]	8 [203]	3.9 [174]	0.4 [174]	–	–
Budding yeast (*S. cerevisiae*)	0.025[203]	32 [203]	1 [210]	0.5 [210]	–	–

Table 4.3: Experimental values plotted in Figs. 4.6 and 4.7 are tabulated. The elastic modulus reported in Ref. [124] is for migratory grasshopper (*M. sanguinipes*) chromosomes, however, due to lack of genomic data on *M. sanguinipes*, we use the genomic data of meadow grasshopper (*C. parallelus*).

Chapter 5

Conclusions

DNA or chromosomes, the native state of DNA inside living cells, bears the genetic information that largely shapes cell function and fate by characterizing gene expression. We discussed a view of DNA and chromosomes as passive objects that are manipulated physically and chemically[1] by various proteins. DNA has a double-helical structure that is stabilized by stacking of complementary bases. Proteins, either by just binding DNA or via facilitating a series of events following DNA association, perturbs DNA structure that has a mechanical response. This mechanical response is often tuned to drive biological processes, making DNA mechanics a topic of interest. We presented theoretical models aimed at understanding the response of DNA and other complex DNA structures, such as protein-bound DNA or chromatin, to biologically relevant perturbations. The cell lives in an aqueous ionic buffer at a finite temperature (≈ 300 K), where screening of charge by counter-ion condensation and thermal fluctuations, i.e., statistical mechanics are important considerations in building theoretical models [Sec. 1.4]. DNA, due to the double-helical wrapping of its two strands, has an inherent topology or linking number, perturbation of which generates a mechanical response [Secs. 1.2, 1.3]. Proteins, cellular machines intended to effectuate biological processes involving DNA, are known to utilize the

[1] We did not focus on how chemical modifications affect DNA, however, DNA structure is intimately related with its chemistry. Chemical modifications to DNA base pairs, such as DNA methylation or post-translational modifications to chromatin, is expected to modify chromosome structure and mechanics, and have important biological functions.

coupling between DNA topology and mechanics to their advantage. Quantitative calibration of DNA statistical mechanics, topology, and their inter-dependence, as provided by the theoretical models described in this dissertation, is important for a fundamental understanding of *in vivo* micromanipulation of DNA, and also for a precise control of biological pathways, possibly leading to biomedical applications.

5.1 Two helically intertwined DNAs subject to tensile stretching

DNA braids, two torsionally unconstrained double helices that are helically intertwined or catenated, is a common *in vivo* occurrence. This is because cellular DNA is long and sufficiently flexible[2] to get intertwined. Under physiological stretching forces[3], DNA intertwines or catenanes adopt a helically-wrapped conformation as it minimizes the bending energy. Akin to wrapped elastic rods, DNA braids are bulkier than the individual double helices. Tight wrapping of DNAs generate a braid torque that increases nonlinearly with the number of turns or catenation [Sec. 2.2.A]. Nonlinear torque makes the twist stiffness in braids a variable quantity that depends on the geometry, such that braid twist rigidity is higher for a higher density of catenations. Braid torque, when higher than a critical value, is also responsible for buckling into a plectonemic structure. Braid plectoneme features self-writhing of the braid, the writhe linking number contribution does not contribute to torque, as a result, increasing the number of braid turns in the buckled state drives an increase in the total plectoneme length.

The braid plectoneme state is characterized by a proliferation of many domains, where a plectoneme domain is distinguished by the presence of a braid end loop, a finite-sized loop structure that associates a nucleation energy cost to a plectoneme domain [Sec. 2.2.A]. This multi-domain proliferation is a direct consequence of the bulky nature of DNA, that destabilizes the plectoneme state due to a high bend stiffness as well as a stronger excluded volume interaction

[2]Contour length of cellular DNA is ~ 1 m, whereas, DNA persistence length is $A \approx 50$ nm.
[3] Physiological forces on DNA are in 10^{-12} Newton or 1 pico-Newton (pN) range.

in DNA braids. The plectoneme nucleation cost, due to the finite-energy end loop, is also responsible for an abrupt buckling in braids that features an experimental signal of discontinuous braid extension, that was predicted and subsequently verified [Sec. 2.3]. We also experimentally verified the prediction of coexistence of multiple plectoneme domains in buckled braid.

Mechanics of braids is also affected by the distance between its two tether points or the intertether distance, such that braids with a higher intertether distance shows a stronger twist stiffening with catenation and consequently buckle at a lower catenation density [Sec. 2.2.C]. This may have important biological consequences, as the intertether distance can be a protein-controlled parameter, thus providing a mechanism for microscopic manipulation of braid mechanics. The ionic concentration of the buffer and the size of the catenated DNA molecules affect the mechanical response of DNA braids, which we also investigated.

Our model provides a statistical mechanical view of braided DNA where the mean-field structure is given by the helical geometry. We treated thermal fluctuations in DNA braids as small perturbations transverse to the mean-field state, that are suppressed under a higher stretching tension. Proximity of two double helices in the braided structure leads to electrostatic coupling of fluctuations [Sec. 2.1.A]. Less electrostatic screening at lower ionic strengths of the solution results in a stronger coupling and a higher net fluctuation. Although, the calculation of electrostatically coupled thermal fluctuations in this dissertation is more accurate than the scaling-like approach utilized in previous studies [12, 66, 68, 76, 80], a better computation is possible. Our computation assumes a uniform electrostatic coupling of all the fluctuation modes of the braid, that is derived from the average electrostatic potential of the helical structure. This is an essential simplification. Using a Debye-Hückel kernel one can, in principle, compute the coupling strengths of each fluctuation wavemode, however, analytical tractability becomes an issue.

Stability of the braid plectoneme state is governed by its bending energy and writhe linking number contribution. We assumed the braid to be a uniform cylinder when computing bending energy and writhe from the geometry [Sec. 2.1.B]. This is a good approximation especially in the

low screening limit, i.e., low salt concentration, where braids tend to behave as thick cylindrical objects. However, at higher salts, where the helical grooves of a braid are more prominent due to strong electrostatic screening, stacking of braids may provide stabilization to the braid plectoneme structure. Exact computation of the braid plectoneme geometry is challenging, an alternative strategy may be to use an effective stacking stabilization, however that interaction needs to be calibrated, possibly using simulations.

Note that our focus was limited to catenated DNAs that are torsionally unconstrained, meaning, individual double helices are not subject to twist [Sec. 2.1.A]. While catenated sister chromatids are torsionally unconstrained, due to the presence of DNA nicks or missing bases where the double helix can swivel to relax any torsional stress, intertwined twistable DNAs may occur *in vivo* due to local twist blocking from protein binding. Braided twistable DNAs also provide an interesting construct to study coupling of inter-DNA linking number or catenation with intra-DNA linking number. Adding twist to the braid Hamiltonian is a possible consideration for future modeling efforts.

5.2 Plectoneme buckling in stretched twisted double-helix DNA

Topology perturbation in double helices leads to DNA twist. Twisted DNA results in torque that increases linearly with the excess linking number [Sec. 3.1], contrasting the nonlinear torque increase with catenation in DNA braids. This is the result of base-stacking interactions that make DNA a solid object featuring a constant DNA twist stiffness, while braids are soft structures that undergo twist stiffening with increased catenation. Sufficient twisting of the double helix results in a DNA torque that drives buckling of DNA [Sec. 3.1.B.i]. The buckled state is a DNA plectoneme where the double helix writhes around itself in a helical fashion[4]. The writhe linking number contribution screens DNA twist and stabilizes the buckled state at higher torque. The buckling transition is abrupt due to nucleation of a finite-sized DNA loop

[4] DNA plectoneme has a similar structure to that of a DNA braid, both are helically wrapped DNAs, however in plectonemes the DNA is twisted and contains an end loop. Braid plectoneme features helical self-wrapping of a braid with itself, a structure bulkier than a plectoneme.

that generates a discontinuity in DNA extension at the critical linking number corresponding to plectoneme buckling. Stretching the double helix causes partial unwrapping of the plectoneme superhelix, and also leads to buckling at a higher critical linking number [Sec. 3.1.B.i]. Long DNA molecules ($>$ 6 kb) show coexistence of multiple plectoneme domains, due to entropic stabilization of plectoneme-domain diffusion along the DNA contour. DNA plectonemes are more stable than braid plectonemes due to their less-bulky structure, as a result, while braids exhibit proliferation of multiple domains in all practical scenarios, proliferation of plectonemic domains in twisted DNA is only favored for DNAs with long length and lower strength of ionic buffer. Low ionic buffers increase plectoneme radii which destabilizes them, leading to coexistence of multiple domains [Sec. 3.1.B.vii].

5.3 Buckling in DNAs with an immobile point defect

Defects on DNA, like a base-unpaired region, or a short single stranded DNA bulge influences its mechanical response. Such defects allow a sharp localized DNA bend at low energy cost and can act as preferred sites of plectoneme nucleation [Sec. 3.2]. A kink in the plectoneme end loop reduces its nucleation energy, which lowers the energy of a plectoneme with a defect placed at its tip. These plectonemes are defect-pinned because of the immobile nature of the defect and lack diffusion stabilization.

We quantified the strength or size of the defect via the amount bending energy reduction associated with the defect [Sec. 3.2.A.ii], e.g., a base-unpaired region with a larger number of unpaired bases is a larger defect as it allows a sharper kink at the defect site . Another important source of DNA defects discussed in our model is protein-mediated kinks that induce a local curvature that reduces the in-situ DNA bending cost. We find that larger defects are more efficient in trapping a plectoneme domain due to a lower nucleation cost of the defect-pinned domain. However, if the defect is too small the entropic stabilization from plectoneme diffusion makes nucleation of a defect-pinned domain less favorable. Our model predicts a three-state coexistence at the buckling point, where the stability of various possible buckling states depend

on the size of the defect [Sec. 3.2.B.i]. Experiments using DNA visualization may be able to verify the coexistence. We also quantitatively explained the observations of existing experiments on DNA with base-unpaired defects [Sec. 3.2.C]

We treated thermal fluctuations perturbatively around a helical mean-field state, similar to the scheme used for braids. A better calculation of electrostatic coupling of thermal fluctuations of the two plectonemic strands may be possible, as discussed above for braids. We considered defects to only affect the local bending stiffness and not the twist stiffness. This is a good approximation for a point defect, however, if the physical dimension of the defect is large, the twist stiffness is also expected to be lowered. An interesting approach can be to treat the large defect as a coexistent state with altered mechanical properties, which is left as a possible future course of work. The kinetics associated with the defect-pinned domain may also be interesting to study. Information regarding the kinetically favored states may be relevant for understanding the dynamics of defect-pinned plectonemes.

The inherent inhomogeneity of DNA may get amplified due to the presence of certain periodic base pair arrays, which may be treated as a series of multiple defects. DNA defects generated by sequence specificity are known to affect nucleosome positioning thus regulating genome access. The framework developed here may be applied to treat a series of multiple defects. Role of defects in buckling of bulkier polymers like braids or single chromatin fibers may also be interesting to study in the future.

5.4 Chromosome-topology simplification driven by loop extrusion

Proper disentanglement and segregation of chromosomes is of paramount importance. The high density of chromosomes inside cell nucleus (or cell volume in case of bacteria) cause strong overlap between chromosomes leading to an entangled topology. TopoII allow topology fluctuations, however, local action of TopoII is not sufficient to disentangle the global topology of chromosomes. We find that lengthwise compaction via loop extrusion in presence of topol-

ogy fluctuations can drive chromosome disentanglement. Loop extrusion leads to a cylindrical brush morphology where chromosomal loops emanate from a flexible backbone, and via extrusion activity the structure can be driven to a highly compact state [Sec. 4.1.A]. The compact state corresponds to a high linear density of loops and leads to a chromosome core along the cylindrical backbone associated with dense packing of nucleosomes, the fundamental unit of chromosome polymer in eukaryotes, near the base of the loops. The cylindrical loop-extruded structure of chromosomes is associated with a transformation of the chromosome polymer viewed a string of nucleosomes or chromatin to a thicker rod-like polymer where the segment length is set by the persistence length of the polymer brush. This transformation drives a decrease in the inter-chromosomal entanglement in confinement [Sec. 4.1.B].

We find that entanglement in confinement is minimized by an optimal loop-extruded configuration where the optimal-loop size scales positively with the total chromosome length [Sec. **??**]. This optimization is a consequence of the finite size of the chromosomes, where the ends of the cylindrical structure that contributes to increase entanglement in the large loop limit. We discussed in detail the implications of loop extrusion in chromosome structure and disentanglement, such as the mechanical rigidity of chromosomes is derived from the densely packed core [Sec. **??**]. One important *in vivo* implication is that organization of interphase chromosomes into loops or chromosome domains, also called Topology Associating Domains (TADs), that have been strongly implicated to affect gene expression, also plays a role in minimizing inter-chromosomal entanglements [Sec. **??**].

It may be interesting to simulate confined polymer brushes and study their entanglement properties. We have focused on the steady-state or quasi-equilibrium properties of chromosomes to show that activity in form of loop extrusion can lead to disentanglement. The dynamical aspects of chromosome disentanglement are yet unexplored, and are very important from the *in vivo* perspective. One challenge in formulating the dynamics will be to accurately account for the polymer viscosity, which sets the fundamental unit of disentanglement time in chromosomes. This unit of disentanglement time is not only set by the concentration of chromosomes but also

the rate of topology fluctuations, such that a faster rate of fluctuating topology will have a less effective viscosity, thus making the viscosity an active parameter. Calculation of dynamics will be able to consolidate the interplay between compaction and segregation and can make testable predictions regarding the time scales expected for mitosis as a function of chromosome length and concentration.

The model of chromosomes presented here is a structural one, i.e., we focused on the structural aspects of chromosome compaction and entanglement, driven by loop extrusion. Manipulation of the structure is important to pack the genome into cylindrical chromosomes that are then divided between the daughter cells. Another equally important facet of chromosomes is its function, which is related to its sequence. DNA sequences code for proteins that are transcribed predominantly during the interphase stage of the cell cycle. Physical proximity of certain non-neighboring DNA sequences is essential for transcription, leading to formation of TAD structures in interphase [157, 158, 163]. The dynamics associated with TAD formation is crucial for gene regulation, and is intimately related to the chromosome structure. The framework developed here may be extended to furnish information regarding the time scales governing TAD formation, which will be relevant in understanding the spatio-temporal regulation of the genome during cell cycle. Organization of the interphase genome into compartments with enriched gene transcription is yet another feature that is poorly understood [163, 182].

Mechanics of cell nuclei has been shown to be dominated by the chromatin mechanics in the weak-stretching regime [211, 212]. The correspondence between the compaction state of the genome and its mechanics, as described by our model, may be explored in the context of nuclei mechanics. Chromosome and nuclei mechanics are affected by chemical modifications to the genome via post-translational modifications to histones [213]. However, the mechanistic understanding of how these chemical modifications affect rigidity is not clear. These chemical modifications, which are connected with gene regulation, may be modeled as higher-order interactions that promote nucleosome stacking. Our model may be extended to provide a quantification of the mechanical response as a function the stacking interactions.

In all, we have developed statistical-mechanical models for DNA and chromosome that enhance our understanding their mechanics and structure. The results of these models have provided quantitative explanations for various experimental observations and their biological implications. The next step with these models is to include the various aspects of biological function, such as methylation and acetylation of the genome that are important markers of biological function, however, they facilitate the function via influencing statistical mechanics.

Bibliography

[1] J. D. Watson and F. H. C. Crick. A structure for deoxyribose nucleic acid. *Nature*, 171:737–738, 1953.

[2] W. Flemming. Contributions to the knowledge of the cell and its vital processes (english translation of original paper of 1880). *Journal of Cell Biology*, 25:1–69, 1965.

[3] A.T. Sumner. *Chromosomes: Organization and Function*. Wiley, 2008.

[4] G. J. Mendel. Experiments in plant hybridization (translated in english. 1866.

[5] W. S. Sutton. The chromosomes in heredity. *Biological Bulletin*, 4:231–251, 1903.

[6] W. Johannsen. The genotype conception of heredity. *The American Naturalist*, 45:129–159, 1911.

[7] 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. *The Journal of Experimental Medicine*, 79:137–158, 1944.

[8] A. D. Hershey and M. Chase. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *The Journal of General Physiology*, 36:39–56, 1952.

[9] J.D. Watson. *Molecular Biology of the Gene*. Always learning. Pearson/CSH Press, 2014.

[10] S B Smith, L Finzi, and C Bustamante. Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science*, 1122:258, 1992.

[11] J. F. Marko and E. D. Siggia. Bending and twisting elasticity of DNA. *Macromolecules*, 27(4):981–988, 1994.

[12] J. F. Marko and Eric D Siggia. Stretching DNA. *Macromolecules*, 28(26):8759–8770, 1995.

[13] M. D. Stone, Z. Bryant, N. J. Crisona, S. B. Smith, A. Vologodskii, C. Bustamante, and N. R. Cozzarelli. Chirality sensing by Escherichia coli topoisomerase IV and the mechanism of type II topoisomerases. *Proc. Natl. Acad. Sci. USA*, 100(15):8654–8659, 2003.

[14] K. C. Neuman, G. Charvin, D. Bensimon, and V. Croquette. Mechanisms of chiral discrimination by topoisomerase IV. *Proc. Natl. Acad. Sci. USA*, 106(17):6986–6991, 2009.

[15] Y. Seol, A. H. Hardin, M. P. Strub, G. Charvin, and K. C. Neuman. Comparison of DNA decatenation by Escherichia coli topoisomerase IV and topoisomerase III: Implications for non-equilibrium topology simplification. *Nucleic Acids Res.*, 41(8):4640–4649, 2013.

[16] K. Terekhova, J. F. Marko, and A. Mondragón. Single-molecule analysis uncovers the difference between the kinetics of DNA decatenation by bacterial topoisomerases i and III. *Nucleic Acids Res.*, 42(18):11657–11667, 2014.

[17] J. C. Wang. Dna topoisomerases: Why so many? *The Journal of Biological Chemistry*, 266(11):6659–6662, 1991.

[18] J. F. Marko. Linking topology of tethered polymer rings with applications to chromosome segregation and estimation of the knotting length. *Phys. Rev. E*, 79:051905, May 2009.

[19] J. F. Marko. Scaling of Linking and Writhing Numbers for Spherically Confined and Topologically Equilibrated Flexible Polymers. *J. Stat. Phys.*, 142(6):1353–1370, 2011.

[20] E. Alipour and J. F. Marko. Self-organization of domain structures by DNA-loop-extruding enzymes. *Nucleic Acids Res.*, 40(22):11202–11212, 2012.

[21] K. Nasmyth. Disseminating the genome: Joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annual Review of Genetics*, 35(1):673–745, 2001. PMID: 11700297.

[22] A. Goloborodko, J. F. Marko, and L. A Mirny. Chromosome compaction via active loop extrusion. *Biophys. J.*, pages 2162–2168, 2016.

[23] D. S. Goodsell. The machinery or life. 1992.

[24] M. Feig, J.F. Marko, and M. Pettitt. 193, 2003.

[25] J. F. Marko. Course 7 Introduction to single-DNA micromechanics. *Les Houches Summer Sch. Proc.*, 82(C):211–270, 2005.

[26] P. G. DeGennes. *Scaling Concepts in Physics.* Cornell, Ithaca, NY, 1977.

[27] V. V. Rybenkov, N. R. Cozzarelli, and A. V. Vologodskii. Probability of dna knotting and the effective diameter of the dna double helix. *Proc. Natl. Acad. of Sci. USA*, 90:53075311, 1993.

[28] J. SantaLucia. A unified view of polymer, dumbbell, and oligonucleotide dna nearest-neighborthermodynamics. *Proc. Natl. Acad. of Sci. USA*, 95:14601465, 1998.

[29] B. Essevaz-Roulet, U. Bockelmann, and F. Heslot. Mechanical separation of the complementary strands ofdna. *Proc. Natl. Acad. of Sci. USA*, 94:1193511940, 1997.

[30] P. Bockelmann, U. Thomen, B. Essevaz-Roulet, V. Viasnoff, and F. Heslot. Unzipping dna with optical tweezers: high sequence sensitivity and force flips. *Biophys. J.*, 82:15371553, 2002.

[31] D. K. Lubensky and David R. Nelson. Pulling pinned polymers and unzipping dna. *Phys. Rev. Lett.*, 85:1572–1575, Aug 2000.

[32] P. Cluzel, A. Lebrun, C. Heller, R. Lavery, J.-L. Viovy, D. Chatenay, and F. Caron. Dna: An extensible molecule. *Science*, 271(5250):792–794, 1996.

[33] 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.

[34] T. Strick, F. Allemand, D. Bensimon, R. Lavery, and V. Croquette. Phase coexistence in a single DNA molecule. *Phys. A Stat. Mech. its Appl.*, 263:392–404, 1999.

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

[36] J. F. Marko and S. 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.

[37] J. F. Marko. Biophysics of protein-dna interactions and chromosome organization. *Physica A*, 418:126–153, 2015.

[38] J.F. Marko. Coupling of intramolecular and intermolecular linkage complexity of two DNAs. *Phys. Rev. E*, 59(1):900–912, 1999.

[39] J. Yan, M.O. Magnasco, and J.F. Marko. A kinetic proofreading mechanism for disentanglement of dna by topoisomerases. *Nature*, 401:932, 1999.

[40] 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*, 142:519, 2010.

[41] G.R. Buck and E.L. Zechiedrich. Dna disentangling by type-2 topoisomerases. *J Mol Biol.*, 340:933, 2004.

[42] T. Chen, Y. Sun, P. Ji, S. Kopetz, and W. Zhang. Topoisomerase ii in chromosome instability and personalized cancer therapy. *Oncogene*, 34:40194031, 2015.

[43] J. F. Marko. Micromechanical studies of mitotic chromosomes. *Chromosom. Res.*, 16(3):469–497, 2008.

[44] M. Doi and S. F. Edwards. *Theory of Polymer Dynamics*. Oxford Univ. Press, Oxford, 1986.

[45] J. F. Marko and E. D. Siggia. Polymer models of meiotic and mitotic chromosomes. *Mol. Biol. Cell*, 8:22217–2231, 1997.

[46] S. Brahmachari and J. F. Marko. Torque and buckling in stretched intertwined double-helix DNAs. *Phys. Rev. E*, 95(5):052401, 2017.

[47] S. Brahmachari, K. H. Gunn, R. D. Giuntoli, A. Mondragón, and J. F. Marko. Nucleation of multiple buckled structures in DNA double helices. *Phys. Rev. Lett.*, 119:188103, 2017.

[48] A. Dittmore, S. Brahmachari, Y. Takagi, J. F. Marko, and K. C. Neuman. Supercoiling DNA locates mismatches. *Phys. Rev. Lett.*, 119:147801, Oct 2017.

[49] S. Brahmachari, A. Dittmore, Y. Takagi, K. C. Neuman, and J. F. Marko. Defect-facilitated buckling in supercoiled double-helix dna. *Phys. Rev. E*, 97:022416, Feb 2018.

[50] H. Peng and K. J. Marians. Decatenation activity of Topoisomerase-IV during oriC and pBR322 DNA replication invitro. *Proc. Natl. Acad. Sci. U.S.A.*, 90(18):8571–8575, 1993.

[51] E. L. Zechiedrich and N. R. Cozzarelli. Roles of topoisomerase IV and DNA gyrase in DNA unlinking during replication in Escherichia coli. *Genes Dev.*, 9:2859–2869, 1995.

[52] P. Cejka, J. L. Plank, C. C. Dombrowski, and S. C. Kowalczykowski. Decatenation of DNA by the S. cerevisiae Sgs1-Top3-Rmi1 and RPA Complex: A Mechanism for Disentangling Chromosomes. *Mol. Cell*, 47(6):886–896, 2012.

[53] H. Bai, M. Sun, P. Ghosh, G. F. Hatfull, N. F. Grindley, and J. F. Marko. Single-molecule analysis reveals the molecular bearing mechanism of DNA strand exchange by a serine recombinase. *Proc. Natl. Acad. Sci. USA*, 108(18):7419–7424, 2011.

[54] T. R. Strick, J. F. Allemand, D. Bensimon, and V. Croquette. Behavior of supercoiled DNA. *Biophys. J.*, 74(4):2016–28, 1998.

[55] G. Charvin, D. Bensimon, and V. Croquette. Single-molecule study of DNA unlinking by eukaryotic and prokaryotic type-II topoisomerases. *Proc. Natl. Acad. Sci. USA*, 100(17):9820–9825, 2003.

[56] G. Charvin, A. Vologodskii, D. Bensimon, and V. Croquette. Braiding DNA: Experiments, Simulations, and Models. *Biophys. J.*, 88(6):4124–4136, 2005.

[57] Y. Seol, A. C Gentry, N. Osheroff, and K. C. Neuman. Chiral discrimination and writhe-dependent relaxation mechanism of human topoisomerase II. *J. Biol. Chem.*, 288(19):13695–13703, 2013.

[58] B. Xiao, M. M. McLean, X. Lei, J. F. Marko, and R. C. Johnson. Controlled rotation mechanism of DNA strand exchange by the Hin serine recombinase. *Sci. Rep.*, 6(April):23697, 2016.

[59] R. A. Keenholtz, N. F Grindley, G. F. Hatfull, and J. F. Marko. Crossover-site sequence and DNA torsional stress control strand interchanges by the Bxb1 site-specific serine recombinase. *Nucleic Acids Res.*, 44(18):gkw724, 2016.

[60] 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.

[61] J. D. Moroz and P. Nelson. Entropic elasticity of twist-storing polymers. *Macromolecules*, 31(18):6333–6347, 1998.

[62] J. F. Marko. Torque and dynamics of linking number relaxation in stretched supercoiled DNA. *Phys. Rev. E*, 76(2):1–13, 2007.

[63] 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.

[64] B. C. Daniels and J. P. Sethna. Nucleation at the DNA supercoiling transition. *Phys. Rev. E*, 83(4):1–16, 2011.

[65] 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, jan 2012.

[66] M. Emanuel, G. Lanzani, and H. Schiessel. Multiplectoneme phase of double-stranded DNA under torsion. *Phys. Rev. E*, 88(2):022706, 2013.

[67] R. Cortini, A. A. Kornyshev, D. J. Lee, and S. Leikin. Electrostatic braiding and homologous pairing of DNA double helices. *Biophys. J.*, 101(4):875–884, 2011.

[68] D. Argudo and P. K. Purohit. DNA superhelical structures with non-constant helical pitch. *Math. Mech. Solids*, (May):1–25, 2013.

[69] D. J. Lee. Effect of undulations on spontaneous braid formation. *Phys. Rev. E*, 88(2), 2013.

[70] D. J. Lee. Collapse and coexistence for a molecular braid with an attractive interaction component subject to mechanical forces. *J. Phys. Condens. Matter*, 145101:145101, 2014.

[71] G. H. M. van der Heijden. The static deformation of a twisted elastic rod constrained to lie on a cylinder. *Proc. R. Soc. London A Math. Phys. Eng. Sci.*, 457(2007):695–715, 2001.

[72] J. Thompson, G.H.M. van der Heijden, and S. Neukirch. Supercoiling of DNA plasmids: mechanics of the generalized ply. *Proc. R. Acad. London A*, 458:959–985, 2002.

[73] G. H.M. Van Der Heijden, J. M.T. Thompson, and S. Neukirch. A Variational Approach to Loaded Ply Structures. *J. Vib. Control*, 9(1-2):175–185, 2003.

[74] J. A. Schellman and D. Stigter. Electrical double layer, zeta potential, and electrophoretic charge of double-stranded DNA. *Biopolymers*, 16(7):1415–1434, 1977.

[75] R. Podgornik and V. A. Parsegian. Molecular fluctuations in the packing of polymeric liquid crystals. *Macromolecules*, 23(24):2265–2269, 1990.

[76] J. Ubbink and T. Odijk. Electrostatic-undulatory theory of plectonemically supercoiled DNA. *Biophys. J.*, 76(5):2502–2519, 1999.

[77] D. Stigter. The charged colloidal cylinder with a gouy double layer. *J. Colloid Interface Sci.*, 53(2):296–306, 1975.

[78] J. R Wenner, M. C Williams, I. Rouzina, and V. A Bloomfield. Salt dependence of the elasticity and overstretching transition of single DNA molecules. *Biophys. J.*, 82(6):3160–3169, 2002.

[79] 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.

[80] T. Odijk. On the Statistics and Dynamics of Confined or Entangled Stiff Polymers. *Macromolecules*, 1344(3):1340–1344, 1983.

[81] T. W. Burkhardt. Free energy of a semiflexible polymer confined along an axis. *J. Phys. A. Math. Gen.*, 28(24):L629, 1995.

[82] J. H. White and W. R. Bauer. Calculation of the twist and the writhe for representative models of DNA. *J. Mol. Biol.*, 189(2):329–341, 1986.

[83] 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.

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

[85] H. Yamakawa. Statistical Mechanics of Wormlike Chains. *Pure Appl. Chem.*, 46(2):135–141, 1976.

[86] I. M. Kulić and H. Schiessel. Nucleosome repositioning via loop formation. *Biophys. J.*, 84(5):3197–3211, 2003.

[87] S. Sankararaman and J. F. Marko. Formation of loops in DNA under tension. *Phys. Rev. E*, 71(2):021911, 2005.

[88] K. H. Gunn, J. F. Marko, and A. Mondragon. *Nat. Struct. Mol. Biol.*, 24:484–490, 2017.

[89] M. T. J. van Loenhout, M. V. de Grunt, and C. Dekker. Dynamics of DNA Supercoils. *Science*, 338(October):94–7, 2012.

[90] C. Matek, T. E. Ouldridge, J. P. K. Doye, and A. A. Louis. Plectoneme tip bubbles: Coupled denaturation and writhing in supercoiled dna. *Sci. Rep.*, 5:7655, 2015.

[91] M. Ganji, S. H. Kim, Jaco Van Der Torre, E. Abbondanzieri, and C. Dekker. Intercalation-based single-molecule fluorescence assay to study DNA supercoil dynamics. *Nano Lett.*, 16(7):4699–4707, 2016.

[92] L. Postow, C. D. Hardy, J. Arsuaga, and N. R. Cozzarelli. . *Genes. Dev.*, 18:1766–79, 2004.

[93] S. J. Spengler, A. Stasiak, and N. R. Cozzarelli. . *Cell*, 42:325, 1985.

[94] M. Dunaway and E. A. Ostrander. Local domains of supercoiling activate a eukaryotic promoter in vivo. *Nature*, 361:746–748, 2 1993.

[95] J. J. Champoux. Dna topoisomerases: Structure, function, and mechanism. *Annual Review of Biochemistry*, 70(1):369–413, 2001.

[96] J. van Noort, S. Verbrugge, N. Goosen, C. Dekker, and R. T. Dame. Dual architectural roles of hu: Formation of flexible hinges and rigid filaments. *Proc. Natl. Acad. Sci. U.S.A.*, 101(18):6969–6974, 2004.

[97] D. Skoko, B. Wong, R. C. Johnson, and J. F. Marko. Micromechanical analysis of the binding of dna-bending proteins hmgb1, nhp6a, and hu reveals their ability to form highly stable dnaprotein complexes. *Biochemistry*, 43(43):13867–13874, 2004. PMID: 15504049.

[98] P. T. Lowary and J. Widom. *J. Mol. Biol.*, 276:19, 1998.

[99] G.-C. Yuan, Y.-J. Liu, M. F. Dion, M. D. Slack, L. F. Wu, S. J. Altschuler, and O. J. Rando. Genome-scale identification of nucleosome positions in s. cerevisiae. *Science*, 309(5734):626–630, 2005.

[100] E. Segal, Y. Fondufe-Mittendorf, L. Chen, A. Thstrm, Y. Field, I. K. Moore, J.-P. Z. Wang, and J. Widom. A genomic code for nucleosome positioning. *Nature*, 442:772–778, 2006.

[101] A. L. Hughes and O. J. Rando. Mechanisms underlying nucleosome positioning in vivo. *Annu. Rev. Biophys.*, 43:41–63, 2014.

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

[103] N. Clauvelin, B. Audoly, and S. Neukirch. Mechanical response of plectonemic dna: An analytical solution. *Macromolecules*, 41(12):4479–4483, 2008.

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

[105] T. C. Boles, J. H. White, and N. R. Cozzarelli. Structure of plectonemically supercoiled DNA. *J. Mol. Biol.*, 213(4):931–51, 1990.

[106] 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.

[107] 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 Letters*, 12(7):3634–3639, 2012.

[108] H. Yamakawa and W. H. Stockmayer. Statistical mechanics of wormlike chains. II. excluded volume effects. *J. Chem. Phys.*, 57:2843, 1972.

[109] H. Zhang and J. F. Marko. Maxwell relations for single-dna experiments: Monitoring protein binding and double-helix torque with force-extension measurements. *Phys. Rev. E*, 77:031916, Mar 2008.

[110] F. Kriegel, N. Ermann, R. Forbes, D. Dulin, N. H. Dekker, and J. Lipfert. Probing the salt dependence of the torsional stiffness of dna by multiplexed magnetic torque tweezers. *Nucl. Acids. Res.*, 45:5920–5929, 2017.

[111] J. J. SantaLucia, H. T. Allawi, and P. A. Seneviratne. *Biochemistry*, 35:3555–3562, 1996.

[112] B. Essevaz-Roulet, U. Bockelmann, and F. Heslot. Mechanical separation of the complementary strands ofdna. *Proc. Natl. Acad. Sci. USA*, 94(22):11935–11940, 1997.

[113] A. Agresti. *An Introduction to Categorical Data Analysis*, page 35. John Wiley & Sons, Inc., 2006.

[114] H. Meng, J. Bosman, T. vanderHeijden, and J. vanNoort. Coexistence of twisted, plectonemic, and melted dna in small topological domains. *Biophys. J.*, 106(5):1174 – 1181, 2014.

[115] D. M. Gowers and S. E. Halford. Protein motion from non-specific to specific dna by three-dimensional routes aided by supercoiling. *EMBO J.*, 22(6):1410–1418, 2003.

[116] M. A. Lomholt, B. van den Broek, S.-M. J. Kalisch, G. J. L. Wuite, and R. Metzler. Facilitated diffusion with dna coiling. *Proc. Natl. Acad. Sci. USA*, 106(20):8204–8208, 2009.

[117] J. Gorman, F. Wang, S. Redding, A. J. Plys, T. Fazio, S. Wind, Eric E. Alani, and Eric C. Greene. Single-molecule imaging reveals target-search mechanisms during dna mismatch repair. *Proc. Natl. Acad. Sci. U.S.A.*, 109(45):E3074–E3083, 2012.

[118] D. J. Hosfield, Y. Guan, B. J. Haas, R. P. Cunningham, and J. A. Tainer. Structure of the dna repair enzyme endonuclease iv and its dna complex: Double-nucleotide flipping at abasic sites and three-metal-ion catalysis. *Cell*, 98(3):397 – 408, 1999.

[119] S. H. Kim, M. Ganji, J. van der Torre, E. Abbondanzieri, and C. Dekker. Dna sequence encodes the position of dna supercoils. *bioRxiv*, 2017.

[120] M. Sun, R. Kawamura, and J. F Marko. Micromechanics of human mitotic chromosomes. *Phys. Biol*, 8(1):1–18, 2011.

[121] S. Almagro, D. Riveline, T. Hirano, B. Houchmandzadeh, and S. Dimitrov. The mitotic chromosome is an assembly of rigid elastic axes organized by structural maintenance of chromosomes (smc) proteins and surrounded by a soft chromatin envelope. *Journal of Biological Chemistry*, 279(7):5118–5126, 2004.

[122] B. Houchmandzadeh, J. F. Marko, D. Chatenay, and A. Libchaber. Elasticity and structure of eukaryote chromosomes studied by micromanipulation and micropipette aspiration. *The Journal of Cell Biology*, 139(1):112, 1997.

[123] M. G. Poirier, S. Eroglu, and J. F. Marko. The bending rigidity of mitotic chromosomes. *Molecular Biology of the Cell*, 13(6):21702179, 2002.

[124] R. B. Nicklas. Measurements of the force produced by the mitotic spindle in anaphase. *The Journal of Cell Biology*, 97(2):542548, 1983.

[125] A. Goloborodko, M. V. Imakaev, J. F. Marko, and L. A Mirny. Compaction and segregation of sister chromatids via active loop extrusion. *eLife*, e14864(5), 2016.

[126] M. Daoud and J. P. Cotton. Star shaped polymers: a model for the conformation and its concentration dependence. *Journal de Physique*, 43:531–38, 1982.

[127] B. Houchmandzadeh and S. Dimitrov. Elasticity measurements show the existence of thin rigid cores inside mitotic chromosomes. *The Journal of Cell Biology*, 145(2):215–223, 1999.

[128] G. Fudenberg, M. Imakaev, C. Lu, A. Goloborodko, N. Abdennur, and L. A. Mirny. Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.*, 15(9), 2016.

[129] J. H Gibcus, K. Samejima, A. Goloborodko, I. Samejima, N. Naumova, J. Nuebler, M. T. Kanemaki, L. Xie, J. R. Paulson, W. C. Earnshaw, L. A. Mirny, and J. Dekker. A pathway for mitotic chromosome formation. *Science*, 6135(January), 2018.

[130] R. C. Ball, Milner S.T. Marko, J. F., and Thomas A. Witten. Polymers grafted to a convex surface. *Macromolecules*, 24:693–703, 1991.

[131] H. Li and T. A. Witten. Polymers grafted to convex surfaces: A variational approach. *Macromolecules*, 27:449–457, 1994.

[132] L. Feuz, F. A. M Leermakers, M. Textor, and O. Borisov. Bending rigidity and induced persistence length of molecular bottle brushes: A self-consistent-field theory. *Macromolecules*, 38(21):8891–8901, 2005.

[133] S. Panyukov, E. B. Zhulina, S. S. Sheiko, G. C. Randall, J. Brock, and M. Rubinstein. Tension Amplification in Molecular Brushes in Solutions and on Substrates. *J Phys. Chem. B*, 113:3750–3768, 2009.

[134] M. Daoud and P.G. De Gennes. Statistics of macromolecular solutions trapped in small pores. *Journal de Physique* , 38:85–93, 1977.

[135] P. Pincus. Excluded Volume Effects and Stretched Polymer Chains. *Macromolecules*, 9(3):386–388, 1976.

[136] A. Rosa and R. Everaers. Structure and Dynamics of Interphase Chromosomes. *PLoS Comp. Biol.*, 4(8):e1000153, 2008.

[137] J. D. Halverson, J. Smrek, K. Kremer, and A. Y. Grosberg. From a melt of rings to chromosome territories : the role of topological constraints in genome folding. *Rep. Prog. Phys*, 77(022601), 2014.

[138] L. C. Green, P. Kalitsis, T. M. Chang, M. Cipetic, J. H. Kim, O. Marshall, L. Turnbull, C. B. Whitchurch, P. Vagnarelli, K. Samejima, W. C. Earnshaw, K. H. A. Choo, and D. F. Hudson. Contrasting roles of condensin i and condensin ii in mitotic chromosome formation. *Journal of Cell Science*, 125(6):1591–1604, 2012.

[139] T. Ono, D. Yamashita, and T.Hirano. Condensin ii initiates sister chromatid resolution during s phase. *The Journal of Cell Biology*, 200(4):429–441, 2013.

[140] T. Terakawa, S. Bisht, J. M. Eeftens, C. Dekker, C. H. Haering, and E. C. Greene. The condensin complex is a mechanochemical motor that translocates along dna. *Science*, 358(6363):672676, 2017.

[141] M. Ganji, I. A Shaltiel, S. Bisht, E. Kim, A. Kalichava, C. H. Haering, and C. Dekker. Real-time imaging of DNA loop extrusion by condensin. *Science*, 7831(February):1–9, 2018.

[142] R. A. Keenholtz, T. Dhanaraman, R. Palou, J. Yu, D. DAmours, and J. F. Marko. Oligomerization and atp stimulate condensin-mediated dna compaction. *Sci. Rep.*, 7(1), 2017.

[143] N. Walther, M. J. Hossain, A. Z. Politi, B. Koch, M. Kueblbeck, Ø. Ødegård-Fougner, M. Lampe, and J. Ellenberg. A quantitative map of human condensins provides new insights into mitotic chromosome architecture. *The Journal of Cell Biology*, 2018.

[144] T. Ono, A. Losada, M. Hirano, M. P. Myers, A. F. Neuwald, and T. Hirano. Differential contributions of condensin i and condensin ii to mitotic chromosome architecture in vertebrate cells. *Cell*, 115(1):109 – 121, 2003.

[145] R. Kawamura, L. H. Pope, M. O. Christensen, M. Sun, K. Terekhova, F. Boege, C. Mielke, A. H. Andersen, and J. F. Marko. Mitotic chromosomes are constrained by topoisomerase II-sensitive DNA entanglements. *J. Cell Biol.*, 188(5):653–663, 2010.

[146] A. Zhiteneva, J. J. Bonfiglio, A. Makarov, T. Colby, P. Vagnarelli, E. C. Schirmer, and W. C. Earnshaw. Mitotic post-translational modifications of histones promote chromatin compaction in vitro. *Open Biology*, 7(9), 2017.

[147] J. J. Siegel and A. Amon. New insights into the troubles of aneuploidy. *Annual Review of Cell and Developmental Biology*, 28:189214, 2012.

[148] D. Zhang and S. Oliferenko. Remodeling the nuclear membrane during closed mitosis. *Current Opinion in Cell Biology*, 25(1):142 – 148, 2013. Cell architecture.

[149] S. Sazer, M. Lynch, and D. Needleman. Deciphering the evolutionary history of open and closed mitosis. *Current Biology*, 24(22):R1099R1103, 2014.

[150] C. P. De Souza and S. A. Osmani. Double duty for nuclear proteins the price of more open forms of mitosis. *Trends in Genetics*, 25(12):545554, 2009.

[151] K. K. Lee, Y. Gruenbaum, P. Spann, J. Liu, and K. L. Wilson. C. elegans nuclear envelope proteins emerin, man1, lamin, and nucleoporins reveal unique timing of nuclear envelope breakdown during mitosis. *Molecular Biology of the Cell*, 11(9):30893099, 2000.

[152] J Richard McIntosh and Eileen T O'Toole. Life cycles of yeast spindle pole bodies: Getting microtubules into a closed nucleus. *Biology of the Cell*, 91(45):305–312, 5 1999.

[153] S. L. Jaspersen and S. Ghosh. Nuclear envelope insertion of spindle pole bodies and nuclear pore complexes. *Nucleus*, 3(3):226236, 2012.

[154] M.R. Paddy, H. Saumweber, D.A. Agard, and J.W. Sedat. Time-resolved, in vivo studies of mitotic spindle formation and nuclear lamina breakdown in drosophila early embryos. *Journal of Cell Science*, 109(3):591–607, 1996.

[155] E. Piskadlo, A. Tavares, and R. A. Oliveira. Metaphase chromosome structure is dynamically maintained by condensin i-directed dna (de)catenation. *eLife*, 6:e26120, 2017.

[156] J. R. Dixon, S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, and B. Ren. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398):376380, 2012.

[157] E. P. Nora, B. R. Lajoie, E. G. Schulz, L. Giorgetti, I. Okamoto, N. Servant, and E. Heard. Spatial partitioning of the regulatory landscape of the x-inactivation center. *Nature*, 485(7398):381385, 2012.

[158] S. S. P. Rao, M. H. Huntley, N. C. Durand, E. K. Stamenova, I. D. Bochkov, J. T. Robinson, A. L. Sanborn, I. Machol, A. D. Omer, E. S. Lander, and E. Lieberman-Aiden. Article A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell*, 159(7):1665–1680, 2014.

[159] T. B. K. Le, M. V. Imakaev, L. A. Mirny, and M. T. Laub. High-resolution mapping of the spatial organization of a bacterial chromosome. *Science (New York, N.Y.)*, 342(6159):731734, 2013.

[160] C. Liu, Y.-J. Cheng, J.-W. Wang, and D. Weigel. Prominent topologically associated domains differentiate global chromatin packing in rice from arabidopsis. *Nature Plants*, 3:742–748, 2017.

[161] C. Hou, L. Li, Z. S. Qin, and V. G. Corces. Gene density, transcription and insulators contribute to the partition of the drosophila genome into physical domains. *Molecular Cell*, 48(3):471484, 2012.

[162] Q. Dong, N. Li, X. Li, Z. Yuan, D. Xie, X. Wang, J. Li, Y. Yu, J. Wang, B. Ding, Z. Zhang, C. Li, Y. Bian, A. Zhang, Y. Wu, B. Liu, and L. Gong. Genome wide hic analysis reveals extensive hierarchical chromatin interactions in rice. *The Plant Journal*, 94(6):1141–1156, 6 2018.

[163] N. Naumova, M. Imakaev, G. Fudenberg, Y. Zhan, B. R. Lajoie, L. A. Mirny, and J. Dekker. Organization of the mitotic chromosome. *Science*, 342(6161):948–953, 2014.

[164] L. H. Pope, C. Xiong, and J. F. Marko. Proteolysis of mitotic chromosomes induces gradual and anisotropic decondensation correlated with a reduction of elastic modulus and structural sensitivity to rarely cutting restriction enzymes. *Molecular Biology of the Cell*, 17(1):104–113, 2006. PMID: 16221892.

[165] B. John and K. R. Lewis. Chromosome heirarchy (p 23). 1975.

[166] B. John and K. R. Lewis. The meiotic mechanism. 1983.

[167] A. H. Sparrow, H. J. Price, and A. G. Underbrink. *Brookhaven Symp. Biol.*, 23:451, 1972.

[168] A. T. Sumner. Chromosome banding. 1990.

[169] M. D. Bennett. *Phil. Trans. Roy. Soc. Lond.*, B277:201, 1977.

[170] J. Darnell, H. Lodish, and D. Baltimore. *Molecular Cell Biology*, 1990.

[171] T. D. Allen, E. M. Jack, and C. J. Harrison. Chromosomes and chromatin, vol ii (p 54). 1988.

[172] G. Holmquist. Chromosomes and chromatin, vol ii (p 88). 1988.

[173] H. Macgregor. An introduction to animal cytogenetics (p 7). 1993.

[174] J. R. Daban. The energy components of stacked chromatin layers explain the morphology, dimensions and mechanical properties of metaphase chromosomes. *Journal of the Royal Society Interface*, 11(92), 2014.

[175] D. G. Booth, M. Takagi, L. Sanchez-Pulido, E. Petfalski, G. Vargiu, K. Samejima, and P. Vagnarelli. Ki-67 is a pp1-interacting protein that organises the mitotic chromosome periphery. *eLife*, 3:e01641, 2014.

[176] S. Cuylen, C. Blaukopf, A. Z. Politi, T. Mller-Reichert, B. Neumann, I. Poser, and D. W. Gerlich. Ki-67 acts as a biological surfactant to disperse mitotic chromosomes. *Nature*, 535(7611):308312, 2016.

[177] M. Takagi, T. Ono, T. Natsume, C. Sakamoto, M. Nakao, N. Saitoh, M. T. Kanemaki, T. Hirano, and N. Imamoto. Ki-67 and condensins support the integrity of mitotic chromosomes through distinct mechanisms. *Journal of Cell Science*, 131(6), 2018.

[178] K. K Swinger and P. A Rice. Ihf and hu: flexible architects of bent dna. *Current Opinion in Structural Biology*, 14(1):28 – 35, 2004.

[179] V. V. Rybenkov, V. Herrera, Z. M. Petrushenko, and H. Zhao. Mukbef, a chromosomal organizer. *Journal of Molecular Microbiology and Biotechnology*, 24:371383, 2014.

[180] S. Jun and B. Mulder. Entropy-driven spatial organization of highly confined polymers: Lessons for the bacterial chromosome. *Proceedings of the National Academy of Sciences*, 103(33):12388–12393, 2006.

[181] U. Eser, D. Chandler-Brown, F. Ay, A. F. Straight, Z. Duan, W. S. Noble, and J. M. Skotheim. Form and function of topologically associating genomic domains in budding yeast. *Proc. Natl. Acad. Sci.*, 114(15):E3061–E3070, 2017.

[182] E. Lieberman-aiden, N. L. Van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, and L. A. Mirny. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science*, 33292(October):289–294, 2009.

[183] A. R. Strom, A. V. Emelyanov, M. Mir, D. V. Fyodorov, X. Darzacq, and G. H. Karpen. Phase separation drives heterochromatin domain formation. *Nature*, 547(7662):241–245, 2017.

[184] J. R. Paulson and U.K. Laemmli. The structure of histone-depleted metaphase chromosomes. *Cell*, 12(3):817 – 828, 1977.

[185] J.B. Rattner and C.C. Lin. Radial loops and helical coils coexist in metaphase chromosomes. *Cell*, 42(1):291 – 296, 1985.

[186] N. Kireeva, M. Lakonishok, I. Kireev, T. Hirano, and A. S. Belmont. Visualization of early chromosome condensation: a hierarchical folding, axial glue model of chromosome structure. *The Journal of Cell Biology*, 166(6):775785, 2004.

[187] M. Milla and J. R. Daban. Self-assembly of thin plates from micrococcal nuclease-digested chromatin of metaphase chromosomes. *Biophysical Journal*, 103(3):567575, 2012.

[188] K. P. Baetcke, A. H. Sparrow, C. H. Nauman, and S. S. Schwemmer. The relationship of dna content to nuclear and chromosome volumes and to radiosensitivity (ld50). *Proceedings of the National Academy of Sciences of the United States of America*, 58(2):533540, 1967.

[189] G. Maul and L. Deaven. Quantitative determination of nuclear pore complexes in cycling cells with differing dna content. *The Journal of Cell Biology*, 73(3):748–760, 1977.

[190] J. Schwarz-Finsterle, H. Scherthan, A. Huna, P. Gonzlez, P. Mueller, E. Schmitt, J. Erenpreisa, and M. Hausmann. Volume increase and spatial shifts of chromosome territories in nuclei of radiation-induced polyploidizing tumour cells. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 756(1):56 – 65, 2013. From DNA Damage to Chromosomal Aberrations.

[191] R. R. Snapp, E. Goveia, L. Peet, N. A. Bouffard, G. J. Badger, and H. M. Langevin. Spatial organization of fibroblast nuclear chromocenters: component tree analysis. *Journal of Anatomy*, 2013.

[192] D. H. Kim, B. Li, F. Si, J. M. Phillip, D. Wirtz, and S. X. Sun. Volume regulation and shape bifurcation in the cell nucleus. *Journal of Cell Science*, 128(18):3375–3385, 2015.

[193] T. A. Dittmer, N. J. Stacey, K. Sugimoto-Shirasu, and E. J. Richards. Little nuclei genes affecting nuclear morphology in arabidopsis thaliana. *The Plant Cell*, 19(9):2793–2803, 2007.

[194] L. Willis, Y. Refahi, R. Wightman, B. Landrein, J. Teles, K. C. Huang, E. M. Meyerowitz, and H. Jönsson. Cell size and growth regulation in the arabidopsis thaliana apical stem cell niche. *Proceedings of the National Academy of Sciences*, 113(51):E8238–E8246, 2016.

[195] Y. Arata, H. Takagi, Y. Sako, and H. Sawa. Power law relationship between cell cycle duration and cell volume in the early embryonic development of caenorhabditis elegans. 2014.

[196] S. Uppaluri, S. C. Weber, and C.P. Brangwynne. Hierarchical size scaling during multicellular growth and development. *Cell Reports*, 17(2):345 – 352, 2016.

[197] Y. Gonzalez, K. Meerbrey, J. Chong, Y. Torii, N. N. Padte, and S. Sazer. Nuclear shape, growth and integrity in the closed mitosis of fission yeast depend on the ran-gtpase system, the spindle pole body and the endoplasmic reticulum. *Journal of Cell Science*, 122(14):2464–2472, 2009.

[198] A. Takemoto, S. A. Kawashima, J.-J. Li, L. Jeffery, K. Yamatsugu, O. Elemento, and P. Nurse. Nuclear envelope expansion is crucial for proper chromosomal segregation during a closed mitosis. *Journal of Cell Science*, 129(6):1250–1259, 2016.

[199] F. R. Neumann and P. Nurse. Nuclear size control in fission yeast. *The Journal of Cell Biology*, 179:593600, 2007.

[200] R. P. Joyner, J. H. Tang, J. Helenius, E. Dultz, C. Brune, L. J. Holt, and K. ... Weis. A glucose-starvation response regulates the diffusion of macromolecules. 2016.

[201] P. Jorgensen, N. P. Edgington, B. L. Schneider, I. Rupe, M. Tyers, and B. Futcher. The size of the nucleus increases as yeast cells grow. *Molecular Biology of the Cell*, 18(9):35233532, 2007.

[202] H. E. Kubitschek and J. A. Friske. Determination of bacterial cell volume with the coulter counter. *Journal of Bacteriology*, 168:14661467., 1986.

[203] National center for biotechnology information (ncbi)[internet]. bethesda (md): National library of medicine (us), national center for biotechnology information; [1988] [cited 2018 apr 06]. available from: https://www.ncbi.nlm.nih.gov/.

[204] E. Crane, Q. Bian, R. P. McCord, B. R. Lajoie, B. S. Wheeler, E. J. Ralston, and B. J. Meyer. Condensin-driven remodeling of x-chromosome topology during dosage compensation. *Nature*, 523, 2015.

[205] H. Tanizawa, K. D. Kim, O. Iwasaki, and K. Noma. Architectural alterations of the fission yeast genome during the cell cycle. *Nature Structural Molecular Biology*, 24(11):965976, 2017.

[206] T. Mizuguchi, G. Fudenberg, S. Mehta, J. M. Belton, N. Taneja, H. D. Folco, and S. I. S. Grewal. Cohesin-dependent globules and heterochromatin shape 3d genome architecture in s. pombe. *Nature*, 516(7531):432435, 2014.

[207] Y. Kakui, A. Rabinowitz, D. J. Barry, and F. Uhlmann. Condensin-mediated remodeling of the mitotic chromatin landscape in fission yeast. *Nat. Genet.*, 49(10):1553–1557, 2017.

[208] K. D. Kim, H. Tanizawa, O. Iwasaki, and K. Noma. Transcription factors mediate condensin recruitment and global chromosomal organization in fission yeast. *Nature Genetics*, 48(10):12421252, 2016.

[209] C. Nikolaou. *Curr Genet*, 64:247, 2018.

[210] R. Koszul, K.P. Kim, M. Prentiss, N. Kleckner, and S. Kameoka. Meiotic chromosomes move by linkage to dynamic actin cables with transduction of force through the nuclear envelope. *Cell*, 133(7):1188 – 1201, 2008.

[211] A. D. Stephens, E. J. Banigan, S. A. Adam, R. D. Goldman, and J. F. Marko. Chromatin and lamin a determine two different mechanical response regimes of the cell nucleus. *Molecular Biology of the Cell*, 28(14):1984–1996, 2017. PMID: 28057760.

[212] A. D. Stephens, E. J. Banigan, and J. F. Marko. Separate roles for chromatin and lamins in nuclear mechanics. *Nucleus*, 9(1):119–124, 2018. PMID: 29227210.

[213] A. D. Stephens, P. Z. Liu, E. J. Banigan, L. M. Almassalha, V. Backman, S. A. Adam, R. D. Goldman, J. F. Marko, and T. Misteli. Chromatin histone modifications and rigidity affect nuclear morphology independent of lamins. *Molecular Biology of the Cell*, 29(2):220–233, 2018. PMID: 29142071.

[214] S. L. Brenner and V. A. Parsegian. A physical method for deriving the electrostatic interaction between rod-like polyions at all mutual angles. *Biophys. J.*, 14:327–334, 1974.

Appendix A

DNA electrostatics: Debye-Hückel theory

Each DNA base pair (~ 0.34 nm) contains two units (1 unit $\equiv e = 1.6 \times 10^{-19}$ C) of negative charge. Hence, interaction of electrostatic origin is important in estimation of free energy of DNA chain conformations. The physiological environment of DNA consists of various ionic species, predominantly ≈ 0.1 M univalent ions (Na^+ or K^+), which results in a non-uniform charge distribution in the vicinity of the negatively charged DNA backbone. The number density of an ionic species can be calculated using the Maxwell-Boltzmann distribution,

$$n_i(\mathbf{r}) = n_i^0 e^{-\beta z_i e \Psi(\mathbf{r})} \tag{A.1}$$

where $\Psi(\mathbf{r})$ is the electrostatic potential we wish to calculate. Using the above number density, we can write the charge distribution in the electrolytic solution surrounding the DNA and arrive at the Poisson-Boltzmann equation.

$$\nabla^2 \Psi = \frac{4\pi e}{\epsilon} \sum_i n_i^0 z_i e^{-\beta z_i e \Psi} \tag{A.2}$$

where z_i is the valency of the i-th ionic species, and ϵ is the dielectric constant of the medium. Eq. (A.2) can be linearized by expanding the exponential, which is valid as long as the electrostatic energy is small compared to the thermal energy ($\beta e\Psi \ll 1$).

$$\lambda_D^2 \nabla^2 \Psi = \Psi \tag{A.3}$$

where $\lambda_D^{-2} = \frac{4\pi\beta e^2}{\epsilon} \sum_i n_i^0 z_i^2$, is called the Debye length of the electrolyte and Eq. (A.3) is also known as the Debye-Huckel equation. λ_D characterizes an emergent length scale over which the Coulomb interaction is shut-off due to rearrangement of charges in the solution.

A.1 Parallel dsDNAs

Eq. (A.3) can be solved for a cylindrical rod where the potential is radially symmetric and varies as a function of the distance from the rod axis. The potential at points far from the axis of the rod can be written as [214],

$$\Psi(r) = (2\nu_h/\epsilon) K_0(r/\lambda_D) \tag{A.4}$$

where ν_h is the *effective* linear charge density of the hypothetical line charge along the axis of the cylindrical rod that causes this potential. Using Eq. (A.4), we can calculate the leading order electrostatic interaction energy in plectonemic DNA by considering pairs of nearest points on the two DNAs as one moves along the plectoneme axis [Fig. 3.1(a)]. An estimation of ν_h readily follows from matching the boundary conditions for the electric field near the surface of the cylinder.

$$-\frac{\partial \Psi}{\partial r}\Big|_{r=a} = \frac{2\nu_h}{\epsilon \lambda_D} K_1(\kappa a) = \frac{4\pi\sigma}{\epsilon} \tag{A.5}$$

where $\sigma = e/2\pi ab$ is the actual surface charge density on the cylinder with radius a and a charge of $1e$ per length b ($b = 0.17$ nm for double helix DNA).

$$\nu_h = \frac{\lambda_D e\gamma(\epsilon, a/\lambda_D)}{abK_1(a/\lambda_D)} \tag{A.6}$$

where γ is a numerical coefficient used to match the far-field Debye-Hückel solution to the exact numerical solution of Eq. (A.2) (Gouy-Chapman) near the surface of the cylinder. γ has been tabulated for various values of salt concentrations[63, 77].

Finally, we write the free energy per unit length due to leading order electrostatic interaction of plectonemic strands $2r$ apart.

$$\mathcal{U}_{el} = \frac{1}{2}\nu_h \Psi(2r/\lambda_D) \tag{A.7}$$

where the factor of half corrects for double counting. Now defining $\nu = \nu_h/e$ and $\ell_B = \beta e^2/\epsilon$, the Bjerrum length we write the following.

$$\beta\mathcal{U}_{el} = \ell_B \nu^2 K_0(2r/\lambda_D) \tag{A.8}$$

The values of ν for a few representative salt concentrations are given in Table 2.1.

A.2 Circular loop of dsDNA

The total Debye-Hückel self electrostatic energy of a circle of perimeter L:

$$U = \frac{\zeta}{2}\int_0^L ds_1 \int_0^L ds_2 \left[\frac{e^{-|\mathbf{r_1}(s_1)-\mathbf{r_2}(s_2)|/\lambda_D}}{|\mathbf{r_1}(s_1)-\mathbf{r_2}(s_2)|} - \frac{e^{-|s_1-s_2|/\lambda_D}}{|s_1-s_2|}\right] \tag{A.9}$$

where $\zeta = k_{\mathrm{B}} T \ell_B \nu^2$, where $\ell = 0.7$nm, is the Bjerrum length of water at $290K$ and ν is the effective linear charge density of DNA. We parameterize the circle using $\xi = L/(\pi\lambda_D)$, such that

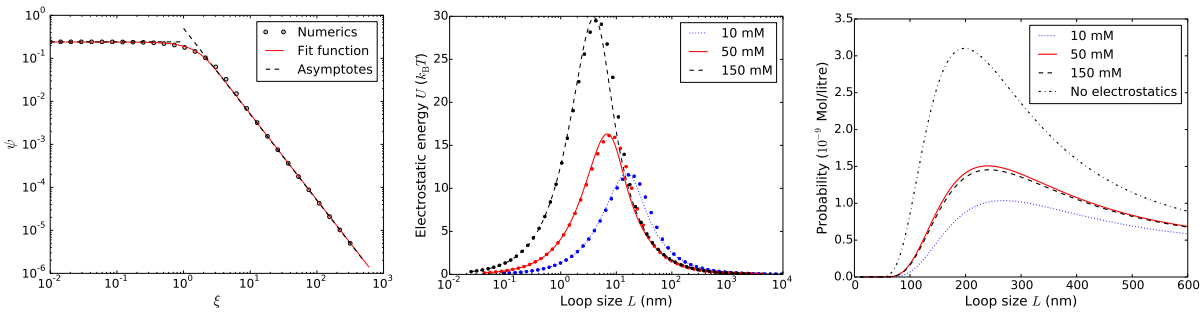

Figure A.1: Electrostatic interactions in a circular DNA loop. (a)Comparison of exact numerical evaluation of ψ [Eq. (A.10)], shown as points) with the proposed fit function [Eq. (A.14), shown as the solid black line] as a function of normalized loop size $\xi = L/(\pi\lambda_D)$. The asymptotic solutions for large loop [$\xi \ll 1$, Eq. (A.11)] and small loop [$\xi \gg 1$, Eq. (A.13)] are shown as red dashed lines. (b)Electrostatic self-energy of a circular DNA loop as a function of the loop size L for 10, 100 and 200 mM monovalent salt concentrations. The points show the exact numerical values whereas the lines show the energy from the fit function [Eq. (A.14)]. (c) Looping probability of double-helix DNA as a function of the size of the circular loop L. The dot-dashed line shows the looping probability considering only elastic bending energy; whereas dotted, solid and dashed lines show the probability for circular looping of DNA with elastic as well as Debye-Hückel electrostatic energy under 10, 50 and 150 mM monovalent salt conditions respectively.

the total electrostatic energy is given by $U(\xi) = \zeta L \psi(\xi)$, where

$$\psi(\xi) = \int_0^{\pi/2} d\phi \left[\frac{e^{-\xi \sin\phi}}{\sin\phi} - \frac{e^{-\xi\phi}}{\phi} \right] \tag{A.10}$$

Asymptotic solutions of ψ:

Case I: Small loop, $\xi \ll 1$.

$$\psi(\xi \ll 1) \approx \sum_{m=0}^{\infty} \frac{\xi^m}{m!} \int_0^{\pi/2} d\phi \left[(\sin\phi)^{m-1} - \phi^{m-1} \right] = \ln(4/\pi) + \sum_{m=0}^{\infty} \alpha_m \xi^{m+2} \tag{A.11}$$

where

$$\alpha_m = \frac{1}{(m+2)!} \left[\frac{\sqrt{\pi}}{2} \frac{\Gamma(\frac{m+3}{2})}{\Gamma(\frac{m+4}{2})} - \frac{(\pi/2)^{m+3}}{m+3} \right] \tag{A.12}$$

Case II: Large loop, $\xi \gg 1$.

We make a transformation $t = \sin\phi$ in Eq. (A.10) and note that the leading order contribution to the integral in the limit $\xi \gg 1$ comes from $t \ll 1$.

$$\psi(\xi) = \int_0^1 dt e^{-\xi t}\left[\frac{t}{2} + \frac{3t^3}{8} + \cdots\right] + Ei(-\xi) - Ei(-\pi\xi/2)$$
$$\Rightarrow \psi(\xi \gg 1) \approx \frac{1}{2\xi^2} + \mathcal{O}(\xi^{-4}) \tag{A.13}$$

where $Ei(x)$ is the exponential integral function such that $Ei(x \to \infty) \sim e^x/x$.

We propose a fit function that smoothly goes from one asymptote to the other:

$$\psi = \frac{a_1}{\xi^2}\left[1 - e^{-a_2\xi^2}\right] \tag{A.14}$$

Numerically fitting Eq. (A.14) to Eq. (A.10) we get: $a_1 = 0.47$ and $a_2 = 0.51$ [Fig. A.1(a)].

Figure A.1 shows the comparison of numerically evaluated values of normalized loop size ξ with the fit function. The electrostatic self energy of the circular loop of size L is given by $U = \zeta L\psi(\kappa L/\pi)$, shown in Fig. A.1(b) for three salt conditions (10, 50 and 150) mM, where the corresponding Debye-Hückel effective linear charge density of DNA used are (2.82, 4.98 and 8.86) nm^{-1}.

We write the probability of length distribution of circular loops [87]:

$$P(L) = \pi^2\left(\frac{2A}{L}\right)^{-6} e^{-\beta(E_b+U)} \tag{A.15}$$

where $A \approx 50$ nm is the bending persistence length of DNA; $\beta E_b = 2\pi^2 A/L - 0.257L/A$, is the elastic bending energy of the circular loop including fluctuations of the ends and U is the electrostatic energy associated with the circular loop.

A.3 Two helically intertwined dsDNAs or DNA braids

Eq. (A.8) ignores the free energy contribution from the interaction between all the length elements of the plectoneme strands but the pair of nearest-neighbors. In Ref. [76], the authors numerically evaluated the correction to Eq. (A.8) due to the presence of the neighboring chain but the authors ignored the effect of the self energy of the helix, which we look at in the following work. First, we need the spherically symmetric solution to the Debye-Hückel equation [Eq. (A.3)].

$$\Psi(\mathbf{r}) = \zeta \frac{e^{-r/\lambda_D}}{r} \tag{A.16}$$

where $\zeta \equiv \ell_B \nu$. Following Ref. [76] we parametrize the plectoneme to write the distance between any two points (M and N) on the opposing strands of the plectoneme,

$$\rho(u) = \sqrt{2r^2(1+\cos u) + p^2 u^2} \tag{A.17}$$

The Debye-Hückel potential exerted by the opposing strand on point M is given by

$$\Psi_n = \zeta \int_{-\infty}^{\infty} ds \frac{e^{-\rho/\lambda_D}}{\rho} = \zeta\sqrt{1+4\mu} \int_0^{\infty} du \frac{e^{-w\phi(u)}}{\phi(u)} \tag{A.18}$$

where $\mu \equiv p^2/4r^2$, $w \equiv 2r/\lambda_D$ and $\phi(u) = \sqrt{(1+\cos u)/2 + \mu u^2}$. The following solution was proposed in Ref. [76] for Eq. (A.18),

$$\Psi_n = \zeta\sqrt{\frac{2\pi}{w}} e^{-w} \left[1 + \frac{m_1}{\mu} + \frac{m_2}{\mu^2}\right] \tag{A.19}$$

where $m_1 = 0.207$ and $m_2 = 0.054$ was chosen to match the numerical solution to Eq. (A.18).

Self-electrostatic energy of DNA braids The minimum energy state of a charged rod in an ionic solution is the straight configuration, such that the electrostatic repulsion between the

length elements of the rod is minimized. External work has to be done against the electrostatic repulsion to keep the rod bent in a helix shape, which is stored as the self energy. Prior theoretical works on braids [45, 68] as well as supercoiled single DNA plectonemes [11, 65, 66, 76] ignore the Debye-Hückel self energy of helically intertwined DNAs. In the following, we numerically compute the self energy in braids and propose an empirical fit function. We find that the energy contribution from Debye-Hückel-type self-interaction of the coaxial helices increases for helices with higher aspect ratio (ratio of radius to pitch), and is non-negligible for braids under 2 pN force at 100 mM monovalent salt concentration [Figs. A.2(c) and A.2(d)].

We parameterize the helical arc length by a rotation angle u such that the distance ρ between any two points on the helix is given by,

$$\rho(u) = \sqrt{2R^2(1-\cos u) + P^2u^2} \equiv 2R\varphi(u), \tag{A.20}$$

where R and $2\pi P$ are radius and pitch of the helix respectively. The authors of Ref. [76] used a similar parameterization scheme to study the enhancement of electrostatic interaction energy between two helically intertwined strands of a plectoneme as a function of the helix angle.

We write Ψ_s, the total self-electrostatic potential for a braid of length L by integrating the spherically symmetric solution of the Debye-Hückel equation:

$$\beta\Psi_s = \frac{\ell_B\nu^2}{2}\int_{-L}^{L} ds_1 \int_{-L}^{L} ds_2 \left[\frac{e^{-\rho(u)/\lambda_D}}{\rho(u)} - \frac{e^{-\rho(u\to 0^+)/\lambda_D}}{\rho(u\to 0^+)}\right]. \tag{A.21}$$

Now, the self-electrostatic potential per unit length A of the braid is given by $\mathcal{U}_s \equiv \Psi_s/(L/A)$, such that

$$\beta\mathcal{U}_s = \frac{\zeta}{2}\int_0^{\infty} du \left[\sqrt{1+a^2}\frac{e^{-w\varphi(u)}}{\varphi(u)} - \frac{2}{u}e^{-\frac{wu}{2}\sqrt{1+a^2}}\right], \tag{A.22}$$

where $a \equiv P/R$, is the inverse of the aspect ratio of the helix, and $w \equiv 2R/\lambda_D$, is the scaled diameter of the helix. We have defined $\zeta = 2A\ell_B\nu^2$, where ℓ_B is the Bjerrum length and ν is the

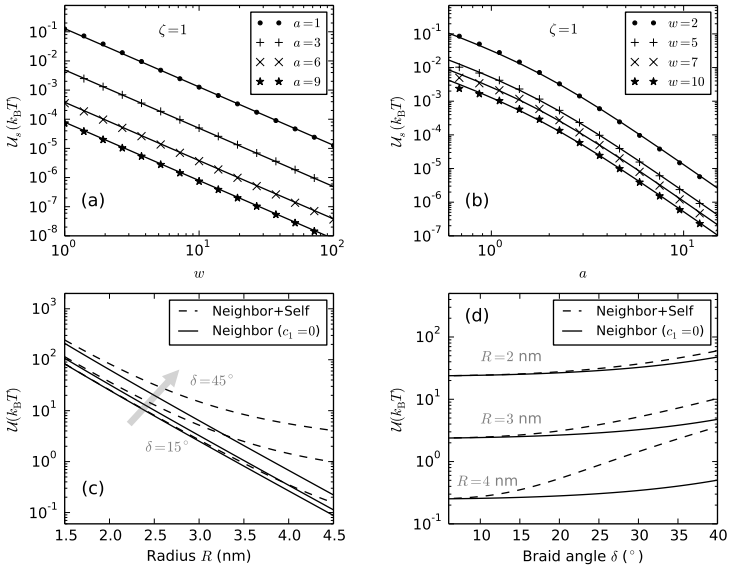

Figure A.2: Electrostatic interactions in a DNA braid. (a) Comparison of numerical evaluation (points) of $\mathcal{U}_s$ (Eq. A.22) with the proposed empirical solution [Eq. (A.23)] (solid lines) with $\zeta = 1$, $c_1 = 0.042$ and $c = 0.312$, as a function of w for various values of $a = 1, 3, 6$ and 9. (b) $\mathcal{U}_s$ versus a for $w = 2, 5, 7$ and 10 showing that the empirical function is a good fit to the numerical solution. (c) Plot of the total electrostatic potential $\mathcal{U}$ per unit braid length A [Eq. (A.24)] with (dashed lines) and without (solid lines) the self-interaction component (the second term in Eq. (A.24) containing the self interaction can be set to zero by putting $c_1 = 0$) versus braid radius (R) over a range of braiding angles $\delta = 15°, 30°$ and $45°$ (Table 4.1). We used $\zeta = 2700$ (corresponding to 100 mM Na^+, Table 4.1), $c_2 = 0.312$ and c_1 was chosen to be either 0 (only neighbor interaction plot, solid lines) or 0.042 (neighbor and self interaction plot, dashed line). (d) Comparison of the self and the neighbor components of the total electrostatic potential as a function of braiding angle δ for braid radii $R = 2, 3$ and 4 nm.

effective linear change density in inverse-length units [Eq. (C.4) and Table 2.1]. We have shifted the reference of the free energy to subtract the contribution of the self-electrostatic energy of a straight rod.

We numerically evaluate the self-energy integral [Eq. (A.22)] for a range of practically relevant values of w and a, and propose the following empirical function that captures the behavior of the self-energy functional,

$$\beta\mathcal{U}_s = \frac{\zeta}{w^2}\left[\frac{4c_1}{a^2 + c_2 a^4}\right]. \tag{A.23}$$

Figure A.2(a)-(b) shows the comparison between the numerically evaluated value of the integral [Eq. (A.22)] and the empirical fit-function [Eq. (A.23)] with $\zeta = 1$, $c_1 = 0.042$ and $c_2 = 0.312$.

Finally, we write the electrostatic potential energy per unit length A of braid with radius R and helix angle δ:

$$\beta\mathcal{U}(R,P) = \zeta K_0\left(\frac{2R}{\lambda_D}\right)(1 + m_1\tan^2\delta + m_2\tan^4\delta) + \frac{\zeta c_1\lambda_D^2\tan^4\delta}{R^2(c_2+\tan^2\delta)}. \tag{A.24}$$

The first term is the interaction potential with $m_1 = 0.828$, $m_2 = 0.864$ [76], and the second term is the self-energy contribution, where $c_1 = 0.042$ and $c_2 = 0.312$.

Figure A.2(c)-(d) shows the comparison between the self and the interaction energy component of the total electrostatic potential for various braid radii (R) and braiding angles (δ). For typical values of the braiding angle $\approx 25^\circ$ (Table 2.1), the self-energy contribution becomes significant in braids with radius > 3 nm, which corresponds to to < 2 pN stretching tension on the braid at 100 mM Na^+ [Fig. 2.3(a)].

Appendix B

Intertwined DNAs (Supplementary figures)

The 6 kb DNA molecules used in braiding were made using one digoxigenin and one biotin 5′-labeled oligonucleotide as primers for PCR [16, 53]. The purified labeled DNA was diluted to 1 ng/μL in phosphate buffered saline (PBS) (17-516Q, Lonza) and incubated with 1 μm streptavidin coated paramagnetic beads (Dynabeads MyOne Streptavidin T1, 65601, Invitrogen) diluted to 0.6 mg/mL in 0.4 mg/mL Bovine Serum Albumin (BSA) (A7030 10G, Sigma-Aldrich). For the incubation, 1 μL of DNA was incubated with 2 μL of beads for 10 min. The beads and DNA were then diluted with 45 μL of PBS and passed into the flow cell [88]. Following a 20 min incubation in the flow cell, the slide was placed on the microscope and magnetic force applied. The bright-field magnetic tweezer microscope used for these experiments has been described before as part of Ref. [88]. To wash the DNA into the experimental buffer (100 mM NaCl, 20mM Tris-HCl pH 8), at least 6 flow cell volumes were passed through the flow cell. Following the identification of a potential DNA tether, the magnets were rotated to introduce turns (catenations) into the DNA. Single tethers would not supercoil, since the DNA only had single attachments to both the bead and glass coverslip, so only double tethers showed a change in extension with magnet rotation. After a double tether was identified, a lookup table was

collected to measure the extension of the tethered bead compared to a fixed surface bead. After exactly centering the double tether using the peak of the extension versus catenation curve as a guide, a force calibration was done using the Brownian motion of the bead [16, 97]. This calibration was used to estimate the force applied to the bead for the experiments. Data were collected at 200 Hz using a CMOS camera (PL-B741U, PixeLink), while the magnets were rotated to introduce or remove catenations in the braid[47].

See Figs. B.1-B.6 for experimental data.

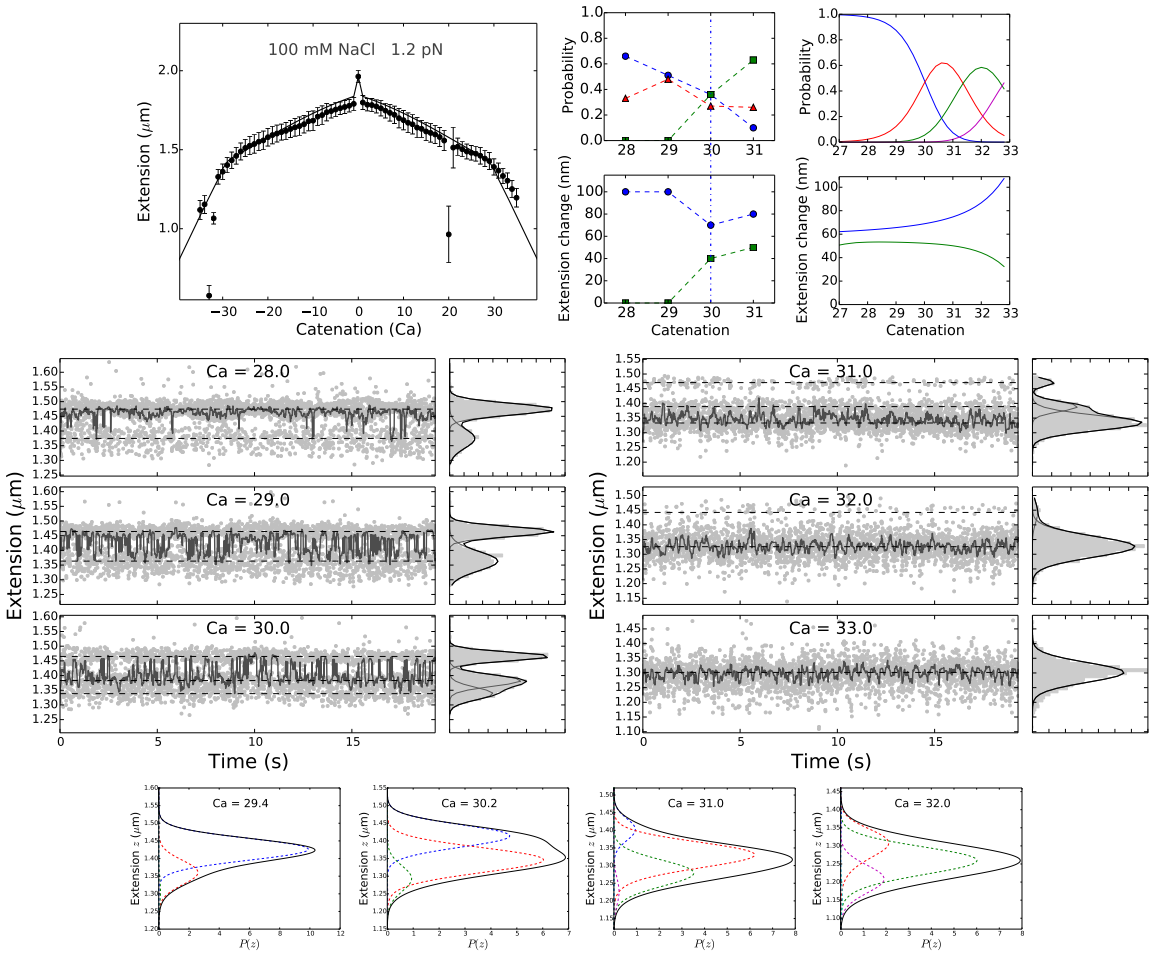

Figure B.1: (a) Extension versus catenation for a braid under 1.2±0.12 pN force under 100 mM NaCl. Points are experimental data and the solid line is the theoretical prediction using $L = 2.4$ μm and an intertether distance of 0.77 μm ($d = 0.32L$). (b) Comparison between theory and experiments for relative probability of occupancy of the three states: highest extension (blue), lower extension (red) and lowest extension (green) as a function of catenation near the buckling transition; and (c) the distance between the successive peaks in extension histograms for the three states as a function of catenation, showing the extension jump associated with nucleation of the first (blue) and the second (green) plectoneme domains. (d) Time series of extension showing dynamic switching near the buckling transition, where the shaded points are experimental data collected at 200 Hz and the dark points are the median-filtered time series using a 0.1 sec time window to emphasize dynamic-switching. Corresponding right panels are histogram of raw data (10 nm bins) which are fitted to a sum of Gaussian distributions showing equilibrium transitions between two or three extension states. (e) Theoretically predicted extension histograms near the buckling transition. The black solid line shows the total extension distribution, while the dashed lines show the contribution from the straight braid (blue) and the buckled braid with one (red), two (green), three (magenta) and four (cyan) plectoneme domains.

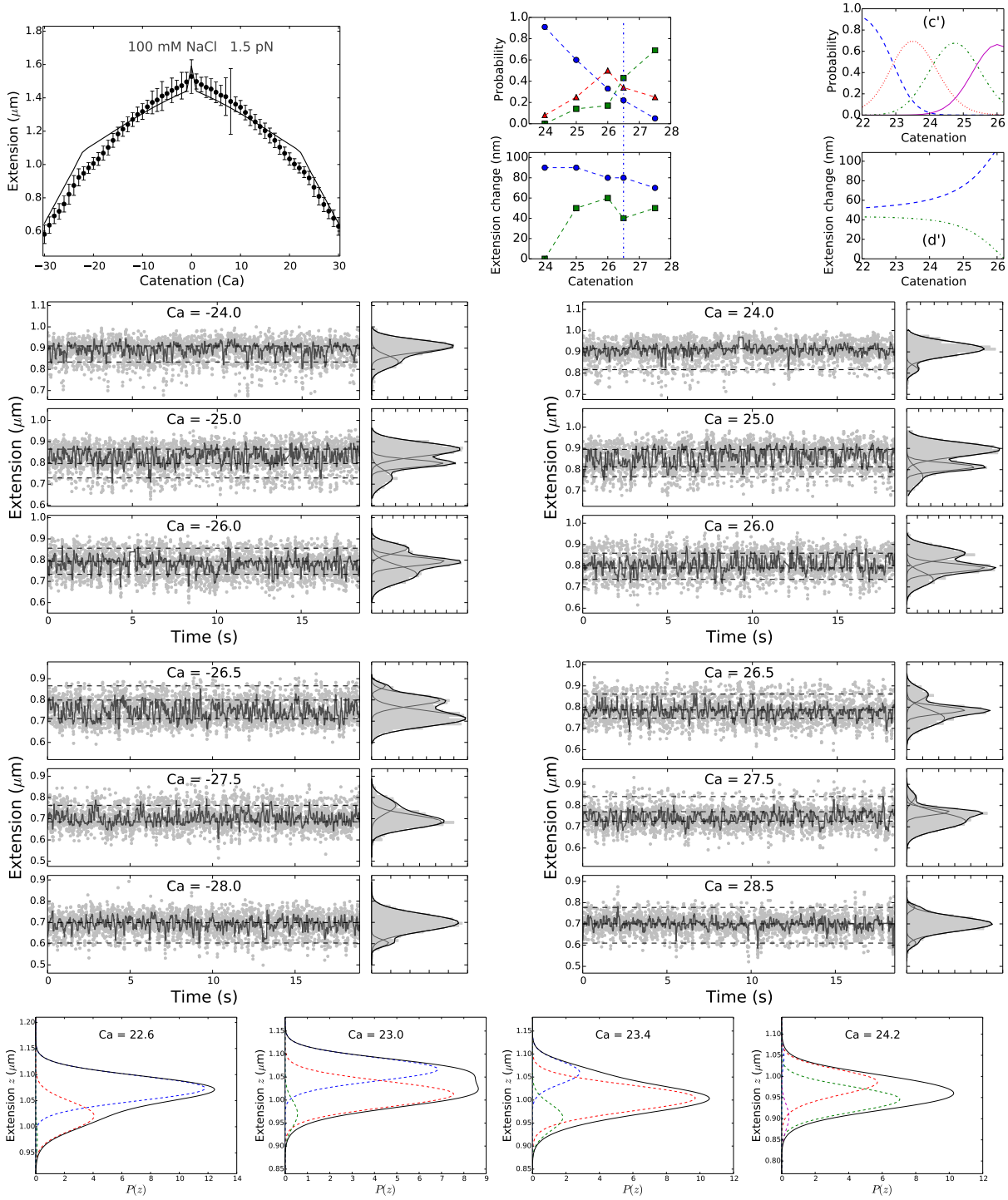

Figure B.2: Experiments for 6 kb DNA braids under 1.5±0.15 pN at 100 mM NaCl. Theoretical predictions are for $L = 1.9$ μm, intertether distance 0.74 μm ($d = 0.39L$), $f = 1.5$ pN and 100 mM monovalent salt.

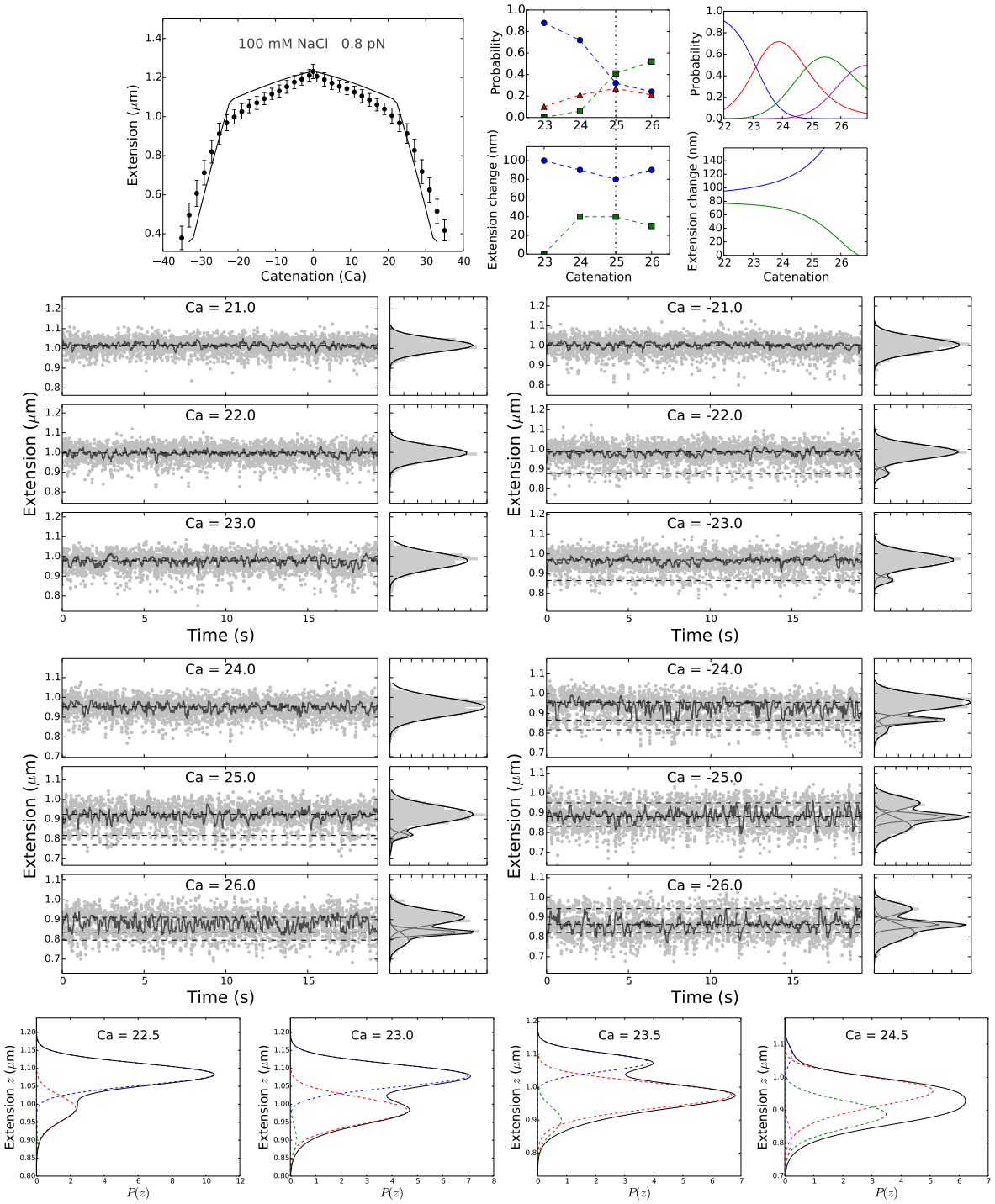

Figure B.3: Experiments for 6 kb DNA braids under 0.8±0.08 pN at 100 mM NaCl. Theoretical predictions are for $L = 1.6$ μm, intertether distance 0.19 μm ($d = 0.12L$), $f = 0.8$ pN and 100 mM monovalent salt.

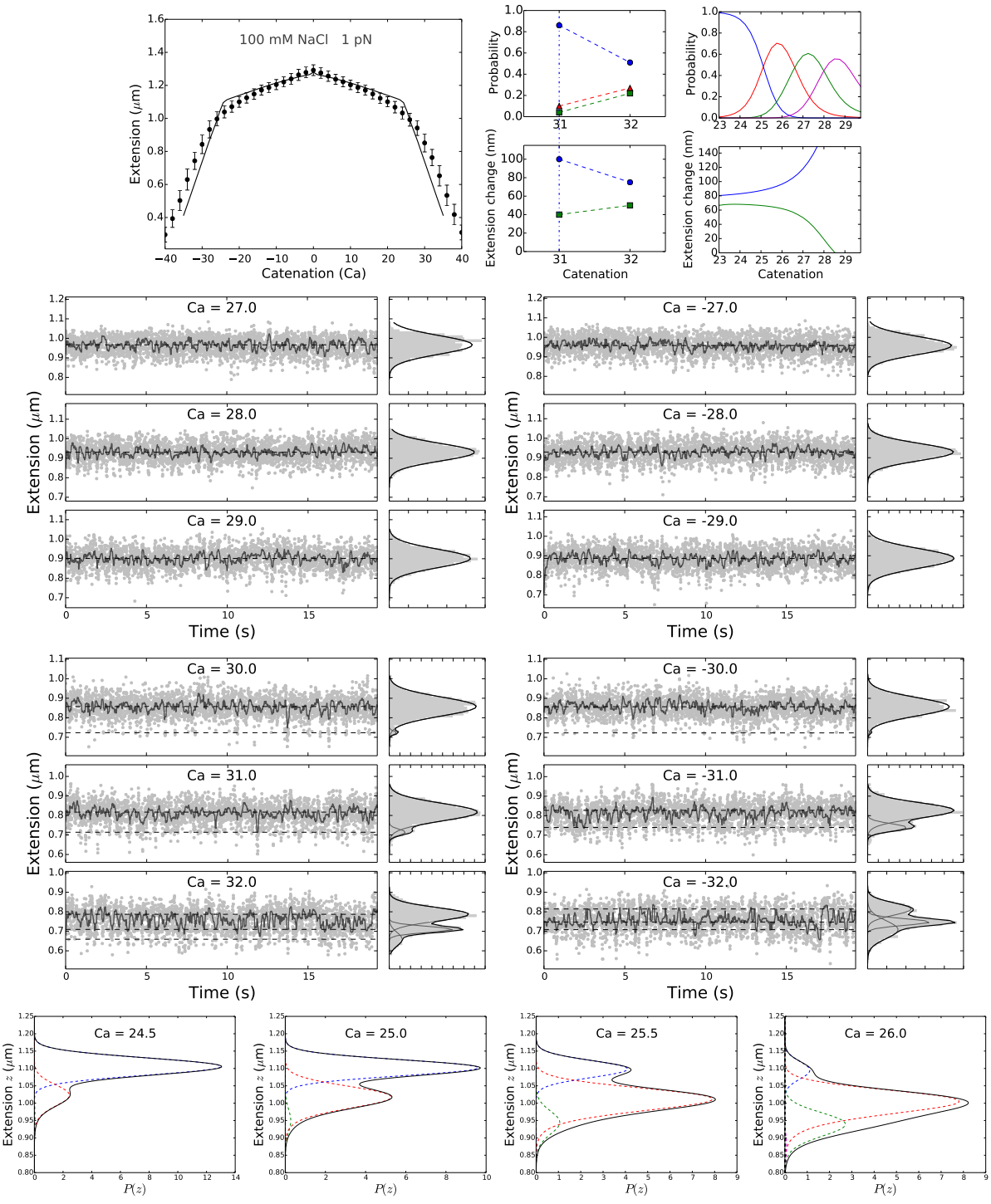

Figure B.4: Experiments for 6 kb DNA braids under 1±0.1 pN at 100 mM NaCl. Theoretical predictions are for $L = 1.6$ μm, intertether distance 0.19 μm ($d = 0.12L$), $f = 1$ pN and 100 mM monovalent salt.

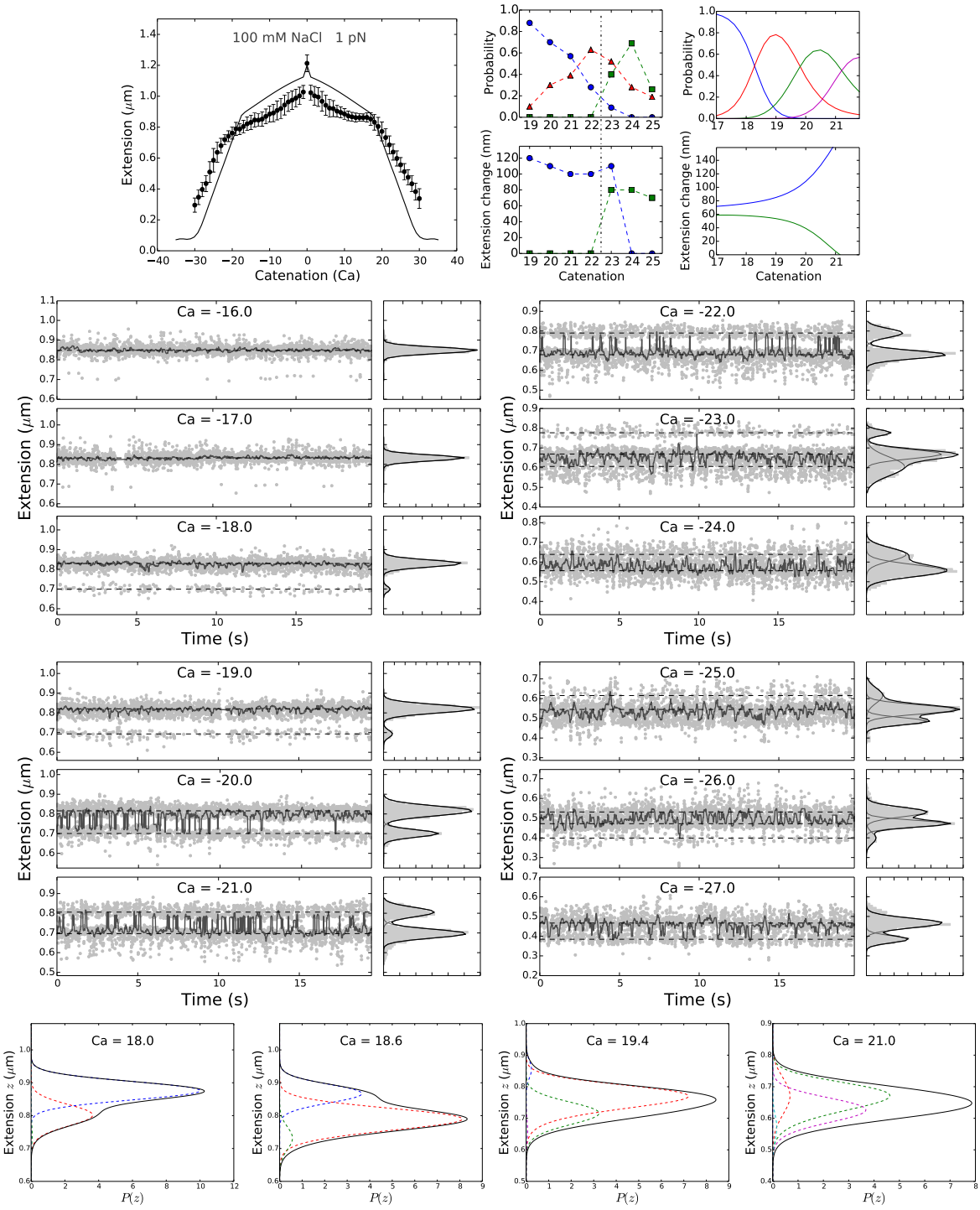

Figure B.5: Experiments for 6 kb DNA braids under 1±0.1 pN at 100 mM NaCl. Theoretical predictions are for $L = 1.5$ μm, intertether distance 0.46 μm ($d = 0.31L$), $f = 1$ pN and 100 mM monovalent salt.

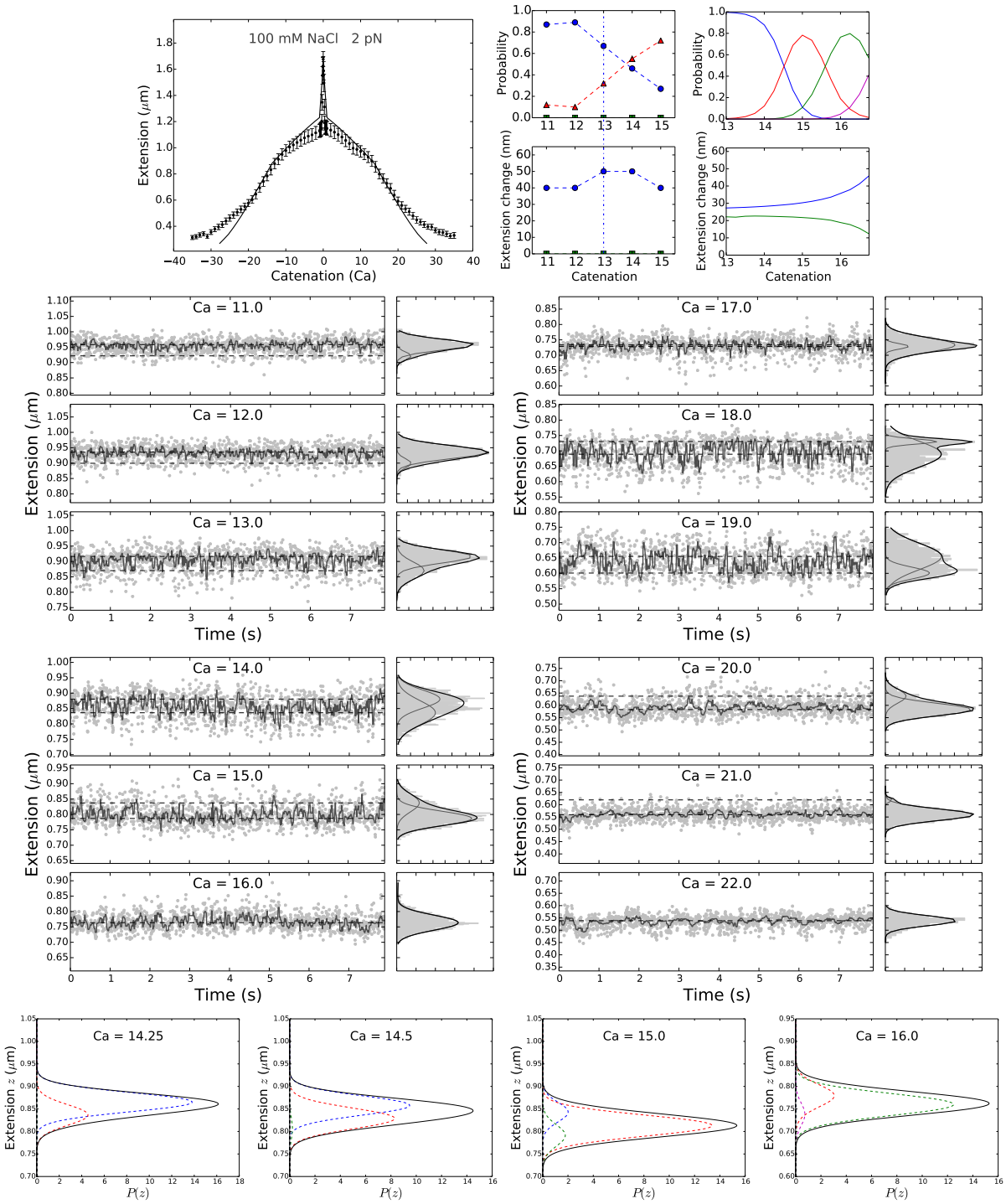

Figure B.6: Experiments for 6 kb DNA braids under 2±0.2 pN at 100 mM NaCl. Theoretical predictions are for $L = 1.9$ μm, intertether distance 1.12 μm ($d = 0.62L$), $f = 2$ pN and 100 mM monovalent salt.

Appendix C

Supercoiled double-helix DNA

C.1 Plectoneme Hamiltonian

A plectoneme structure made up of total DNA length L_p, can be considered as two helices of length $L_p/2$ wrapped around each other. The total Hamiltonian of the plectoneme structure is given by,

$$\beta\mathcal{H} = \int_0^{L_p/2A} d\xi \left[\frac{1}{2} \left(\left| \frac{d\hat{\mathbf{t}}_1}{d\xi} \right|^2 + \left| \frac{d\hat{\mathbf{t}}_2}{d\xi} \right|^2 \right) + U(\mathbf{r}_1, \mathbf{r}_2) - \beta A f \left\{ (\hat{\mathbf{t}}_1 - \hat{\mathbf{t}}_{1o}) \cdot \hat{\mathbf{e}} + (\hat{\mathbf{t}}_2 - \hat{\mathbf{t}}_{2o}) \cdot \hat{\mathbf{e}} \right\} \right] \quad \text{(C.1)}$$

where an external force f is applied at the two ends of the plectoneme in a direction perpendicular to the axis of the plectonemic helix. This force can be interpreted as the stretching tension required at the end of a plectoneme that keeps the helices from unwinding themselves. Since, the force is in a direction perpendicular to the axis of the plectoneme, there is no corresponding force-extension energy contribution, however, the tension plays an important role in controlling the transverse fluctuations of the DNA inside the plectoneme.

In Eq. (C.1), the first parenthesized term containing the square of the local curvature of the two intertwining helices inside the plectoneme corresponds to the total elastic bending energy. The second term contains the total electrostatic energy contribution from close proximity of

the DNAs inside the plectonemic structure [12, 46, 75, 76]. And, the third term contains the coupling of the transverse fluctuations of the DNAs to the external force, where $\hat{\mathbf{e}}$ is the direction of the plectoneme axis.

The above Hamiltonian is similar to the one analyzed for two intertwined DNAs or braids in Ref. [46], except for that braids have a force-extension energy [46], while plectonemes do not. Note in Eq. (C.1), the external force is only coupled to the transverse tangents.

C.1.A Oscillating Reference Frame

We consider two sets of orthonormal triads: $(\hat{\mathbf{t}}_{io}, \hat{\mathbf{t}}_{i\perp r}, \hat{\mathbf{t}}_{i\perp\theta})$, where $i \in \{1, 2\}$ corresponds to the two intertwining helices in the plectoneme. $\hat{\mathbf{t}}_{io}$ points along an average helical shape defined by two helix parameters: radius (r) and pitch ($2\pi p$); $\hat{\mathbf{t}}_{i\perp r}$ points towards the axis of the plectoneme; whereas $\hat{\mathbf{t}}_{i\perp\theta} \equiv \hat{\mathbf{t}}_{io} \times \hat{\mathbf{t}}_{i\perp r}$. See Ref. [46] for a similar calculation done in the context of intertwined DNAs or DNA braids.

We expand the tangent vectors to harmonic order, about a mean helical shape $\hat{\mathbf{t}}_{io}$:

$$\hat{\mathbf{t}}_i = \left[1 - \frac{\mathrm{t}_{i\perp}^2}{2} + \mathcal{O}(\mathrm{t}_{i\perp}^4)\right] \hat{\mathbf{t}}_{io} + \mathbf{t}_{i\perp} \tag{C.2}$$

where the mean helical shape depends on two helix parameters radius (r) and pitch ($2\pi p$):

$$\hat{\mathbf{e}} \cdot \hat{\mathbf{t}}_{io} = \cos\alpha \qquad \hat{\mathbf{e}} \cdot \hat{\mathbf{t}}_{i\perp\theta} = \sin\alpha \qquad \hat{\mathbf{e}} \cdot \hat{\mathbf{t}}_{i\perp r} = 0 \tag{C.3}$$

Here, $\alpha \equiv \arctan(r/p)$, is the braiding angle.

C.1.B Electrostatic Interactions

The electrostatic energy contribution due to DNA-DNA repulsion in a helical structure $\mathcal{U}$ computed from a Debye-Hückel-type interaction is as follows.

$$\mathcal{U}(r, \alpha) \equiv \zeta K_0\left(\frac{2r}{\lambda_D}\right) v(\tan\alpha) + \frac{\zeta\lambda_D^2 c_1 \tan^4\alpha}{r^2(c_2 + \tan^2\alpha)} \tag{C.4}$$

The first term in Eq. (C.4) is the contribution from electrostatic interaction with the neighboring strand in the superhelix. Here, $K_0(x)$, the modified Bessel function of the first kind, denotes the solution for two parallel strands [12, 75], and $v(x) = 1+(0.828)x^2+(0.864)x^4$, is an enhancement factor that accounts for the effect of helical curvature [76]. The second term accounts for the self interaction of the helically-bent polymer in the plectoneme, where $c_1 = 0.042$ and $c_2 = 0.312$ [46]. λ_D is the Debye length of the ionic solution, and we define $\zeta = 2A\ell_B\nu^2$, a parameter that depends on the effective linear charge density of DNA ν [63, 74, 76, 77], and $\ell_B \approx 0.7$ nm is the Bjerrum length of water at $290K$. We have used $\nu = 1.97$, 6.24, 8.85, 10.23, and 26.6 nm^{-1} respectively corresponding to 0.01, 0.1, 0.15, 0.2, and 0.5 M monovalent salt concentrations [46, 63, 65, 76].

We define the electrostatic part of the Hamiltonian as:

$$U(\mathbf{r_1}, \mathbf{r_2}) \equiv \mathcal{U}(r + A\mathrm{w}(\xi), \alpha) = \mathcal{U}_0 + g\mathrm{w} + \eta\mathrm{w}^2 + \mathcal{O}(\mathrm{w}^3) \tag{C.5}$$

where $\mathcal{U}_0 \equiv \mathcal{U}(r, \alpha)$; $g \equiv A\partial\mathcal{U}/\partial r$, and $\eta \equiv (A^2/2)(\partial^2\mathcal{U}/\partial r^2)$ is the effective modulus of the electrostatic potential. The first term gives the mean electrostatic energy per unit length A of plectoneme with fixed radius and pitch, while the subsequent terms are corrections for small uniform deviation in the braid radius. Here, $\mathbf{w}(\xi)$ is the normalized radial deformation:

$$\mathbf{w}(\xi) = \hat{\mathbf{t}}_{1\perp r} \int_0^{\xi} \frac{1}{2}[\mathrm{t}_{1\perp r} + \mathrm{t}_{2\perp r}]d\bar{\xi} \tag{C.6}$$

where $\mathrm{t}_{i\perp r}$ are given by Eq. (C.2) and we assume the boundary condition $\mathbf{w}(0) = 0$. Note, the above definition of normalized radial deformations $\mathbf{w}(\xi)$ assumes a parallel configuration of the two strands.

C.1.C Thermal Fluctuations

Perturbative expansion of the Hamiltonian. Following the above equations, the plectoneme Hamiltonian [Eq. (2.1)] can be expanded as a contribution from the mean-field helix

structure and thermal fluctuations [46]:

$$\beta\mathcal{H} = \frac{L_p}{2A}\left[\kappa^2 + \mathcal{U}_0\right] + \frac{1}{2}\int_0^{L_p/2A} d\xi \left[\left|\frac{d\hat{\mathbf{t}}_{1\perp}}{d\xi}\right|^2 + \left|\frac{d\hat{\mathbf{t}}_{2\perp}}{d\xi}\right|^2 + (\mu - \kappa^2)\left(|\hat{\mathbf{t}}_{1\perp}|^2 + |\hat{\mathbf{t}}_{2\perp}|^2\right) + 2\eta \mathrm{w}^2\right] \tag{C.7}$$

where $\mu \equiv \beta A f \cos\alpha$, is the effective tension in the plectoneme; $\kappa \equiv Ar/(r^2+p^2)$, is the total mean-field curvature per unit persistence length of the strands. Note that we have truncated the expansion to the harmonic order, i.e., we ignore $\sim \mathcal{O}(\mathrm{t}_{1\perp}^3, \mathrm{t}_{2\perp}^3)$ terms. The mean-field term is the sum total of bending and electrostatic energy in the plectoneme [Eqs. (C.11) and (C.17)]. The sub-leading order term is the contribution from thermal fluctuations. We set the reference of the fluctuation free energy by setting the amplitude of the zero-momentum mode of transverse fluctuations to zero [46].

Free energy of fluctuations We construct a partition function for the plectoneme via a path integral over all the transverse tangent conformations:

$$Z_p = \int \mathcal{D}\mathbf{t}_{1\perp} \int \mathcal{D}\mathbf{t}_{2\perp} e^{-\beta\Delta\mathcal{H}}, \tag{C.8}$$

and get the free energy contribution of thermal fluctuations from the partition function [46],

$$-\ln Z_p = \frac{L_p}{2A}\left[\frac{3}{2}\sqrt{\mu} + \eta^{1/4}\cos\left(\frac{1}{2}\tan^{-1}\sqrt{\frac{4\eta}{\mu^2}-1}\right)\right]. \tag{C.9}$$

The first term gives the fluctuation free energy that depends only on the external tension, whereas, the second term depends on both the external tension and the salt concentration of the solution.

Note that there are four independent degrees of transverse fluctuations in a plectoneme structure: $\hat{\mathbf{t}}_{1\perp r}, \hat{\mathbf{t}}_{2\perp r}, \hat{\mathbf{t}}_{1\perp\theta}$, and $\hat{\mathbf{t}}_{2\perp\theta}$ [Fig. (**??**a)]. Three of them [$\hat{\mathbf{t}}_{1\perp\theta}, \hat{\mathbf{t}}_{2\perp\theta}$, and $(\hat{\mathbf{t}}_{1\perp r} + \hat{\mathbf{t}}_{2\perp r})/2$]are solely controlled by the tension μ; while, the fluctuations in the direction $(\hat{\mathbf{t}}_{1\perp r} - \hat{\mathbf{t}}_{2\perp r})/2$ are

controlled by both the external tension μ and the salt concentration via the parameter η. We have used the above expression for the fluctuation free energy inside the plectoneme structure in Eq. (3.5).

Radial fluctuations Fluctuations in the radius of the plectoneme are generated by displacement of one plectonemic strand relative to the other. As is the case for Gaussian fluctuations, the two-point correlation function of radial deformations decays exponentially: $\langle \mathrm{w}(\xi)\mathrm{w}(0)\rangle \sim \exp(-k\xi)$, where $k \sim O(\sqrt{\mu})$ [46]. The zero-distance behavior of the two-point correlation gives the radial fluctuations in the plectonemic superhelix (See Appendix B in Ref. [46]):

$$\sigma_r \approx A\langle|\mathrm{w}(0)|^2\rangle^{1/2} \approx A\eta^{-3/8}. \tag{C.10}$$

This suggests that a stronger electrostatic repulsion reduces fluctuations in the plectoneme radius.

C.2 Finite-Sized Supercoiled DNA

C.2.A Defect-free DNA

Buckled state: plectonemes and end loops The total free energy of the buckled state composed of m domains of plectoneme:

$$\beta E_p = \frac{2\pi^2 C(\mathrm{Tw}_p)^2}{L_p + m\gamma} + L_p\frac{A\sin^4\alpha}{2r^2} + 2m\sqrt{\rho\beta A f} + [(L_p + m\gamma)/2A]\mathcal{U}(r,\alpha) - \ln\Omega(m) \tag{C.11}$$

where the first term corresponds to DNA-twist energy contribution in the buckled state. The second and third terms respectively correspond to the net elastic energy of plectoneme superhelices and m end loops. The fourth term contains the total mean-field electrostatic contribution from the buckled state of the DNA. Finally, the last term in Eq. (C.11) corresponds to configuration

entropy of m plectoneme domains ($m \geq 1$) [46, 65], where

$$\Omega = \frac{(L_u/D)^m}{m!} \cdot \frac{(L_p/D)^{m-1}}{(m-1)!} \tag{C.12}$$

is the total number of energetically degenerate but distinguishable configurations arising from: (1) one-dimensional diffusion of a domain along the DNA contour length [the first combinatorial term in Eq. (C.12)]; and (2) the exchange of DNA length among the plectoneme domains [the second combinatorial term in Eq. (C.12)]. $D \equiv \sqrt{A/(\beta f)}$, is the force-induced correlation length, which we have used as the distinguishable length for plectoneme sliding.

Force-extended state The total energy of the force-coupled state is given by

$$\beta E_u = \frac{2\pi^2 C_f}{L_u} \mathrm{Lk}_u^2 - \beta f (L - L_p) \tag{C.13}$$

The first term corresponds to the total twist energy, where $C_f = C\left[1 - C/(4A\sqrt{\beta A f})\right]$ is the renormalized twist persistence length [60]. The second term is the extension energy of the DNA under external force f.

Thermal fluctuations The total fluctuation contribution is obtained from summing the contributions from the force-extended and plectoneme states:

$$-\ln Z = -\ln Z_p - \ln Z_u, \tag{C.14}$$

where Z_p corresponds to the plectoneme [Eq. (C.9)]; and Z_u corresponds to the force-extended state, which is computed by taking the $\eta \to 0$ limit of Eq. (C.8). $-\ln Z_u$ is the second term in Eq. (3.5).

Numerical scheme Various quantities of interest can be numerically computed from the partition function $\mathcal{Z}$ in Eq. (3.6) as follows:

$$\langle z\rangle = -\frac{1}{\mathcal{Z}}\Big[\frac{\partial \mathcal{F}}{\partial f}e^{-\beta\mathcal{F}(0,0)} + \sum_{m,L_p}\frac{\partial \mathcal{F}}{\partial f}e^{-\beta\mathcal{F}(L_p,m)}\Big], \tag{C.15a}$$

$$\langle m\rangle = \frac{1}{\mathcal{Z}}\sum_{m,L_p} m e^{-\beta\mathcal{F}(L_p,m)}, \tag{C.15b}$$

$$\beta\langle\tau\rangle = -\frac{1}{2\pi}\frac{\partial \ln \mathcal{Z}}{\partial(\Delta\mathrm{Lk})}, \tag{C.15c}$$

$$\langle L_p\rangle = \frac{1}{\mathcal{Z}}\sum_{m,L_p} L_p e^{-\beta\mathcal{F}(L_p,m)}. \tag{C.15d}$$

Probability distributions For a given coexistence state (L_p, m), the probability distribution of $X \in \{z, \tau\}$ is given by

$$P_{L_p,m}(X) = \frac{1}{\sqrt{2\pi\sigma_X^2}}\exp\left[-\frac{(X-\overline{X})^2}{2\sigma_X^2}\right],$$

where the mean $\overline{X}$ and the standard deviation σ_X are obtained as follows:

$$\overline{z}(L_p,m) = -\frac{\partial \mathcal{F}}{\partial f}, \quad \overline{\tau}(L_p,m) = \frac{\partial \mathcal{F}}{\partial(2\pi\Delta\mathrm{Lk})},$$
$$\sigma_z^2(L_p,m) = -\frac{\partial^2 \mathcal{F}}{\beta\partial f^2}, \quad \sigma_\tau^2(L_p,m) = \frac{\partial^2 \mathcal{F}}{\beta\partial(2\pi\Delta\mathrm{Lk})^2}.$$

Now, the total probability distribution of X for a given linking number and force is obtained by summing the contributions from all the states considered in the partition sum:

$$P(X) = \frac{P_{0,0}}{\mathcal{Z}}e^{-\beta\mathcal{F}(0,0)} + \sum_{m=1,2,\cdots}\sum_{L_p}\frac{P_{L_p,m}}{\mathcal{Z}}e^{-\beta\mathcal{F}(L_p,m)} \quad = \mathcal{P}_0(X) + \sum_{m=1,2,\cdots}\mathcal{P}_m(X). \tag{C.16}$$

In the above equation, the probability distribution is written as a sum of contributions from the force-extended state, and the buckled domain containing m end loops.

C.2.B DNA with a Defect

Free energy of the buckled state The free energy of the plectoneme state, now including m mobile plectoneme domains and $m^\dagger$ pinned plectoneme domains, is given by

$$\beta E_p = \frac{2\pi^2 C(\mathrm{Tw}_p)^2}{(L_p + m\gamma + m^\dagger\gamma^\dagger)} + L_p\frac{A\sin^4\alpha}{2r^2} + 2\sqrt{\rho\beta A f}\left[m^\dagger(1-\varepsilon)+m\right] \\ +[(L_p + m\gamma + m^\dagger\gamma^\dagger)/2A]\mathcal{U} - \ln\left[\frac{(L_u/D)^m}{m!}\frac{(L_p/D)^{m+m^\dagger-1}}{(m+m^\dagger-1)!}\right], \tag{C.17}$$

where the first and the second terms respectively correspond to the total twist and bending energy in the plectoneme state. The third term gives the total elastic energy associated with m mobile and $m^\dagger$ pinned plectoneme end loops. The fourth term corresponds to net electrostatic energy of the buckled state. And the fifth term is associated with the total configuration entropy of plectoneme domains.

Theta function. We have used the usual definition of Theta function in the partition sum [Eq. (3.12)].

$$\Theta(L_p - 2L^*) = \begin{cases} 1, & \text{when } L_p > 2L^* \\ 0, & \text{when } L_p < 2L^* \end{cases} \tag{C.18}$$

Probability distributions Similar to Eq. (C.16), we write the probability distribution of X at a fixed linking number and fixed force as:

$$P(X) = \frac{P_{0,0,0}}{\mathcal{Z}^\dagger}e^{-\beta\mathcal{F}(0,0,0)} + {\sum_{L_p,m^\dagger,m}}' \frac{P_{L_p,m^\dagger,m}}{\mathcal{Z}^\dagger}e^{-\beta\mathcal{F}(L_p,m^\dagger,m)} = \mathcal{P}_{00} + {\sum_{m^\dagger,m}}'\mathcal{P}_{m^\dagger m}, \tag{C.19}$$

where the sum is now a restricted one as shown in the partition sum [Eq. (3.12)]. The contribution from the buckled state with m and $m^\dagger$ mobile and pinned plectonemes respectively is $\mathcal{P}_{m^\dagger m}$.

Appendix D

Chromosome structure and topology

D.1 Cylindrical loop-extruded structure

Radial profile of monomer volume fraction We consider a cylindrical shell of height d, inner radius r, and outer radius $r+\xi(r)$. Following the arguments originally proposed by Daoud and Cotton [126] for star polymers, the blobs corresponding the loops diffuse radially outward due to the monomer concentration gradient, and there is, on average, α blobs in the considered shell. Hence, the volume fraction of monomers in the shell, $\varphi(r)$, is equal to that inside the blob in that shell.

$$\varphi(r) = \frac{\alpha g(r) a^3}{rd\xi(r)} = \frac{g(r)a^3}{\xi(r)^3} \quad \Rightarrow \xi(r) = \sqrt{rd/\alpha} \tag{D.1}$$

where $g(r)$, the number of monomers per blob of size $\xi(r)$, obeys self-avoiding statistics: $\xi = ag^{3/5}$.

Osmotic pressure. In a semidilute solution, the osmotic pressure scales as, $\Pi \sim 1/\xi^3$ [26], suggesting, a radially decreasing profile for osmotic pressure.

$$\Pi(r) = \frac{k_B T}{a^3}\frac{\varphi(r)}{g(r)} = \frac{k_B T}{(rd/\alpha)^{3/2}} \tag{D.2}$$

Loop extension. The radial extension of a loop, R, containing n monomers can be obtained

from integrating the volume fraction $\varphi(r)$ [45, 132–134].

$$\int_0^R \mathrm{d}r\ \varphi(r) r d = n a^3 \quad \Rightarrow R = a \frac{n^{3/4}}{\alpha^{1/2}} \left(\frac{a}{d}\right)^{1/4} \tag{D.3}$$

Loop free energy. Free energy per loop is given by the number of blobs per loop, because each blob contributes $O(1)\ k_B T$ [26]. Equivalently, the free energy per loop may also be obtained from integrating the total osmotic pressure in the cylindrical volume accessible to each loop.

$$F = \int_0^R \mathrm{d}r\ \Pi(r) r d = k_B T \alpha^{5/4} n^{3/8} \left(a/d\right)^{5/8} \tag{D.4}$$

Tension along the backbone. Overlapping loops generate a tension along the backbone which we estimate from the free energy per unit length of the backbone.

$$f = -\partial F / \partial d = \frac{k_B T}{a} \alpha^{5/4} n^{3/8} \left(a/d\right)^{13/8} \tag{D.5}$$

High monomer density along the backbone. The volume fraction of the monomers along the backbone, φ_m, is uniform, and is given by,

$$\varphi_m = \frac{g_f a^3}{\xi_f^3} = \frac{m a^3}{\xi_f^2 d} \quad \Rightarrow d = (m/g_f)\xi_f \tag{D.6}$$

where g_f is the number of monomers in a blob of size $\xi_f = k_B T / f$, such that $\xi_f = a g_f^{3/5}$. The number of monomers per backbone segment between two loops is m and the linear distance between two adjacent loop anchors is d.

Persistence length A thermally-excited bend generates a curvature κ along the cylindrical brush axes, that has a convex and a concave side. The volume accessible to the loop monomers in the concave (convex) side is smaller (larger) than the unperturbed case by $\kappa R \ll 1$. This perturbs the monomer volume fractions: $\langle\varphi\rangle(1 \pm \kappa R)$, where the upper/lower signs are for the

concave/convex sides respectively, and $\langle\varphi\rangle = na^3/(R^2d)$ is the average unperturbed volume fraction.

The free energy of a loop depends on the average volume fraction as, $F = k_BTn\langle\varphi\rangle^{5/4}$. The perturbation energy due to a curvature κ for a cylindrical brush with persistence length ρ is given by, $k_BT\rho\kappa^2 d$. Hence, we get the persistence length:

$$\rho = n\langle\phi\rangle^{5/4}R^2/d = a\alpha^{1/4}n^{15/8}\,(a/d)^{17/8} \tag{D.7}$$

D.2 Entanglements

Entanglements in confinement The correlation length in confinement ξ_c, scales with g_c, the number of chromosome monomers in confinement: $\xi_c = \rho^{2/5}(w\rho^2R)^{1/5}g_c^{3/5}$. The number density of chromosome monomers is uniform in the confined volume.

$$\frac{N'}{D^3} = \frac{g_c}{\xi_c^3} \quad \Rightarrow \xi_c = a\phi^{-3/4}\left(\frac{R_F(N)}{R_F(N')}\right)^{5/4} \tag{D.8}$$

The number of confinement blobs

$$N_{blobs} = N'/g_c = N\phi^{5/4}\left(\frac{R_F(N')}{R_F(N)}\right)^{15/4} \tag{D.9}$$

scales positively with the chromatin volume fraction ϕ.

D.3 Optimal chromosomes

The optimal loop size for loops with maximum valency,

$$(n_\star)_{\alpha_{max}} = (d/a)^{2/3}(N/\mathrm{k})^{5/9} \tag{D.10}$$

Maximum valency for optimal loops,

$$(\alpha_{max})_{n_\star} = (d/a)^{5/6}(N/\mathrm{k})^{5/18} \tag{D.11}$$

The maximum core size for optimal loops of size $n_\star$ is given as

$$(r_0)_{n_\star,\alpha_{max}} = \sqrt{(n_\star)_{\alpha_{max}} a/d} = (a/d)^{1/6}(N/\mathrm{k})^{5/18} \tag{D.12}$$

Inter-chromosomal entanglements per chromosome for optimal loops is as follows.

$$\left(\langle \mathrm{Ca}^2 \rangle\right)_{n_\star} = \frac{\phi^{5/4} w^{3/4} (d/a)^{57/256}}{N_e\ \alpha^{15/128}} \left(\frac{N}{\mathrm{k}}\right)^{211/256} \tag{D.13}$$

For saturated chromosomes (loops with maximum valency α_{max}),

$$\left(\langle \mathrm{Ca}^2 \rangle\right)_{n_\star,\alpha_{max}} = \frac{\phi^{5/4}}{N_e} w^{3/4} \left(\frac{d}{a}\right)^{1/8} \left(\frac{N}{\mathrm{k}}\right)^{19/24} \tag{D.14}$$

Stretching modulus. The maximum stretching modulus of saturated chromosomes,

$$(f_0)_{n_\star,\alpha_{max}} = (k_B T/a)(a/d)^{1/3}(N/\mathrm{k})^{5/9} \tag{D.15}$$

Contour length and width. The contour length and width per chromosome for optimal loops

$$\left(L'/\mathrm{k}\right)_{n_\star} = a\alpha^{1/4}(d/a)^{1/8}(N/\mathrm{k})^{3/8} \tag{D.16}$$

$$(R)_{n_\star} = \frac{a(d/a)^{13/32}}{\alpha^{11/16}} \left(\frac{N}{\mathrm{k}}\right)^{15/32} \tag{D.17}$$

and for saturated chromosomes,

$$\left(L'/\mathrm{k}\right)_{n_\star,\alpha_{max}} = a(d/a)^{1/3}(N/\mathrm{k})^{4/9} \tag{D.18}$$

$$(R)_{n_\star,\alpha_{max}} = a(a/d)^{1/6}\,(N/\mathrm{k})^{5/18} \tag{D.19}$$

Bacterial DNA: cylindrical monomers

$$R_F(N) = a^{2/5}(wa^2b)^{1/5}N^{3/5} \tag{D.20}$$

$$\frac{R}{a} = \left(\frac{b}{a}\right)^{1/4}\left[\frac{n^{3/4}w^{1/4}}{\alpha^{1/2}}\left(\frac{a}{d}\right)^{1/4}\right] \tag{D.21}$$

$$\frac{\rho}{a} = \left(\frac{b}{a}\right)^{5/8}\left[n^{15/8}w^{5/8}\alpha^{1/4}\left(\frac{a}{d}\right)^{17/8}\right] \tag{D.22}$$

$$\frac{R_F(N')}{R_F(N)} = \frac{(a/b)^{1/40}\,(d/a)^{1/8}}{w^{1/40}\alpha^{1/20}n^{3/40}}\left(1 + \frac{(b/a)^{2/5}\,w^{1/5}\alpha^{2/5}n^{8/5}}{(d/a)^{7/5}N} + \cdots\right) \tag{D.23}$$

$$n_\star = \left(\frac{a}{b}\right)^{1/4}\left[\frac{(d/a)^{7/8}}{w^{1/8}\alpha^{1/4}}\left(\frac{N}{\mathrm{k}}\right)^{5/8}\right] \tag{D.24}$$